

Systemy wizyjne w Przemysle 4.0

✎ Wojciech Kaczmarek, Jarosław Panasiuk

1. Wprowadzenie

Systemy wizyjne stają się coraz popularniejszym elementem stanowisk zarówno tych zautomatyzowanych, jak i zrobotyzowanych, pozwalającym z jednej strony podnieść elastyczność rozwiązania, a z drugiej wpłynąć na jego efektywność. Wiąże się to nierozdzielnie ze znacznym uproszczeniem konstrukcji samego stanowiska oraz z uzyskiwaniem dodatkowych informacji wykorzystywanych w procesie produkcyjnym. Systemy wizyjne, a szczególnie najnowsze rozwiązania oparte na elementach uczenia maszynowego i sztucznej inteligencji, stały się wizytówką i nieodłącznym elementem rozwiązań Przemysłu 4.0. Początki rozwiązań z zakresu wizji maszynowej opracowano na przełomie lat 40. i 50. XX w. wraz ze wstępnymi badaniami nad sztuczną inteligencją. Zagadnienie to było szczególnie atrakcyjne z punktu widzenia rozwiązań militarnych, gdzie zamierzano wykorzystać systemy wizyjne i analizę obrazu zarówno w zakresie odnajdywania istotnych informacji na obrazie (znajdowanie zamaskowanych wyrzutni pocisków międzykontynentalnych na bezkresnych obszarach ZSRR), jak i naprowadzania pocisków na cele. Koncepcja systemów wizyjnych została uprzemysłowiona dopiero w latach 60. i 70. XX w., kiedy naukowcy z Massachusetts Institute of Technology opracowali pierwsze wystarczająco wydajne algorytmy analizy obrazu zawierające matematyczny opis tego, co rejestruje oko kamery, a co miało być wykorzystane do sterowania ramieniem robota w zastosowaniach przemysłowych.

W latach 80. XX w. poświęcono wiele wysiłku na sformułowanie skutecznych metod matematycznych do analizy i przetwarzania cyfrowego obrazu w celu zapewnienia możliwości efektywnego wykorzystania informacji z kamery. Dzięki rosnącym możliwościom jednostek obliczeniowych i coraz wydajniejszym algorytmom skracającym czas analizy obrazu możliwe było w końcu zastosowanie widzenia maszynowego na skalę przemysłową. Konieczność uzyskania powtarzalnych wyników pomiarów oraz rosnące wymagania co do jakości produktów sprawiły, że wizja maszynowa stała się nieodzownym elementem wielu linii produkcyjnych.

Lata 90. XX w. przyniosły dalszy szybki rozwój branży wizji maszynowej, czego motorem był szybki postęp technologiczny w dziedzinie komputerów i jednostek obliczeniowych. Pojawiły się pierwsze zintegrowane rozwiązania w postaci czujników wizyjnych oraz inteligentnych kamer, w których zaimplementowano algorytmy umożliwiające lokalizację obiektów w przestrzeni obrazu, pomiar podstawowych cech tych obiektów oraz porównywanie ich z wcześniej zapisanymi wzorcami. Pojawiły się również pierwsze

zrobotyzowane aplikacje wykorzystujące systemy wizyjne. Okazało się, że połączenie to znacznie zwiększa możliwości stosowania robotów przemysłowych, które dotychczas nie były brane pod uwagę przy realizacji niektórych zadań z racji sztywnych algorytmów ich działania. Implementacja standardowego sterowania pozycyjnego zazwyczaj okazywała się niewystarczająca, szczególnie gdy zmiany otoczenia musiały wpływać na pracę robota. Do rozpowszechnienia systemów wizyjnych w robotyce przyczynił się także rozwój przemysłowych standardów komunikacji oraz metod kalibracji systemów wizyjnych, co pozwoliło w prosty sposób łączyć dane pochodzące z układu współrzędnych systemu wizyjnego z układami współrzędnych, w jakich porusza się robot.

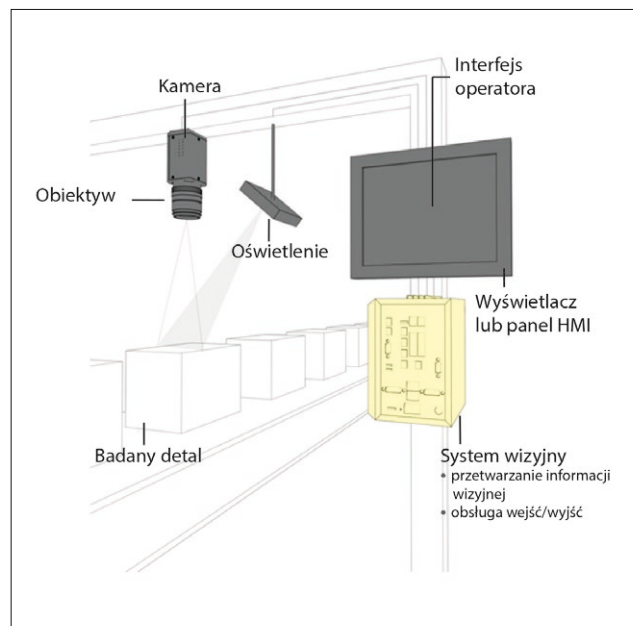
Obecnie główne wysiłki w zakresie rozwoju systemów wizyjnych nakierowane są na wykorzystanie algorytmów uczenia maszynowego oraz widzenie przestrzenne. Tak jak zastosowanie systemów wizyjnych na stanowiskach zrobotyzowanych pozwoliło na dodanie robotom dodatkowych źródeł informacji, tak zastosowanie algorytmów sztucznej inteligencji stanowi kolejny poziom nadania jeszcze większej elastyczności, a przede wszystkim prostoty w przygotowaniu aplikacji wizyjnej. Coraz krótsze serie produkcyjne, częste zmiany profilu działalności firm czy też różnorodność wytwarzanych produktów powodują, że dąży się do tego, aby robot przemysłowy, który zastąpił człowieka na linii produkcyjnej, stał się nie tylko szybki i precyzyjny, ale również choć trochę inteligentny. Inteligencja w tym przypadku wiąże się nierozdzielnie z koniecznością odbierania informacji z otoczenia, a w przypadku stanowiska zrobotyzowanego – informacji istotnych z punktu widzenia realizowanego procesu technologicznego. Jednym z najbardziej istotnych źródeł informacji są w tym przypadku właśnie systemy wizyjne dostarczające informacji o otoczeniu robota w dwóch lub trzech wymiarach.

Mówiąc o systemach wizyjnych, warto zdać sobie sprawę, że w literaturze często można spotkać się z terminami: wizja maszynowa oraz wizja komputerowa. Koncepcja maszyn, które mogą zbierać informację z otoczenia i działać na ich podstawie, nie jest nowa. Na początku automatyzacji przemysłu maszyny wyposażano w pojedyncze czujniki wykrywające określone oddziaływanie fizyczne – nacisk, ciśnienie, temperaturę (czujniki zamieniające wielkość fizyczną na analogowy sygnał elektryczny, a obecnie coraz częściej na cyfrową postać sygnału, co znacząco ułatwia jej dalsze wykorzystanie w cyfrowych systemach sterowania). Od początku jednak właśnie nauczanie maszyn widzenia stanowiło swojego

rodzaju Świętego Grała, którego zdobycie stanowiłoby radykalną zmianę zasad działania tych maszyn. Wynikało to z faktu, że korzystając z jednego zarejestrowanego obrazu, można bezkontaktowo pozyskać tę samą ilość informacji co przy użyciu kilku klasycznych czujników rejestrujących np. położenie czy inne cechy badanego obiektu. Wizja maszynowa (ang. *machine vision* – MV) to termin, którym określa się rozwiązania polegające na wykorzystaniu istniejących technologii i narzędzi, umożliwiające przekazywanie określonych danych lub informacji pochodzących z kamer w taki sposób, aby mogły być użyte do rozwiązywania problemów związanych z postrzeganiem cech obiektów znajdujących się w ich polu widzenia. Jednym z najczęstszych, praktycznych zastosowań systemów wizyjnych maszyn jest kontrola produktu w procesie produkcji oraz pozycjonowanie obiektów w przestrzeni roboczej. Istotą widzenia maszynowego jest przetwarzanie obrazów dostarczonych do komputera z kamer i ich interpretacja na potrzeby sterowania kinematyką robota. Tu pojawia się pojęcie widzenia komputerowego (ang. *computer vision* – CV), czyli rozpoznawania obrazu. Techniki widzenia komputerowego polegają na naśladowaniu ludzkiego systemu widzenia w celu wyodrębnienia przydatnych informacji z cyfrowych obrazów statycznych lub ruchomych, w tym filmów. CV to dziedzina elektroniki i informatyki obejmująca technologie i narzędzia, dzięki którym komputery widzą i interpretują otaczający je świat.

Wizja maszynowa jest wykorzystywana do czterech zasadniczych zadań. Są to:

- **Pozycjonowanie** – zadanie polegające na wykrywaniu i lokalizowaniu obiektów, a następnie zgłaszaniu liczby wykrytych obiektów oraz ich współrzędnych i orientacji każdego z nich. Zadanie to jest szczególnie popularne w przypadku zastosowań zrobotyzowanych podczas operacji pobierz i odłóż (ang. *pick and place*), kiedy robot dzięki zintegrowanemu systemowi wizyjnemu może odnaleźć detal w polu widzenia systemu wizyjnego, a następnie udostępnić te dane, aby robot, z którym zintegrowany jest system wizyjny, mógł pobrać element i odłożyć w miejsce zgodnie z zaprogramowanym algorytmem. Pozycjonowanie może być też wykorzystywane w operacjach depaletyzacji – wówczas stanowisko wykorzystujące taki system jest w stanie w sposób w pełni automatyczny przeprowadzić depaletyzację towaru.
- **Inspekcja** – zadanie polegające na sprawdzeniu jakości produktu, m.in. sprawdzeniu obecności wszystkich części zespołu lub znalezieniu wady bądź odchylenia parametrów od zadanych wartości. Zadanie to może być realizowane zarówno na stanowiskach zrobotyzowanych, jak i niezrobotyzowanych. W przypadku stanowisk zrobotyzowanych najczęściej procesy inspekcyjne wykorzystuje się na stanowiskach montażowych oraz obsługi maszyn. Przykładem stanowisk niezrobotyzowanych są wszelkiego rodzaju zastosowania związane z szybkim przepływem towarów jak przy produkcji żywności o długim okresie przydatności czy też napojów (sprawdzenie poprawności nadruków, stanu opakowania, poprawności zamknięcia opakowania itp.).



↑ Rys. 1. Podstawowe komponenty przemysłowego systemu wizyjnego [164]

- **Pomiar** – zadanie polegające na określeniu wymiarów geometrycznych obiektu, takich jak długość, szerokość, wysokość, powierzchnia i objętość. Zadanie to jest często elementem zadań wchodzących w skład inspekcji wizyjnej. Pomiarów są realizowane zarówno przez najprostsze czujniki wizyjne, jak i przez kamery inteligentne czy bardziej zaawansowane systemy wizyjne. W tym przypadku istotny jest zakres dokładności i jednostki, w jakich określane są wyniki pomiarów. W przypadku większości czujników wyniki pomiarów są podawane w pikselach, a w bardziej zaawansowanych systemach istnieje możliwość wybrania jednostek.
- **Odczyt** – zadanie dające możliwość dekodowania i rozpoznawania obrazów oraz tekstów, takich jak kod 1D, kod 2D i OCV/OCR. Zadania te często mogą stanowić wstęp do innych zadań procesów wizyjnych – na podstawie odczytanego kodu system wizyjny jest w stanie np. przestroić się na inspekcję obiektu opisanego określonym kodem. Zadanie czytania jest jednak najbardziej rozpowszechnione w zastosowaniach związanych z logistyką, gdzie kodom są przyporządkowane całe zbiory informacji mówiące o życiu produktu. O istotności tego zadania może świadczyć również fakt, że główni producenci rozwiązań z zakresu systemów wizyjnych mają dedykowane linie produktów do zadań związanych z odczytem kodów 1D i 2D.

W wielu przypadkach system wizyjny dla określonej aplikacji przemysłowej składa się z kombinacji kilku takich zadań, gdzie często w zależności od rozpoznanego detalu, realizowane mogą być różne zestawy zadań.

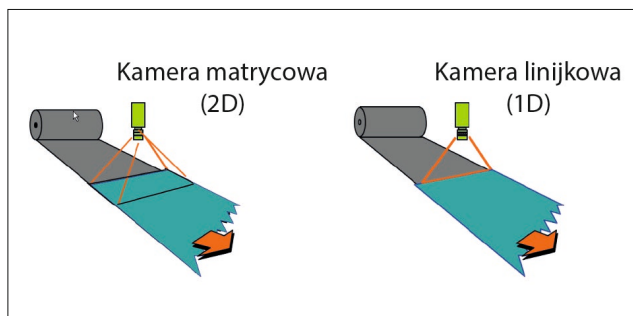
W skład praktycznie każdego systemu wizyjnego (rys. 1) wchodzi urządzenie (kamera lub zestaw kamer w przypadku systemów stereowizyjnych lub systemów realizujących operacje będące złożeniem kilku obrazów z kamer) służące do

rejestracji obrazu (natężenia oświetlenia pochodzącego ze sceny wizyjnej) oraz jego przetworzenia na sygnał cyfrowy, jak również urządzenia przeznaczone do dalszego przetwarzania i analizy informacji cyfrowej, zgodnie z opracowanym i uruchomionym algorytmem.

Aby system wizyjny mógł pracować wydajnie, konieczne jest zapewnienie odpowiednich warunków oświetlenia. Zmiana parametrów oświetlenia może bowiem prowadzić do sytuacji, kiedy system wizyjny po prostu nie będzie w stanie rozpoznać elementów na scenie wizyjnej: kiedy jest za jasno, system jest niejako oślepiony, a gdy jest za ciemno, nie może niczego dostrzec w mroku. Człowiekowi rozszerzająca się lub zwężająca źrenica umożliwia dostosowanie ilości światła padającego na siatkówkę oka, czyli dostosowanie do zmiennych warunków ekspozycji w otaczającym świecie. W przypadku systemów wizyjnych elementem odpowiadającym źrenicy jest przesłona obiektywu. W wielu systemach stopień otwarcia przesłony jest regulowany automatycznie. W innych zmienne warunki oświetlenia są kompensowane zmianą czasu ekspozycji. Niestety każda zmiana tych parametrów zajmuje pewien czas (czas trwania procesu regulacji), a zakres zmiany tych parametrów jest ograniczony.

2. Systemy wizyjne 1D, 2D, 2,5D i 3D

Najczęściej spotykane są systemy wizyjne 2D i 3D (obrazowanie w dwóch wymiarach lub trzech wymiarach – przestrzenne). W ostatnich latach szczególnie systemy 3D stały się bardzo popularne ze względu na pojawienie się wielu nowych technologii implementujących obrazowanie przestrzenne. Znaczenie poszczególnych kategorii systemów w przypadku stanowisk zrobotyzowanych jest jednak nieco inne. System wizyjny 2D jest rozwiązaniem najczęściej spotykanym. Polega on na zastosowaniu jednej kamery lub zestawu kamer, jeśli wymagane jest uzyskanie obrazu z większej powierzchni podczas realizacji procesu wizyjnego. Większość rozwiązań dostępnych na rynku wykorzystuje standardowe kamery 2D przetwarzające obraz rzeczywistego świata na płaski obraz.



↑ Rys. 2. Sposób przetwarzania obrazu w kamerze matrycowej i liniowej [165]

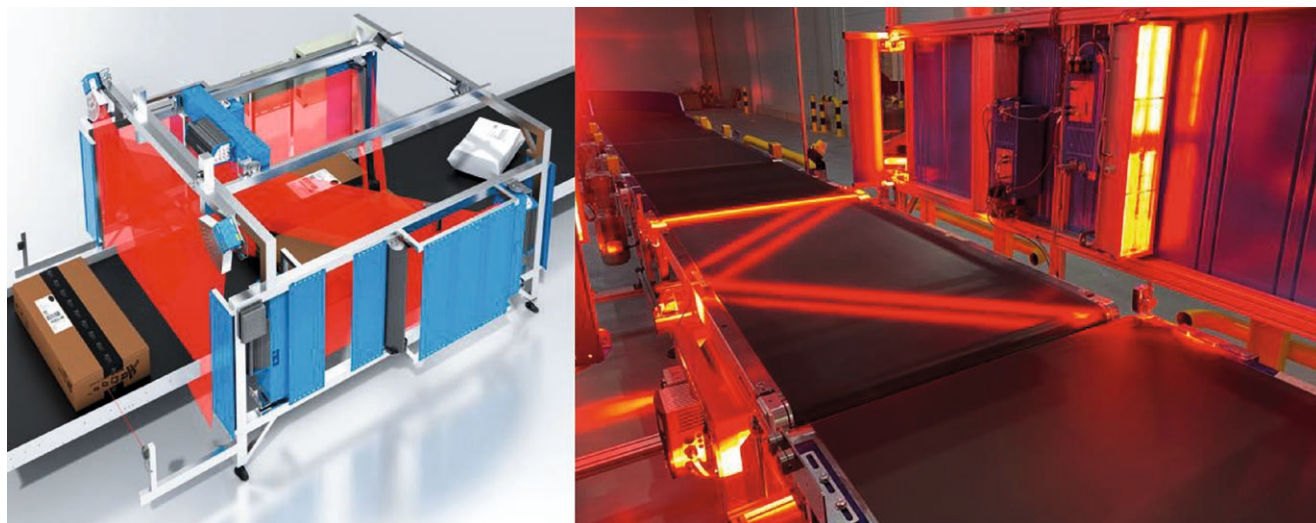
Takie rozwiązanie jest z jednej strony dość wygodne, lecz z drugiej strony ma dość znaczące ograniczenia. Nie zapewnia ono bowiem pełnej informacji o położeniu i orientacji obiektu w przestrzeni, co zapewniają dopiero systemy 3D. Warto również zauważyć, że choć może to się wydawać nie naturalne, kiedy mówi się o systemach wizyjnych, są jeszcze systemy oparte na kamerach liniowych, często w literaturze nazywane systemami 1D.

Czym są więc poszczególne technologie widzenia maszynowego i jakie są ich główne cechy?

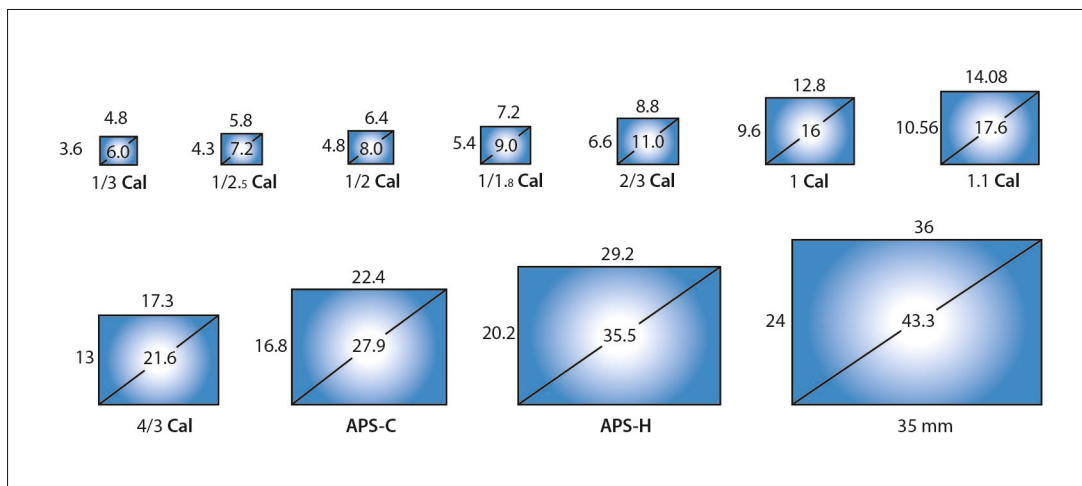
2.1. Technologia wykorzystująca kamery liniowe (1D)

Rozwiązanie 1D polega na analizowaniu sygnału rejestrowanego przez kamerę, której matryca ma postać prostokąta o krótszym boku równym najczęściej jeden do trzech pikseli (w literaturze możemy spotkać się ze stosowanymi zamiennie nazwami kamera liniowa jak również kamera liniowa). Obraz, na którym mogą być przeprowadzane operacje analizy obrazu, jest składany z kolejnych linijek obrazu (rys. 2).

Tego rodzaju systemy są z powodzeniem wykorzystywane do ciągłej kontroli przy produkcji takich materiałów jak papier, tworzywa sztuczne, metale i inne włókna w arkuszach lub rolkach. Technologia wykorzystująca kamery



↑ Rys. 3. Przykład rozwiązania firmy SICK wykorzystującego kamery liniowe do skanowania kodów paczek [166]



← Rys. 4. Wielkości matryc stosowanych w standardowych kamerach wykorzystywanych w systemach wizyjnych [167]

linijkowe jest powszechnie używana do wykrywania i klasyfikacji defektów materiałów wytwarzanych w procesie ciągłym. Głównymi zadaniami, dla których warto zastanowić się nad zastosowaniem kamer linijkowych, są:

- kontrola obiektów o kształtach cylindrycznych;
- inspekcja w ograniczonej przestrzeni – tam, gdzie nie ma możliwości uzyskania pełnego obrazu kontrolowanego obiektu;
- kontrola parametrów obiektów przy zachowaniu wysokiej rozdzielczości;
- kontrola obiektów i towarów w ruchu ciągłym.

W przypadku technologii 1D proces skanowania detalu może być realizowany albo poprzez ruch detalu (kolejne linie przemieszczającego się detalu są rejestrowane przez linijkę detektorów matrycy), albo poprzez ruch zwierciadła skanującego linijką detektorów powierzchnię obiektu (detal jest nieruchomy, a przemieszczany jest jedynie obraz fragmentu detalu rzutowany na linijkę detektorów). Na rysunku 3 przedstawiono system śledzenia oparty na kamerze linijkowej ICR89x firmy SICK. System wizyjny wykorzystuje sześć kamer z przetwornikami liniowymi. Cztery z nich są umieszczone po bokach podajnika (po dwie na stronę), jedna znajduje się nad podajnikiem, a kolejna w szczelinie pomiędzy podajnikami. Takie rozwiązanie pozwala na zeskanowanie każdej z sześciu stron kartonu, a następnie na odczyt kodu na zarejestrowanym obrazie. Idealnie dla wszystkich zaawansowanych aplikacji do odczytu kodów liniowych i 2D w procesach transportowych i logistycznych. Wyjątkowa jakość obrazu zintegrowanej kamery ICR89x sprawia, że nadaje się ona do zastosowania w OCR, kodowaniu wideo i aplikacjach. Rozwiązanie jest idealne dla wszystkich zaawansowanych aplikacji do odczytu kodów liniowych i dwuwymiarowych w procesach transportowych i logistycznych. Wyjątkowa jakość obrazu zintegrowanej kamery ICR89x sprawia, że nadaje się ona do zastosowania w rozpoznawaniu tekstu, kodowaniu wideo i aplikacjach wizyjnych.

2.2. Technologia kamer 2D

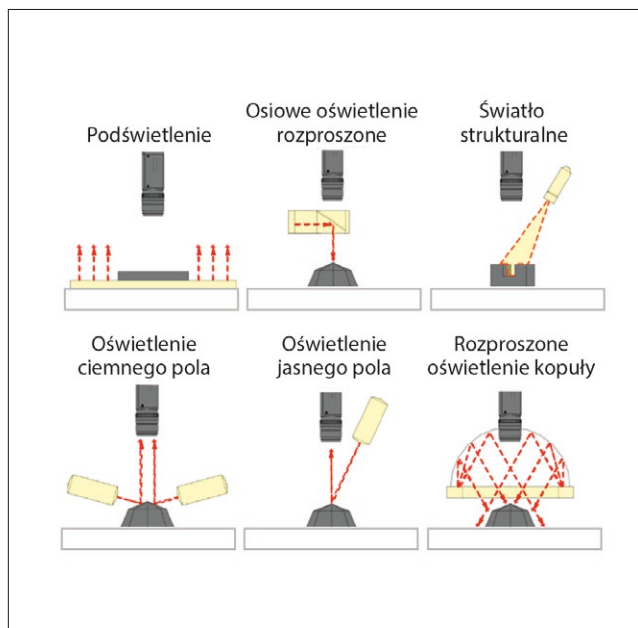
Najpopularniejszym rodzajem technologii wizyjnej jest obrazowanie 2D. Jest ona szczególnie przydatna w przypadku

detekcji i śledzenia obiektów o dużym kontraście, określonej teksturze lub kolorze. Z wykorzystaniem wizji 2D można rozwiązać praktycznie wszystkie cztery główne zadania wizyjne, choć trzeba tutaj zwrócić uwagę, że jeśli chodzi o operację pozycjonowania na potrzeby stanowisk zrobotyzowanych, to nie dostarcza ona pełnych informacji o orientacji detalu i nadaje się bardziej do pozycjonowania na zadanej płaszczyźnie.

W systemach wizyjnych 2D do akwizycji obrazu obiektu wykorzystuje się standardowe kamery analogowe, a coraz częściej cyfrowe. W przypadku widzenia maszynowego 2D rejestrowana jest (i przetwarzana zależnie od przyjętego algorytmu) dwuwymiarowa mapa ($x-y$) natężenia odbitej wiązki światła. Należy pamiętać, że obraz 2D nie dostarcza bezpośrednich informacji w trzecim wymiarze. Przetwarzanie zazwyczaj polega na porównywaniu zmian natężenia oświetlenia (kontrastu). Systemy wizyjne 2D są stosowane w branży automatyki przemysłowej w szerokim zakresie zadań, w tym w zadaniach weryfikacji cech i położenia, sprawdzaniu wymiarów, odczytywaniu kodów kreskowych, rozpoznawaniu znaków, weryfikacji etykiet i kontroli jakości. Typowymi ograniczeniami technologii widzenia maszynowego 2D są: błędy paralaksy, mała głębia ostrości, wpływ światła otoczenia i zmian kontrastu. Często ograniczenia te mogą być eliminowane za pomocą odpowiednich algorytmów obliczeniowych, choć zazwyczaj wiąże się to z koniecznością spowolnienia procesu wizyjnego.

W obrazowaniu 2D scena wizyjna jest rejestrowana w czasie rzeczywistym przez kamerę matrycową. Wynikową reprezentacją sceny jest obraz monochromatyczny (oddający zmiany natężenia oświetlenia w skali szarości) lub obraz kolorowy (często wartości RGB). Parametrem mającym wpływ na czułość przetwornika i poziom szumu jest wielkość matrycy (rys. 4). Kamery z większym przetwornikiem zazwyczaj charakteryzują się większą czułością i mniejszym poziomem szumu na rejestrowanym obrazie, co ma bezpośredni wpływ na jakość uzyskiwanych obrazów.

Kluczowe dla akwizycji obrazu 2D o jakości odpowiedniej do zadania są, poza matrycą, wybór obiektywu (o odpowiedniej ogniskowej i przesłonie) oraz oświetlenia. Wybór



↑ Rys. 5. Najpopularniejsze rozwiązania oświetlenia stosowane w systemach wizyjnych 2D [168]

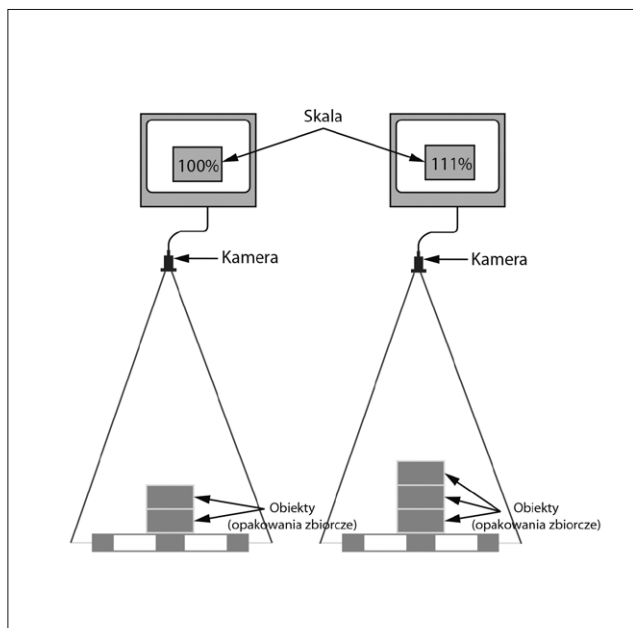
obiektywu wiąże się z takimi parametrami optycznymi jak jasność oraz ogniskowa. Dobierając układ optyczny do kamery systemu wizyjnego, należy mieć na uwadze zastosowanie, czyli pole widzenia oraz odległość kamery od rejestrowanej powierzchni. Od parametrów oświetlenia w dużej mierze zależy jakość obrazu rejestrowanego przez system wizyjny. Prawidłowe oświetlenie (rys. 5) pozwala na uwydatnienie cech istotnych obiektu poddawanego analizie, zapewnia wysoką jakość obrazu i zapewnia stabilne warunki pracy, niezależnie od światła otoczenia, które często potrafi zmieniać się w cyklu dobowym. W przypadku nieodpowiedniego oświetlenia sceny wizyjnej uzyskany obraz będzie wymagał skorygowania przez wydłużenie czasu ekspozycji bądź zastosowanie odpowiednich narzędzi programowych, co może nie dać takich wyników jak dobrze oświetlona scena. Przy odpowiednio dobranym oświetleniu i ustabilizowaniu warunków często można znacząco uprościć i przyspieszyć wdrożenie systemu wizyjnego.

Systemy wizyjne 2D są głównie wykorzystywane do:

- identyfikacji części, rozpoznawania położenia;
- pomiaru wielkości geometrycznych;
- kierowania robotem w płaszczyźnie;
- programowania i/lub korekcji centralnego punktu narzędzia;
- operacji pick and place;
- składania (zestawianie) części;
- rozpoznawania wad na płaszczyznach (powierzchniach);
- rozpoznawania pisma (OCR) i kodów.

2.3. Technologia widzenia 2,5D

Termin widzenia 2,5D wiąże się z dodaniem trzeciego wymiaru, a więc wysokości elementu, którego obraz jest rejestrowany klasyczną kamerą 2D. Korzystając ze standardowej



↑ Rys. 6. Idea pracy systemu wizyjnego 2,5D wykorzystywanego w procesie depaletyzacji [169]

kamery i systemu wizyjnego, jak to ma miejsce w przypadku standardowych systemów wizyjnych, nie ma możliwości określenia wysokości obiektów. Można ją jednak obliczyć w sposób pośredni, korzystając z faktu, że obiekt znajdujący się bliżej kamery daje większy obraz, a im obiekt jest dalej, tym jego wymiar na obrazie jest mniejszy.

Aby móc jednak obliczyć wysokość, na której znajdują się kolejne warstwy opakowań (najczęściej są to kartony zbiorcze), konieczna jest znajomość wysokości opakowania podawanego rozpoznaniu. Podczas konfiguracji aplikacji należy zarejestrować obraz dwóch kolejnych warstw i wprowadzić wysokość warstwy (wysokość opakowania). Następnie algorytm automatycznie, znając parametry systemu optycznego, przelicza procentową zmianę skali obrazu i na tej podstawie interpretuje wysokość warstwy, na której znajduje się wykryte opakowanie. W przypadku zbiorczych opakowań kartonowych stosunek wysokości do pozostałych wymiarów jest na tyle duży, że system wizyjny jest w stanie policzyć i zinterpretować wysokość, na jakiej znajduje się dane opakowanie (rys. 6).

Najważniejszą korzyścią implementacji rozwiązania 2,5D jest brak konieczności stosowania dodatkowych czujników w celu wykrywania poziomu warstwy, z której mają być pobierane elementy. Dobry stosunek jakości do ceny spowodował, że systemy 2,5D stały się powszechne w operacjach paletyzacji i depaletyzacji.

2.4. Technologia widzenia 3D

Najbardziej zaawansowanym rozwiązaniem z zakresu technologii systemu wizyjnego jest system 3D, umożliwiający zarówno rozpoznawanie cech i parametrów przestrzennych, jak i położenia (X , Y , Z) oraz orientacji (W , P , R) detali. W takim przypadku możliwe jest wykorzystanie robota nie

tylko do pobierania elementów umieszczonych na dwuwymiarowej powierzchni podajnika, ale również wtedy, gdy ich położenie jest losowe oraz gdy znajdują się one w pojemniku lub kontenerze.

Systemy wizyjne 3D z reguły są konstrukcjami dość złożonymi w stosunku do systemów 2D. Zwykle składają się z wielu kamer lub laserowych czujników przemieszczenia zamontowanych w różnych miejscach, umożliwiając triangulację obiektu w przestrzeni 3D. Systemy wizyjne 3D mogą być wykorzystywane do bardzo wielu zadań, w tym do pomiaru grubości i wysokości, obliczania objętości, weryfikacji kształtów, pomiaru parametrów otworów, kątów i krzywych, wymiarowania i zarządzania przestrzenią, wykrywania wad powierzchniowych lub wad montażu, kontroli jakości, operacji typu pick and place, pakowania lub montażu, skanowania i digitalizacji obiektów jak również prowadzenia robota i śledzenia powierzchni (np. podczas operacji spawania, klejenia, gratowania).

Techniki obrazowania 3D mogą być realizowane na wiele różnych sposobów. Każda z nich ma swoje zalety i wady, które mogą ją predestynować do określonych zastosowań bądź ją z nich wykluczać. W ogólności techniki obrazowania 3D można podzielić na dwie zasadnicze kategorie:

- metody skanowania;
- metody migawkowe.

W technice wykorzystującej proces skanowania obrazy 3D są generowane profil po profilu albo przez przesuwanie obiektu w obszarze pomiarowym, albo przez przesuwanie kamery nad obiektem. Warunkiem uzyskania poprawnego obrazu 3D jest zapewnienie pełnej synchronizacji przemieszczenia z rejestracją kolejnych profili obrazu obiektu. Może być to realizowane bądź przy użyciu enkodera rejestrującego liniowy ruch obiektu, bądź przez rejestrowanie ruchu robota przemieszczającego obiekt w przestrzeni skanowania. Uzyskiwane w ten sposób obrazy zazwyczaj charakteryzują się dużą dokładnością. Przykładem tego typu metody obrazowania 3D są systemy wykorzystujące triangulację laserową. W przypadku technik migawkowych system wykonuje często jedno lub kilka ujęć z jednej lub kilku kamer. W zależności od zastosowanej techniki migawkowej widzenia 3D możemy wyróżnić metody oparte na stereowizji, na wykorzystaniu światła strukturalnego oraz metody polegające na pomiarze czasu przelotu sygnału świetlnego (ang. *Time-of-Flight* – ToF).

2.5. Techniki obrazowania 3D wykorzystujące triangulację laserową

W triangulacji laserowej do obrazowania 3D jest używany laser i kamera służąca do rejestracji linii laserowej rozkładającej się w poprzek na powierzchni skanowanego obiektu. Uzyskane w ten sposób profile są następnie łączone w specjalistycznym oprogramowaniu, aby stworzyć obraz 3D. Przykładem może tu być kamera TriSpectorP1000 firmy SICK (rys. 7). Ponieważ pozyskiwanie profilu wysokości wymaga ruchu obiektu lub kamery, metoda jest określana jako technologia skanowania. Triangulacja laserowa charakteryzuje



↑ Rys. 7. System wizyjny TriSpectorP1000 firmy SICK wykorzystujący metodę triangulacyjną [170]

się większą dokładnością pomiaru niż ma to miejsce w przypadku takich technologii jak ToF, ma jednak bardziej ograniczony zakres pomiarowy, co w wielu przypadkach może stanowić istotne ograniczenie jej zastosowania.

2.6. Techniki obrazowania 3D wykorzystujące stereowizję

W stereowizji dane obrazu pozyskiwane z dwóch kamer są przetwarzane za pomocą specjalnego algorytmu, w wyniku czego uzyskuje się przestrzenny obraz każdego punktu w przestrzeni. W ten sposób, znając wzajemne położenie kamer, można odtworzyć głębię obrazu. Technika ta jest odzwierciedleniem ludzkiego procesu widzenia, co ma szereg zalet, takich jak możliwość osiągnięcia wysokiej precyzji dla małych odległości czy możliwość wykorzystania dwóch kamer 2D zamiast kamery 3D. Metoda ta wymaga jednak skomplikowanych algorytmów, dużej mocy obliczeniowej oraz precyzyjnej metody pobierania obrazów. Sprawia to, że stereowizja jest droga, złożona w obsłudze i trudna w zastosowaniach do aplikacji wymagających pracy w czasie rzeczywistym.

W odróżnieniu od triangulacji laserowej i ToF w metodzie stereowizji nie używa się zazwyczaj dedykowanego źródła światła. Jednak, aby znaleźć korelacje, dwa obrazy muszą mieć wystarczającą liczbę szczegółów, a obiekty odpowiednią teksturę lub niejednorodność. Z tego powodu nadaje się ona do zastosowań przy dużym polu widzenia i do zastosowań na zewnątrz.

Aby system mógł działać poprawnie, konieczne jest wykonanie kalibracji. Polega ona na ustawieniu wewnętrznych i zewnętrznych parametrów kamery. Wewnętrzne parametry określają geometryczne, cyfrowe i optyczne cechy kamery, takie jak ogniskowa kamery, współrzędne środka obrazu i rozmiar piksela matrycy kamery. Zewnętrzne parametry to względna pozycja kamery i orientacja kamer względem siebie, a także parametr stanowiący o odległości między środkami rzutowania.

Proces akwizycji i obróbki danych rozpoczyna się od korekcji zniekształceń geometrycznych wynikających z optyki układu oraz przekształcenia obrazów z niekanonicznego w kanoniczny układ kamer (osie optyczne kamer są

równoległe oraz współrzędne z ognisk kamery są takie same). Procesy te są określane mianem rektyfikacji. Jest to możliwe dzięki wstępnej kalibracji układu optycznego. W kolejnym etapie na obu obrazach są wyszukiwane odpowiadające sobie piksele i wyznaczana jest wartość dysparycji (odległość między obrazami tego samego punktu na płaszczyznach dwóch lub więcej kamer) dla każdej stereopary (obrazy tego samego punktu widzianego z różnych perspektyw). Umożliwia to wykonanie mapy dysparycji.

Aby możliwe było dopasowanie obrazów stereo, przyjmuje się założenia upraszczające:

- jeden z obrazów przyjmuje się jako obraz odniesienia;
- pikselowi z pierwszego obrazu przyporządkowuje się dokładnie jeden piksel z drugiego obrazu;
- zakłada się symetryczność obrazów (zamiana obrazu lewego z prawym nie wpływa na przyporządkowanie);
- jasność pikseli na jednym obrazie jest zbliżona do jasności na drugim obrazie;
- wolnozmiennosc dysparycji w całym obszarze obrazu.

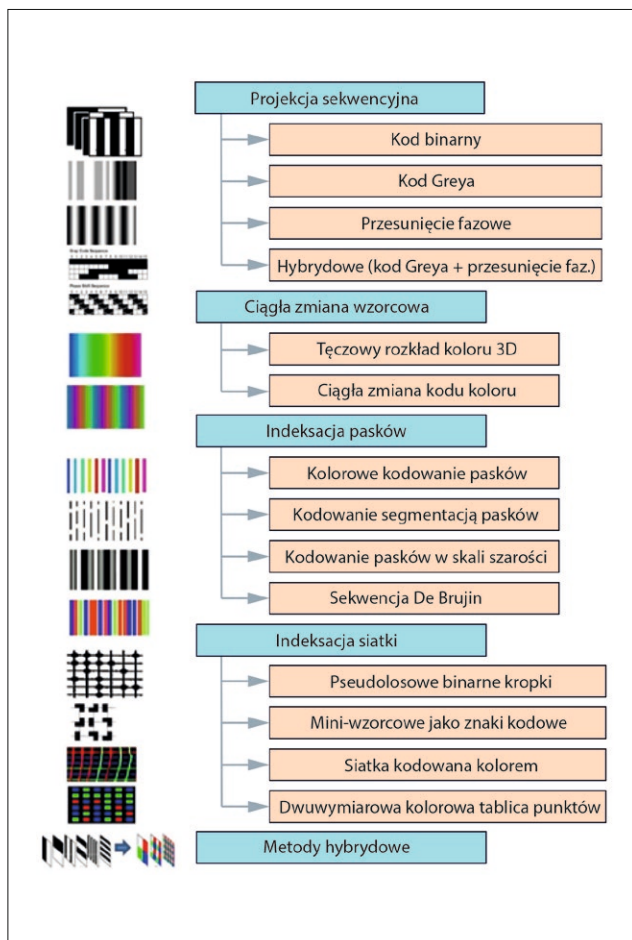
Na podstawie danych z kalibracji oraz mapy dysparycji, wykorzystując zjawisko triangulacji, tworzy się mapę głębi. Dzięki temu znane są nie tylko współrzędne x i y analizowanego obiektu, ale także jego odległość od kamery. Aby uzyskać lepsze wyniki, możliwe jest podniesienie szczegółowości poprzez oświetlenie sceny światłem strukturalnym. Stereowizja znajduje bardzo szerokie zastosowania, poczynając od aplikacji zrobotyzowanych, poprzez systemy rzeczywistości wirtualnej, aż po diagnostykę medyczną.

2.7. Techniki obrazowania 3D wykorzystujące światło strukturalne

Obecnie dostępnych jest wiele migawkowych technik obrazowania powierzchni za pomocą światła strukturalnego (rys. 8). W zależności od tego, ile obrazów jest wykorzystywanych i jaki rodzaj światła strukturalnego jest użyty, można te rozwiązania podzielić na sekwencyjne (wielokrotne) lub jednokrotne. Jeśli docelowy obiekt 3D podlegający analizie jest statyczny, a zastosowanie nie nakłada surowych ograniczeń na czas akwizycji, można zastosować techniki wielokrotne (ang. *sequential projections*), uzyskując dokładniejsze i bardziej wiarygodne wyniki. Jeśli obiekt się porusza, konieczne jest zastosowanie metod wykorzystujących jeden obraz – migawkowy obraz powierzchni 3D obiektu w określonym czasie. Techniki wykorzystujące jeden obraz możemy podzielić na trzy szerokie kategorie w zależności od rodzaju wykorzystanego światła strukturalnego rzutowanego na obiekt. Są to metody wykorzystujące:

- stale zmieniające się wzorce światła strukturalnego (ang. *continuous varying pattern*);
- schematy kodowania 1D (ang. *stripe indexing*);
- schematy kodowania 2D (ang. *grid indexing*).

Każda z nich ma swoje zalety i wady, w zależności od konkretnych zastosowań. Istnieje również możliwość łączenia różnych technik w celu osiągnięcia zamierzonych korzyści, co musi być jednak zweryfikowane już do konkretnego zastosowania.



↑ Rys. 8. Klasyfikacja technik obrazowania 3D wykorzystujących światło strukturalne [171]

2.8. Techniki obrazowania 3D wykorzystujące ToF

Kamery 3D Time-of-flight (ToF) tworzą obrazy 3D w technice migawki. Oznacza to, że nie jest potrzebny ruch obiektu ani ruch kamery. Metoda polega na pomiarze czasu przelotu sygnału świetlnego pomiędzy źródłem światła, stanowiącym element systemu, powierzchnią, na którą to światło pada i od której się odbija, a kamerą rejestrującą obraz. Pomiar realizowany jest dla każdego punktu obrazu. Znając przesunięcie fazowe czasu nadejścia sygnału w stosunku do sygnału początkowego, można wyznaczyć odległość między urządzeniem a obiektem i sceną. W wyniku uzyskuje się natychmiast (niemal w czasie rzeczywistym) obraz 3D sceny. Metoda z pomiarem czasu przelotu jest odpowiednia do zastosowań z dużym polem widzenia i odległością roboczą powyżej 0,5 m.

Pierwsze wprowadzone na rynek kamery ToF opierały się na pomiarze głębokości przy użyciu techniki znanej jako modulacja fali ciągłej (CW), która polega na obliczaniu różnic fazowych między światłem emitowanym ze źródła światła a światłem uchwyconym przez matrycę kamery. Najnowsze rozwiązania wykorzystują tak zwaną technikę modulacji impulsowej. W rozwiązaniu tym kamera uruchamia jednostkę oświetlającą, która emituje impulsy światła o dużym natężeniu skierowane na obiekty znajdujące się na scenie. Impulsy odbite od obiektów znajdujących się bliżej wracają do czujnika czasu przelotu w aparacie przed obiektami

znajdującymi się dalej. Na podstawie kilku kolejnych przechwyconych klatek możliwe jest wyodrębnienie głębi obrazu ze sceny. Liczba klatek na sekundę rejestrowanych przez kamerę odpowiada docelowemu zakresowi pracy kamery.

Teoretycznie pulsacyjna kamera ToF może wykorzystywać do oświetlenia sceny światło o dowolnej długości fali. Jednak rozdzielczość głębi obrazu kamer wykorzystujących bliską podczerwień jest czterokrotnie większa niż ich odpowiedników wykorzystujących światło widzialne. Światło podczerwone, jako niewidoczne dla ludzkiego oka, jest również mniej inwazyjne niż światło widzialne, co skutkuje mniejszym zmęczeniem operatora. Ważne jest też to, że kamery wykorzystujące światło podczerwone są odporne na natężenie oświetlenia otoczenia i mogą nawet działać w ciemnym otoczeniu. Kamery ToF mają przewagę nad pasywnymi i aktywnymi systemami triangulacji, ponieważ nie wymagają, aby system wizyjny określał, które punkty obrazu uchwyconego przez jedną kamerę odpowiadają tym samym punktom w innej kamerze. Nie ma również potrzeby mierzenia parametrów zewnętrznych, takich jak względna pozycja i orientacja dwóch kamer, ponieważ używana jest tylko jedna kamera. Kolejną zaletą rozwiązania ToF jest fakt, iż nie ma potrzeby integracji dodatkowego zewnętrznego źródła światła w systemie, ponieważ obiekty są automatycznie oświetlane przez źródło światła, które jest zsynchronizowane z kamerą.

Niestety dostępne na rynku systemy mają również istotne ograniczenia. Kamery o modulacji fali ciągłej charakteryzowały się małymi czujnikami o bardzo małej rozdzielczości w porównaniu z kamerami częściej używanymi w przemyśle widzenia maszynowego, co czyni je nieodpowiednimi do użytku w wielu zastosowaniach przemysłowych. W przypadku kamer z modulacją impulsową rozdzielczość głębi obrazu, którą oferują przy szybkości 30 klatek na sekundę, nie przekracza ± 1 cm w odległości roboczej od 0,5 do 5 m. Kamery ToF nie są również rozwiązaniem idealnym szczególnie w przypadku zastosowań przemysłowych, gdzie ważna jest powtarzalność i odporność na warunki zewnętrzne. Zdecydowanie największą wadą wynikającą z tego, że do obrazowania wykorzystywane są fale świetlne powracające do czujnika, jest to, że każde inne światło obecne w przestrzeni roboczej kamery będzie negatywnie wpływać na dokładność pomiaru.

Dowodem świadczącym jednak o skuteczności i łatwości zastosowania technologii ToF jest to, że urządzenie z niej korzystające sprzedawało się w milionach egzemplarzy i do dzisiejszego dnia jest używane. Urządzeniem tym jest kamera Kinect do Xboxa 360 oraz jego nowsza wersja Kinect 2.0, a obecnie technologia zaczyna pojawiać się także w smartfonach.

Analizując rozwiązania z zakresu systemów wizyjnych 3D, warto mieć na uwadze czynniki, które charakteryzują

↓ **Tabela 1.** Porównanie systemów wizyjnych 3D

System	Cechy	Przykład
Triangulacja laserowa (kamera + projektor laserowy)	<ul style="list-style-type: none"> • źródło światła: laser do projekcji liniowej • nie ma potrzeby oświetlenia otoczenia • wysoka rozdzielczość i dokładność szczegółów • stosunkowo krótki zakres pomiarowy • okluzja możliwa, gdy kamera nie widzi lasera, gdy jest ukryta za obiektem • technologia skanowania 	TriSpectorP1000 firmy SICK
Triangulacja laserowa na potrzebę określenia orientacji obiektu (kamera + promiennik laserowy)	<ul style="list-style-type: none"> • źródło światła: laser do projekcji krzyżowej + pasywne/aktywne światło otoczenia • wysoka rozdzielczość i dokładność szczegółów • stosunkowo krótki zakres pomiarowy • technologia migawki 	Fanuc iRVision 3DL Vision System
Stereowizja (dwie kamery)	<ul style="list-style-type: none"> • źródło światła: pasywne/aktywne światło otoczenia • stosunkowo niska rozdzielczość i dokładność szczegółów • duży zakres pomiarowy • nadaje się do zastosowań zewnętrznych jak i wewnętrznych • technologia migawki 	Fanuc iRVision 3D Area Sensor Vision System
Technologia wykorzystująca światło strukturalne	<ul style="list-style-type: none"> • źródło światła: projektor światła strukturalnego • nie ma potrzeby oświetlenia otoczenia • duży zakres pomiarowy • stosunkowo duża rozdzielczość • technologia migawki 	
Time-of-flight (TOF)	<ul style="list-style-type: none"> • źródło światła: modulowane czasowo • nie ma potrzeby oświetlenia otoczenia • duży zakres pomiarowy • stosunkowo mała rozdzielczość i dokładność szczegółów • technologia migawki 	Kinect

wydajność techniczną systemu obrazowania powierzchni 3D. Z punktu widzenia zastosowania najistotniejszymi wskaźnikami efektywności, które należy wykorzystać do oceny przydatności systemów obrazowania 3D, są:

- **dokładność** – dokładność pomiaru oznacza maksymalne odchylenie zmierzonej wartości uzyskanej przez system obrazowania powierzchni 3D od rzeczywistego wymiaru obiektu 3D; dość często system obrazowania 3D może mieć różne dokładności w różnych (x , y , z) kierunkach ze względu na nieodłączne właściwości projektowe systemów;
- **rozdzielczość** – rozdzielczość optyczna jest definiowana jako zdolność systemu optycznego do rozróżniania poszczególnych punktów lub linii na obrazie. Podobnie rozdzielczość obrazu 3D oznacza najmniejszą część powierzchni obiektu, którą może rozróżnić system obrazowania 3D. Jednak w systemach obrazowania 3D termin rozdzielczość obrazu czasami oznacza również maksymalną liczbę punktów pomiarowych, które system jest w stanie uzyskać w pojedynczej klatce;
- **szybkość akwizycji** – cecha ważna przy obrazowaniu 3D szczególnie w przypadku poruszających się obiektów. W przypadku systemów obrazowania 3D wykorzystujących pojedynczy obraz liczba klatek na sekundę oznacza ich zdolność do generowania pełnej klatki w krótkim odstępie czasu. W przypadku sekwencyjnych systemów obrazowania 3D (np. systemów skanowania laserowego) oprócz szybkości pozyskiwania klatek obrazu należy wziąć pod uwagę jeszcze jedną kwestię: czy obiekt porusza się podczas

wykonywania sekwencyjnej akwizycji. W związku z tym uzyskany pełnoklatkowy obraz 3D może nie przedstawiać migawki obiektu 3D w jednym miejscu, ale stanowić integrację punktów pomiarowych uzyskanych w kolejnych chwilach czasowych, dlatego wynikowy kształt 3D może być zniekształcony w stosunku do oryginalnego kształtu obiektu 3D. Istnieje jeszcze jedna różnica między szybkością akwizycji a szybkością obliczeń. Na przykład, niektóre systemy są w stanie pozyskiwać obrazy 3D z szybkością 30 klatek na sekundę, ale pozyskane obrazy są poddawane obróbce końcowej ze znacznie mniejszą szybkością. Zatem czas generowania danych 3D również musi być uwzględniony podczas implementacji takiego systemu w konkretnym zastosowaniu.

2.9. Podsumowanie

Przedstawiając technologie wykorzystywane w systemach wizyjnych 3D, należy podkreślić, że wybór konkretnego rozwiązania zależy od tego, która technologia będzie najbardziej odpowiednia ze względu daną aplikację i jej środowisko. W tabeli 1 podsumowano kluczowe cechy wspomnianych rozwiązań technologii 3D. □

Fragment pochodzi z książki: *Robotyzacja i automatyzacja* dr inż. Wojciech Kaczmarek, Jarosław Panasiuk
Wydawnictwo Naukowe PWN SA

Reklama



Wspieramy naukę i rozwój!

napędy i sterowanie
miesięcznik naukowo-techniczny

Jesteś studentem, wykładowcą lub doktorantem?
Nawiąż współpracę z naszą redakcją i zaprezentuj wyniki swoich badań w miesięczniku „Napędy i Sterowanie”.

Napisz do nas i zaproponuj temat: redakcja.nis@industrypublisher.com