

Rozpoznawanie obrazów w systemie wizyjnym robota przemysłowego

W artykule omówiono budowę systemu wizyjnego robota przemysłowego dokonującego operacji montażu układów scalonych w pewnym urządzeniu elektronicznym. Głównym celem systemu wizyjnego jest rozpoznawanie układów scalonych na podstawie oznaczeń na ich obudowach. System wizyjny został zrealizowany w oparciu o wieloprocesorowy układ DSP, którego budowę pokrótce opisano. W dalszej części omówiono szczegółowo wszystkie operacje związane ze wstępnym przetwarzaniem obrazów, ich segmentacją, ekstrakcją cech i rozpoznawaniem. Zaprezentowano uzyskane wyniki rozpoznawania oraz dokonano dyskusji oferowanych przez wieloprocesorowy układ TMS320C80 sposobów zrównoleglenia obliczeń.

Image recognition in an industrial robot vision system

In the paper an industrial robot vision system was described. The robot is to perform a montage operation of integrated circuits in some electronic device. The main purpose of the vision system is to recognise the integrated circuits basing on the descriptions on their packages. The vision system was realised on a multiprocessor DSP chip that key features were presented. Further all the image processing operations such as pre-processing, segmentation, feature extraction and image recognition were described in details. The image recognition results were presented and the computation parallelization methods offered by the TMS320C80 multiprocessor chip were discussed.

1. WSTĘP

W artykule omówiono budowę i zasadę działania skonstruowanego w Laboratorium Zastosowań Informatyki Katedry Automatyki Akademii Górniczo-Hutniczej w Krakowie systemu wizyjnego robota przemysłowego. W proponowanym stanowisku badawczym robot przemysłowy ma być wykorzystany do realizacji operacji montażu układów scalonych na płycie drukowanej, stanowiącej fragment pewnego urządzenia elektronicznego. Na wstępie założono, że liczba układów scalonych podlegających operacji montażu wynosi osiem, przy czym każdy z tych układów scalonych jest innego typu, natomiast posiadają one identyczny typ obudowy o czternastu wyprowadzeniach. Ponieważ w takim przypadku nie jest możliwe rozpoznanie użytych tutaj układów scalonych na podstawie kształtu, bądź wymiarów ich obudowy, założono, że będą one rozpoznawane na podstawie napisów, bądź oznaczeń umieszczonych na górnej części ich obudowy. W celu rozwiązania postawionego wyżej zadania należało wyposażyć robota w system wizyjny, dzięki czemu zyskał on możliwość „oglądania” swojego stanowiska pracy. W tym konkretnym przypadku system wizyjny robota został zrealizowany w oparciu o kamerę dostarczającą obrazu stanowiska pracy robota oraz kartę z systemem wielo-

procesorowym DSP firmy Texas Instruments typu TMS320C80. Kamera produkcji Japońskiej nosząca oznaczenie EVI G-21, jest kamerą zrealizowaną w oparciu o technikę CCD, dostarczającą kolorowego obrazu w standardzie PAL. Posiada on zdolność samoogniskowania się obiektywu (funkcja autofocus), przez co unika się komplikowania i tak trudnych zagadnień przetwarzania i rozpoznawania obrazów poprzez wprowadzenie do systemu wizyjnego obrazu nieostrego, bądź zniekształconego przez system optyczny kamery. Ponadto kamera ta została wyposażona w urządzenie zmieniające wielkość jej ogniskowej (funkcja autozoom), dzięki czemu możliwa jest dobra praca systemu wizyjnego w przypadku różnego oddalenia obserwowanych obiektów. Oprócz tego kamera posiada ruchomą głowicę ze światłoczułą matrycą CCD o możliwościach ruch ± 30 stopni w płaszczyźnie horyzontalnej i ± 15 stopni w płaszczyźnie wertykalnej, co znacznie zwiększa zasięg jej pola widzenia. Możliwa jest również regulacja przesłony kamery, którą można ustawić w dwóch położeniach odpowiadających jaśniejszemu i ciemniejszemu obrazowi. Wszystkie dostępne funkcje kamery, takie jak regulacja położenia światłoczułej głowicy, zmiana wielkości ogniskowej, zmiana jasności obrazu mogą być zmieniane przy pomocy działającego na podczerwień pilota, bądź mogą być ustalone przez komputer PC poprzez łącze transmisji szeregowej typu RS232. Obraz pozyskany przez kamerę doprowadzany jest kablem ekranowanym do układów akwizycji obrazu karty z systemem wieloprocessorowym TMS320C80, gdzie podlega on kwantyzacji jako obraz monochromatyczny o rozdzielczości 768 na 576 pikseli i 256 poziomach szarości (ang. gray levels) oraz dalszym operacjom przetwarzania, segmentacji i rozpoznawania. Fakt posiadania przez system TMS320C80 czterech pomocniczych 32-bitowych procesorów DSP może mieć duży wpływ na przyspieszenie czasochłonnych i wymagających dużej mocy obliczeniowej operacji przetwarzania i rozpoznawania obrazów.

2. SYSTEM WIELOPROCESOROWY TMS320C80

Układ TMS320C80 MVP (ang. multimedia video processor) jest urządzeniem wieloprocessorowym wykonanym w postaci pojedynczego układu scalonego, zaprojektowanym i wykorzystywanym w aplikacjach takich jak przetwarzanie obrazów, dwu i trójwymiarowa grafika oraz grafika tworzona dla potrzeb wirtualnej rzeczywistości, a także kompresja cyfrowych sygnałów wizyjnych i audio. [3] Moc obliczeniowa systemu MVP potrafi sprostać także wymaganiom stawianym przez aplikacje z dziedziny systemów nadzorczych, rozpoznawania obrazów i różnorodnych zastosowań, związanych z rynkiem telekomunikacyjnym. W układzie TMS320C80 zintegrowano w postaci pojedynczego obwodu scalonego następujące podzespoły:

- pięć odznaczających się dużą mocą obliczeniową, w pełni programowalnych procesorów
- rozbudowany kontroler bezpośredniego dostępu do pamięci DMA (ang. direct memory controller) wraz z interfejsem zewnętrznej pamięci DRAM, SRAM i VRAM
- 50 KB pamięci SRAM
- układ nadzorujący proces akwizycji sygnałów video (ang. video timing control)

3. LABORATORYJNE STANOWISKO ROZPOZNAWANIA OBRAZÓW

Laboratoryjne stanowisko montażu zostało zorganizowane w ten sposób, że w regularnym odstępie czasu τ w przestrzeni roboczej robota przemysłowego, obserwowanej przez kamerę pojawia się nowy układ scalony, o nie znanym z góry typie. Przy czym układ ten może znajdować się w dowolnym punkcie przestrzeni roboczej o uprzednio ściśle określonym kształcie i rozmiarze. W przypadku rozważanego stanowiska laboratoryjnego założono, że przestrzeń robocza jest kwadratem. Układ scalony pojawiający się w przestrzeni roboczej musi zostać

tam umieszczony napisem znajdującym się na jego obudowie do góry, tak aby był on widoczny przez kamerę, natomiast jego orientacja przestrzenna może być dowolna, to znaczy może być on obrócony o dowolny kąt. Zatem robot, od momentu pojawienia się układu scalonego posiada czas τ na rozpoznanie typu układu scalonego i dokonanie operacji jego montażu, ponieważ po upływie czasu τ w przestrzeni roboczej pojawi się nowy układ scalony wymagający analogicznych operacji. Taki tryb pracy nakłada na system komputerowy sterujący pracą robota ostre ograniczenia czasowe, gdyż rozpoznanie układów scalonych i wyznaczenie trajektorii ruchu ramienia robota musi zostać dokonane w czasie rzeczywistym, w nieprzekraczalnym limicie czasu τ .

4. POZYSKANIE OBRAZU I JEGO WSTĘPNE PRZETWARZANIE

Dostarczony przez kamerę obraz w standardzie PAL poddawany jest w układzie akwizycji obrazu karty z systemem wieloprocesorowym DSP kwantyzacji z rozdzielczością przestrzenną 768 na 576 pikseli i 256 poziomami szarości. Z tak pozyskanego obrazu do dalszej obróbki brany jest jedynie jego fragment w postaci kwadratu o rozmiarze 512 na 512 pikseli umieszczony w prawym dolnym rogu obrazu pozyskanego przez układ akwizycyjny karty. Pozyskany w postaci cyfrowej obraz umieszczony w odpowiednim obszarze pamięci na początku poddawany jest operacji filtracji, która ma za zadanie eliminację towarzyszącego sygnałowi wizyjnemu szumu termicznego generowanego w strukturze CCD oraz zmniejszenie wpływu powstających w torze transmisji sygnału zakłóceń. Spośród wielu możliwych rozwiązań wybrano filtrację medianową o pięcioelementowej masce, obejmującej rozważany piksel oraz jego czterech najbliższych sąsiadów, czyli piksele leżące bezpośrednio nad, pod oraz z jego prawej i lewej strony.



Rys. 1 Obraz po binaryzacji

Zaletą filtracji medianowej w porównaniu z innymi metodami np. filtracji konwolucyjnej polega na tym, że filtr medianowy wygładzając obraz zachowuje jednocześnie krawędzie i kontury obiektów, które w każdej innej metodzie filtracji obrazu ulegają rozmyciu i zamaza-

niu. Działanie filtru medianowego polega na posortowaniu pikseli od najmniejszej do największej wartości ich stopni szarości, a następnie wybraniu wartości leżącej dokładnie pośrodku, czyli właśnie mediany tak posortowanych liczb i przypisaniu jej rozważanemu pikselowi. [1] Po zakończeniu filtracji uzyskany obraz poddawany jest operacji binaryzacji. Binarizacja obrazu polega na przypisaniu każdemu z pikseli jednej z tylko dwóch możliwych poziomów jasności odpowiadających czerni i bieli, w zależności od tego czy stopień szarości rozważanego piksela jest mniejszy czy też większy od pewnej wartości progowej. Z procesem binaryzacji związana jest znaczna redukcja informacji niesionych przez obraz. Zatem niezwykle ważną kwestią jest odpowiedni dobór wartości progowej dla procesu binaryzacji, w taki sposób aby nie utracić istotnej informacji zawartej w obrazie. Obraz powstały po operacji binaryzacji można zobaczyć na rys.1. Otrzymany w procesie binaryzacji obraz podlega następnie operacji filtracji logicznej. Najpierw stosowana jest filtracja logiczna pozioma, a następnie pionowa, obie o trójelementowych maskach filtru. Celem tak przeprowadzonej filtracji logicznej jest usunięcie z obrazu niejednorodności występujących w postaci pojedynczych białych punktów na czarnym tle lub pojedynczych czarnych punktów na białym tle. Eliminacji podlegają także poziome i pionowe linie o szerokości jednego piksela występujące na tle o odmiennym poziomie jasności. Dzięki zastosowaniu operacji filtracji logicznej białe napisy znajdujące się na czarnych obudowach układów scalonych stają się bardziej wyraźne i czytelne. Obraz otrzymany po zakończeniu operacji filtracji logicznej można zobaczyć na rys. 2.



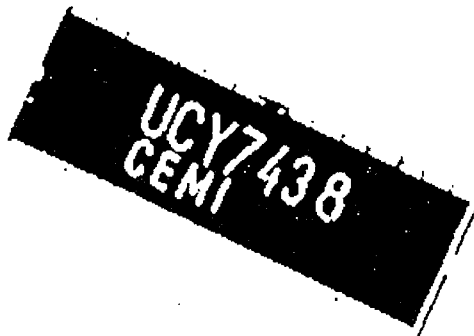
Rys. 2 Obraz po filtracji logicznej

Po zakończeniu operacji wstępnego przetwarzania obrazu, których celem było poprawienie jego jakości i uwypuklenie istotnych z punktu widzenia rozpoznawania informacji obraz poddawany jest segmentacji, w celu wydzielenia jego fragmentu odpowiadającego obrazowi obudowy układu scalonego.

5. OBRÓT OBRAZU I JEGO SEGMENTACJA

Układ scalony podlegający operacji rozpoznawania może w ogólnym przypadku posiadać dowolną orientację przestrzenną, to znaczy może być on obrócony o dowolny kąt φ . Pomiar kąta skręcenia układu scalonego dokonywany jest w ten sposób, że na zbinaryzowanym obra-

zie analizowane są kolejno piksele należące do dwóch pionowych linii, o odciętych odpowiednio 192 i 320, aż do napotkania pierwszych czarnych pikseli. Rzędne tych czarnych pikseli reprezentujących brzeg obudowy układu scalonego są zapamiętywane i następnie wykorzystywane do obliczenia kąta skrócenia układu. Ponieważ linia brzegowa obudowy układu scalonego nie jest idealnie gładka, ale posiada liczne rozgałęzione wypustki, koniecznym stało się poddanie obrazu, przed pomiarem kąta skrócenia operacji erozji, która sprawia, że brzeg obudowy układu staje się bardziej podobny do linii prostej. Brak tej operacji powodował, że wartość kąta skrócenia układu obarczona była znacznym błędem. Obraz po zastosowaniu operacji erozji można zobaczyć na rys. 3.



Rys. 3 Obraz po erozji

W kroku następnym dokonywana była operacja obrotu obrazu o kąt β względem punktu należącego do dolnego brzegu obrazu, charakteryzującego się współrzędnymi o odciętej 256 i rzędnej 512 pikseli. W wyniku tak przeprowadzonej operacji obrotu całego obrazu linie brzegowe układów scalonych stają się równoległe względem osi skojarzonego z obrazem układu współrzędnych. Rozwiązanie to posiadało jednak dość poważną wadę polegającą na tym, że w obrazie przetworzonym powstawały liczne dziury, będące pikselami o współrzędnych nie wyliczonych przez procedurę dokonującą operacji obrotu obrazu. Rozwiązanie takie prowadziło do znacznego pogorszenia jakości rozpoznawanego obrazu oraz było zupełnie nie do zaakceptowania w systemie wizyjnym robota. W celu ominięcia powyższych trudności należało posłużyć się transformacją odwrotną operacji obrotu obrazu. [2] W tym przypadku należało wziąć po kolei wszystkie piksele należące do obrazu przetworzonego, o współrzędnych x_{new}, y_{new} , a następnie znaleźć współrzędne x_{old}, y_{old} odpowiadających im w obrazie pierwotnym pikseli. Obliczając wartości współrzędnych x_{old}, y_{old} należało najpierw sprawdzić czy nie leżą on poza brzegiem obrazu. Po zakończeniu operacji obrotu obrazu linie brzegowe obudowy układu scalonego były już równoległe do osi układu współrzędnych, pokrywających się z lewym i dolnym brzegiem obrazu. W tym momencie można już było przystąpić do wykonania operacji segmentacji obrazu, czyli wydzielenia z całości obrazu jedynie jego fragmentu reprezentującego widok obudowy układu scalonego. W kroku kolejnym przeglądane są po kolei, począwszy od górnej krawędzi obrazu jego wiersze, aż do napotkania pierwszego czarnego pikselu, reprezentującego najbardziej wysunięty ku górze punkt należący

do obudowy układu scalonego. Podobnie, począwszy od lewej krawędzi układu scalonego badane są kolejne kolumny obrazu, aż do napotkania pierwszego czarnego piksela, odpowiadającego najbardziej wysuniętemu ku lewej stronie punktowi obudowy układu. Następnie w oparciu o wartości współrzędnych tych dwóch punktów na obraz obudowy układu scalonego nakładana jest prostokątna maska, o wymiarach 200 na 75 pikseli, wycinająca z obrazu jedynie fragment obudowy układu scalonego, zawierający oznaczenia, na podstawie których układ ten może zostać rozpoznany. Obraz układu scalonego po operacji maskowania zamieszczono na rys. 4.



Rys. 4 Obraz po operacji maskowania

6. POMIAR WARTOŚCI CECH OBRAZU

Fragment obrazu o wymiarach 200 na 75 pikseli reprezentujący fragment obudowy układu scalonego, z widocznymi na niej oznaczeniami, dzielony jest następnie dwoma pionowymi liniami na trzy sektory, z których dwa skrajne, lewy i prawy posiadają wymiary 70 na 75 pikseli, a sektor środkowy posiada wymiary 60 na 75 pikseli. Jako cechy opisujące obraz przyjęto liczby białych pikseli w poszczególnych sektorach. Należy zwrócić uwagę na fakt, że na tym etapie przetwarzania obrazów nastąpiła drastyczna redukcja informacji zawartej w obrazie, ponieważ zamiast wielu tysięcy pikseli, cały obraz opisywany jest jedynie przez trzy liczby niosące jednak całą istotną, z punktu widzenia danej aplikacji, informację o obrazie. Jak zostanie pokazane w punkcie następnym opis obrazu jedynie w postaci trzech liczb naturalnych, jest, w przypadku rozpoznawania układów scalonych w systemie wizyjnym robota, całkowicie wystarczający do jego prawidłowego rozpoznania. Przykładowe zmierzone wartości cech dla układu scalonego typu UCY7453 przedstawiono w tab. 1.

	pierwszy sektor obrazu	drugi sektor obrazu	trzeci sektor obrazu
UCY7453	524	914	784

Tab.1 Przykładowe wartości cech dla układu UCY7453

7. ROZPOZNAWANIE OBRAZÓW

Spośród wielu możliwych rozwiązań problemu rozpoznawania obrazów zdecydowano się zastosować metodę minimalnoodległościową, głównie ze względu na prostotę wykorzystywanego w niej aparatu matematycznego oraz wzięto pod uwagę jej przejrzystość i intuicyjną jasność. Istota tej metody polega na pomiarze odległości w przestrzeni cech, jakie dzieli rozpoznawany obiekt od wszystkich możliwych klas jego przynależności. Na koniec typowana jest ta klasa, którą z nieznanym obiektem dzieli najmniejsza odległość. Spośród wielu możliwych wariantów tej metody, związanych głównie z doбором odpowiedniej metryki, wybrano

znowuż wariant najprostszy, czyli metrykę euklidesową. W rozważanym tutaj przypadku pomiar trzech cech dla każdego z układów scalonych, generuje euklidesową przestrzeń metryczną o dwudziestu czterech wymiarach. Odległość pomiędzy dwoma punktami takiej przestrzeni oblicza się wyciągając pierwiastek z sumy kwadratów różnic odpowiednich współrzędnych. W celu skrócenia czasu obliczeń zdecydowano się obliczać tylko kwadraty odległości, które dzieląc układ scalony o nieznaney przynależności od każdej z ośmiu klas. Następnie wykonano szereg eksperymentów mających na celu umożliwienie oceny poprawności działania systemu wizyjnego wykorzystującego powyższą metodę rozpoznawania. Otrzymane wyniki rozpoznawania dla wszystkich ośmiu typów układów scalonych zamieszczono w tab. 2.

typ układu scalonego	liczba rozpoznań prawidłowych	liczba rozpoznań błędnych
7453	100	0
7406N	100	0
7404	100	0
7403N	100	0
7417	99	1
7438	100	0
6410	98	2
1201	100	0

Tab. 2 Wyniki rozpoznawania układów scalonych

Ogółem przeprowadzono 800 prób rozpoznania typu układów scalonych z czego 797 zakończyło się sukcesem, a tylko w 3 przypadkach została wytypowana niewłaściwa klasa przynależności układu, co daje odsetek ponad 99% rozpoznań prawidłowych.

8. ZASTOSOWANE METODY ZRÓWNOLEGLENIA OBLICZEŃ

Architektura wieloprocessorowego systemu TMS320C80 umożliwia implementację wielu różnych metod zrównoleglenia obliczeń, dzięki czemu można uzyskać znaczne ich przyśpieszenie.

8.1. Architektura macierzowa

Wszystkie procesory PP wykonywały te same zadania, ale każdy na innej ćwiartce obrazu:

- filtracja medianowa
- binaryzacja obrazu
- filtracja logiczna pozioma i pionowa
- erozja obrazu
- obrót obrazu

Z kolei procesor MP dokonywał ekstrakcji cech, obliczał odległość w przestrzeni cech i typował klasę przynależności.

8.2. Architektura potokowa

Z procesorów utworzono strukturę potokową o pięciu stacjach, w których wykonywane były następujące operacje:

1. MP: filtracja medianowa

2. PP0: binaryzacja obrazu, filtracja logiczna pozioma
3. PP1: filtracja logiczna pionowa, erozja obrazu
4. PP2: obrót obrazu
5. PP3: ekstrakcja cech, pomiar odległości w przestrzeni cech i typowanie klasy przynależności

8.3. Architektura hybrydowa

Z procesorów utworzono strukturę potokową o czterech stacjach, przy czym stacja trzecia składała się z dwóch procesorów PP pracujących współbieżnie na połówkach obrazu. W poszczególnych stacjach potoku wykonywane były następujące operacje:

1. MP: filtracja medianowa
2. PP0: binaryzacja obrazu, filtracja logiczna pozioma i pionowa, erozja obrazu
3. PP1, PP2: obrót obrazu
4. PP3: ekstrakcja cech, pomiar odległości w przestrzeni cech i typowanie klasy przynależności

9. WNIOSKI KOŃCOWE

Zaprezentowany w pracy system wizyjny robota może znaleźć praktyczne zastosowanie w przypadku automatycznego montażu układów scalonych przez robota, bądź też system taki może posłużyć do weryfikacji prawidłowości wykonania operacji montażu, to znaczy określenia, czy w danym miejscu obwodu drukowanego został umieszczony właściwy układ scalony. Spośród przetestowanych architektur wieloprocessorowych najkrótszy czas obliczeń uzyskano dla konfiguracji macierzowej, gdzie wszystkie procesory PP wykonywały te same operacje, ale na różnych zbiorach danych i były przez cały czas równomiernie obciążone. Zdecydowanie najgorsza okazała się konfiguracja potokowa, ze względu na różny czas trwania operacji wykonywanych w poszczególnych stacjach potoku. Operacja obrotu obrazu wymagała zdecydowanie najdłuższego czasu obliczeń, co powodowało konieczność wstrzymania pracy pozostałych stacji potoku w celu zapewnienia synchronizacji. Pewną poprawę przyniosło tutaj zastosowanie konfiguracji hybrydowej, gdzie stacje potoku dokonującą obrotu obrazu zrealizowano w postaci dwóch pracujących współbieżnie na połówkach obrazu procesorów PP.

LITERATURA

- [1] Tadeusiewicz R., *Systemy wizyjne robotów przemysłowych*, WNT Warszawa 1992
- [2] Watkins Ch., Sadun A., Marenka S., *Nowoczesne metody przetwarzania obrazu*, WNT Warszawa 1995
- [3] *TMS320C80 System-Level Synopsis*, materiały katalogowe firmy Texas Instruments

Pracę wykonano w ramach grantu KBN 8T 11A 01014 „Konstrukcje algorytmów i programów dla wieloprocessorowych układów DSP dla wybranych systemów czasu rzeczywistego.”