

Ryszard S. Choraś

Instytut Telekomunikacji i Elektrotechniki
ATR Bydgoszcz

STRUKTURALNE METODY ANALIZY OBRAZÓW W SYSTEMIE WIZYJNYM ROBOTA PRZEMYSŁOWEGO

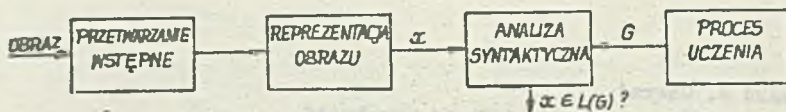
Streszczenie. W referacie przedstawiono strukturalne metody analizy obrazów w systemie wizyjnym robota. Podstawową cechą odróżniającą podejście strukturalne od innych jest bezpośrednie wykorzystanie informacji o strukturze obrazu w procesie rozpoznawania. Umożliwia to opis dużych zbiorów złożonych obrazów przy wykorzystaniu małych zbiorów prymtywów obrazowych i reguł gramatycznych.

1. Wstęp

System strukturalnego rozpoznawania obrazów składa się z : bloku przetwarzania wstępnego obrazu, bloku opisu obrazu, bloku realizującego analizę syntaktyczną i bloku realizującego proces uczenia /rys.1/.

Blok przetwarzania wstępnego obrazu realizuje operacje, które w ogólności można sprowadzić do: a/ kodowania i b/ filtracji i/lub restauracji obrazu. Kodowanie obrazu umożliwia wyeliminowanie redundancji informacji i efektywne zapamiętywanie danych obrazu. Filtracja umożliwia poprawianie jakości obrazu poprzez redukcję szumów i zniekształceń. Umożliwia ona restaurację obrazu zakłóconego efektami wywołanymi, np. ruchem kamery, korekcję zniekształceń geometrycznych itp. Kodowanie i filtracja i/lub restauracja obrazu są problemami wykraczającymi poza wąskie ramy tego referatu [2,3,5].

Blok opisu obrazu realizuje operację wydzielenia z obrazu prymtywów obrazowych wykorzystywanych następnie do jego opisu. Istnieją dwa podstawowe podejścia umożliwiające podział obrazu na części. Pierwsze bazuje na obserwacji, że linie konturowe zawierają ważne informacje o obiekcie, więc możliwa jest klasyfikacja obiektu na podstawie jego konturów. Drugie podejście oparte jest na stwierdzeniu, że nie zawsze możliwe jest wydzielenie linii konturowych /tę i obiekt nie różnią się lub różnią nieznacznie pod względem jasności/ i nie zawsze informacja o konturach obiektu jest wystarczająca do dalszego przetwarzania. Drugie podejście w systemach wizyjnych robotów /zwłaszcza przemysłowych/ nie znajduje zastosowania ze względu: 1/na niepraktyczną, skomplikowaną implementację hardware'ową oraz skomplikowane algorytmy obliczeniowe, 2/na to, że prob-



Rys. 1. System strukturalnego rozpoznawania obrazów

lezy wynikające z niewidoczności linii konturowych można rozwiązać stosując specjalne oświetlenie. Z tych powodów podejście takie nie będzie uwzględniane w dalszych rozważaniach.

Następnie obraz przedstawiany jest za pomocą struktur językopodobnych /np. łańcuch lub drzewo, graf/. Obraz jest reprezentowany przez zbiór prymitywów oraz specyficznych syntaktycznych operacji, np. przy wykorzystaniu operacji konkatenacji obraz przedstawiany jest w postaci łańcucha prymitywów.

Język opisu obrazu L jest generowany przez gramatykę $G = \langle V_N, V_T, P, S \rangle$, gdzie V_N i V_T są odpowiednio zbiorami symboli nieterminalnych /zmiennych/podobrzów i symboli terminalnych /stałych/ lub prymitywów obrazowych; $S \in V_N$ jest symbolem początkowym; P jest skończonym zbiorem reguł produkcji /reguł gramatycznych/ postaci $p_i: \alpha \rightarrow \beta, \alpha \in V_N \cup V_T^*$; $\beta \in V_N \cup V_T^*$; V^* jest zbiorem wszystkich możliwych sentencji symboli $\in V$, łącznie z symbolem pustym.

Stosując reguły produkcji z P i startując od S generujemy łańcuch lub sentencję symboli terminalnych $v \in V_T^*$, co oznaczamy $S \Rightarrow v$. Zbiór wszystkich sentencji symboli terminalnych generowanych za pomocą reguł produkcji tworzy język $L/G/$ generowany przez gramatykę G

$$L/G/ = \{ v \mid S \Rightarrow v, v \in V_T^* \}$$

Analizę syntaktyczną /parsing/ sentencji opisujących obraz można przeprowadzić dwoma sposobami: bottom-up lub top-down. W pierwszym przypadku rozpoczynamy analizę od sentencji v i próbując stosować reguły produkcji w odwrotnej kolejności dążymy do osiągnięcia symbolu S . Jeżeli jest to nie możliwe to, v nie jest sentencją $L/G/$ i jest odrzucane. W przypadku drugiego sposobu analizę syntaktyczną rozpoczynamy od symbolu S i stosując reguły produkcji dążymy do uzyskania sentencji v . Jeżeli jest to niemożliwe to $v \notin L/G/$.

Należy zauważyć, że analizę obrazu rozpoczynamy od macierzy obrazu, a nie od łańcucha /sentencji/ symboli. Symbole są wydzielane z obrazu. Analiza sentencji symboli pozwala na poprawienie błędów lub na ponowne wydzielenie zbioru symboli.

Proces uczenia w przypadku syntaktycznego rozpoznawania i analizy obrazów może wprowadzać nowe gramatyki, tj. wprowadzać nowe reguły produkcji bazując na obserwacji skończonego zbioru sentencji /łańcuchów/.

2. Wydzielenie konturów obiektów

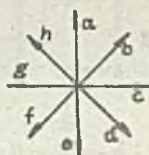
Kształty obiektów, np. części maszynowych mogą być adekwatnie opisane przez proste geometryczne pojęcia takie jak linia prosta, łuk itp. Ponieważ obiekty są dowolnie położone w obrazie, pierwszym krokiem jest określenie położenia obiektu i wydzielenie linii konturowych obiektów. Wydzielenie linii konturowych w obrazie można zrealizować stosując różne techniki, np. opisaną w [1,2,3]. Uzyskany obraz zawiera oprócz linii konturowych również pojedyncze punkty lub grupy punktów nie tworzących żadnego konturu. Wprowadzając operator filtracji [2,3,4] uzyskujemy obraz w którym wszystkie punkty obrazu są punktami linii konturowych. Na rys.2 przedstawiono obraz wejściowy /a/ oraz obraz konturowy /b/ pewnej części maszyny do pisania.

Obraz będzie reprezentowany przez zbiór prymitywów obrazowych i specyficznych operacji syntaktycznych. Wybór prymitywów ma decydujące znaczenie dla uzyskania poprawnego opisu syntaktycznego. Prymitywy powinny być podstawowymi elementami obrazu, łatwe wydzielanymi lub rozpoznawalnymi metodami nie-lingwistycznymi umożliwiającymi zwięzły lecz adekwatny opis danych obrazu.

Prymitywy obrazowe tworzą segmenty linii prostych lub krzywych. Zbiór prymitywów obrazowych przedstawiono na rys.3.



Rys.2. Obrazy wejściowy /a/, konturowy /b/



Rys.3. Zbiór prymitywów obrazowych

3. Strukturalny opis i rozpoznawanie obrazów

W syntaktycznym podejściu do rozpoznawania obrazów, obrazy są reprezentowane przez sentencje języka charakteryzowanego przez gramatykę. Sentencje mogą być drzewami. Jeżeli obraz jest opisywany przez drzewa, to jest łatwe generowanie przez gramatykę drzew. Gramatyka drzew jest płatką $G_t = \langle V_N, V_T, r, P, S \rangle$, gdzie: V_T i V_N są zbiorami symboli terminalnych /prymitywów obrazowych/ i nieterminalnych /pedeobrazów/, r jest liczbą gałęzi wychodzących

z punktu a, /iii/ P jest skończonym zbiorem reguł produkcji postaci $\alpha \rightarrow \beta$, gdzie α i β są drzewami i S jest symbolem początkowym.

Język generowany przez gramatykę drzew G_t jest następujący

$$L/G_t/ = \left\{ \alpha \in t \mid \text{jeżeli istnieje } Y \in S \text{ takie że } Y \Rightarrow \alpha \text{ jest w } G_t \right\}$$

W procesie uczenia wykorzystuje się znane struktury obiektów dla których tworzony jest szereg początkowych reguł gramatyki drzew. Obraz jest przetwarzany przy wykorzystaniu początkowego zbioru reguł gramatycznych. Mając geometryczną interpretację konturu obiektu porównujemy ją z wynikami przetwarzania. Dla tych struktur obiektów dla których istnieje geometryczna interpretacja / której nie można uzyskać na podstawie wyników przetwarzania / reguły gramatyczne generujące te struktury są dodawane do początkowego zbioru reguł gramatycznych, po czym obraz jest przetwarzany ponownie w celu sprawdzenia tych reguł. Po skończonej liczbie kroków otrzymujemy końcowy zbiór reguł gramatycznych.

Dla konturu przedstawionego na rys.2 gramatyka ucząca jest następująca:

$$G = /V, r, P, S/$$

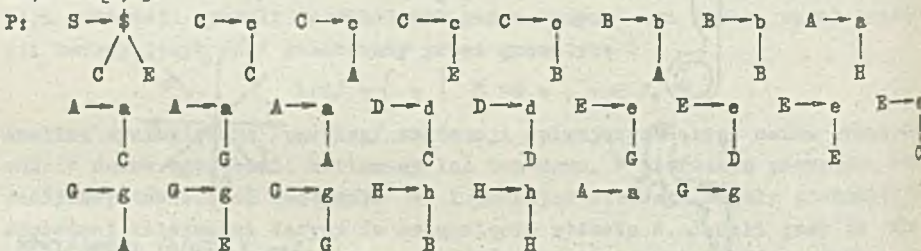
$$V = \{S, A, B, C, D, E, F, G, H, a, b, c, d, e, f, g, h, \$\}$$

$$V_T = \{\$, a \uparrow, b /, c \rightarrow, d \setminus, e \downarrow, f /, g \leftarrow, h \setminus\}$$

$$r/a/ = \{2, 1, 0\} ; r/b/ = \{2, 1, 0\} ; r/c/ = \{2, 1, 0\} ; r/d/ = \{2, 1, 0\}$$

$$r/e/ = \{2, 1, 0\} ; r/f/ = \{2, 1, 0\} ; r/g/ = \{2, 1, 0\} ; r/h/ = \{2, 1, 0\}$$

$$r/\$/ = \{2, 1\}$$



Po procedurze uczenia opisanej w [5] otrzymujemy następujące reguły gramatyki drzew do syntaktycznego rozpoznawania zbioru części maszyny do pisania:

$$G = /V, r, P, S/$$

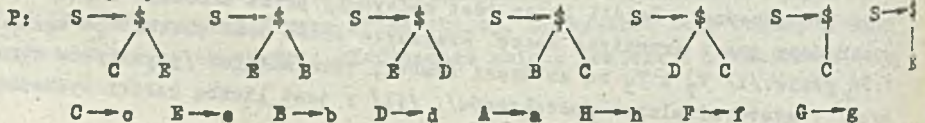
$$V = \{S, A, B, C, D, E, F, G, H, a, b, c, d, e, f, g, h, \$\}$$

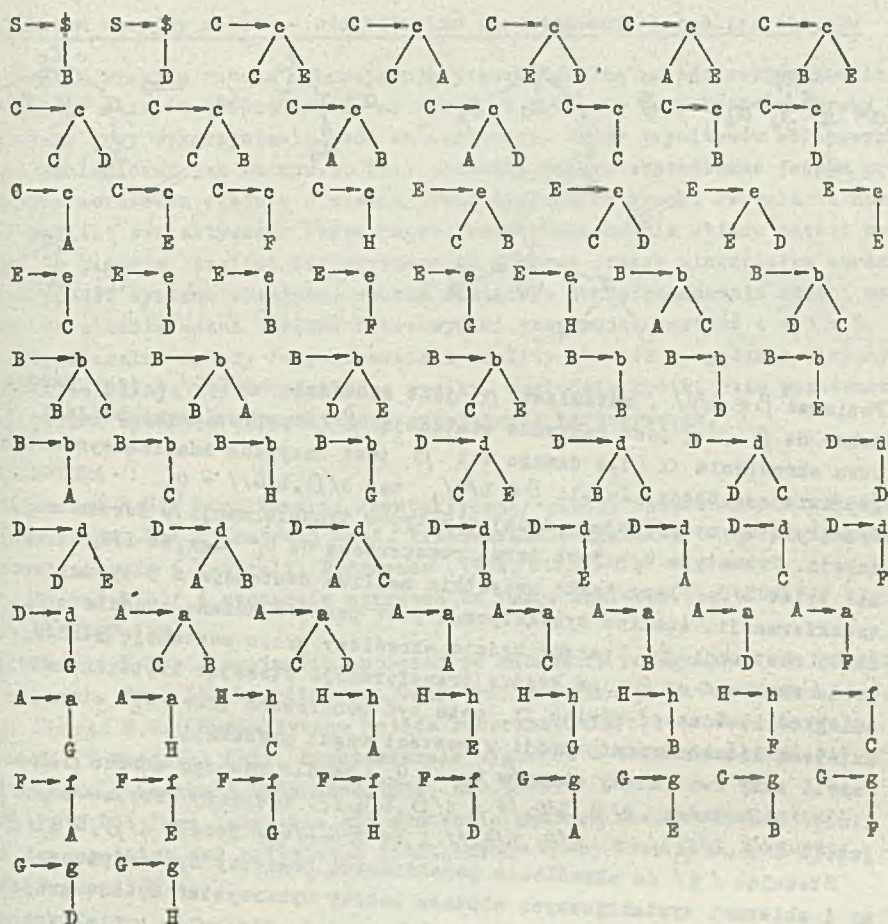
$$V_T = \{\$, a \uparrow, b /, c \rightarrow, d \setminus, e \downarrow, f /, g \leftarrow, h \setminus\}$$

$$r/a/ = \{2, 1, 0\} ; r/b/ = \{2, 1, 0\} ; r/c/ = \{2, 1, 0\} ; r/d/ = \{2, 1, 0\}$$

$$r/e/ = \{2, 1, 0\} ; r/f/ = \{2, 1, 0\} ; r/g/ = \{2, 1, 0\} ; r/h/ = \{2, 1, 0\}$$

$$r/\$/ = \{2, 1\}$$





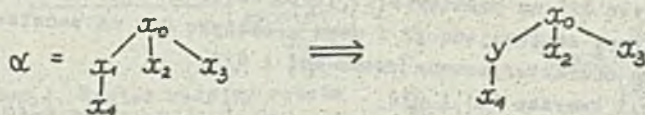
Odległość pomiędzy drzewami /syntaktycznymi opisami dwóch obrazów/ można zdefiniować w terminach najmniejszej liczby operacji podstawień, wstawień, usunięć wymaganych do transformacji jednego drzewa w drugie.

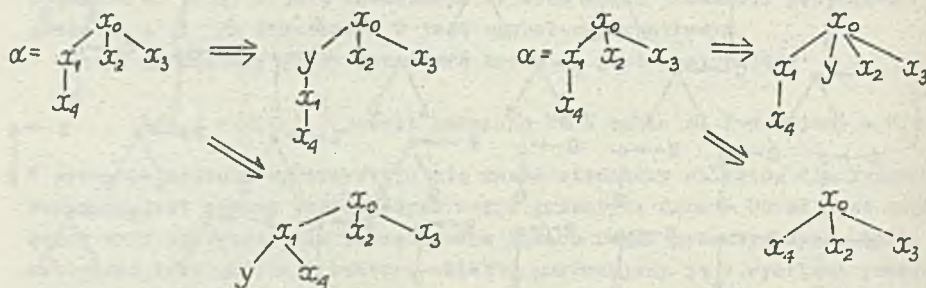
Jeżeli L jest językiem generowanym przez gramatykę drzew, to drzewo $\beta \notin L$ można otrzymać na podstawie drzewa $\alpha \in L$ przez zastosowanie pewnej liczby transformacji. Odległość między β i L określamy przez

$$d/L, \beta / = \min \{ d/\alpha, \beta / \mid \alpha \in L \}$$

111

gdzie operacje transformacji są przedstawiane poniżej





Ponieważ $\beta \notin L/G$, natomiast α jest sentencją $\in L/G$ i jest bardzo podobne do β , α jest nazywane sentencją β z poprawionymi błędami. Procedura określenia α dla danego G i β jest nazywana analizą syntaktyczną poprawiającą błędy. Jeżeli $\beta \in L/G$, to $d/\beta, L/G// = 0$.

Gramatykę G_t modyfikujemy dodając 5 typów transformacji w formie reguł produkcji. Gramatyka G_t jest teraz rozszerzona do G'_t tak, że $L/G'_t//$ zawiera nie tylko $L/G_t//$ lecz także wszystkie możliwe sentencje z 5 typami operacji transformacji. Analiza syntaktyczna jest przeprowadzana zgodnie z G'_t . Dla danej sentencji β , parser będzie określony przez sentencję $\alpha \in L/G_t//$, którą można utworzyć od β na drodze transformacji stosując kryterium minimalnej odległości. Inaczej mówiąc β może być generowane przez G'_t stosując najmniejszą liczbę transformacji w postaci reguł produkcji.

Jeżeli mamy dwie klasy obrazów G_{t1} i G_{t2} to dla pewnego obrazu β możemy obliczyć odległości $d/\beta, L/G_{t1}//$ i $d/\beta, L/G_{t2}//$. Będziemy przyjmować, że $\beta \in L/G_{t1}//$ jeżeli $d/\beta, L/G_{t1}// < d/\beta, L/G_{t2}//$ i że $\beta \in L/G_{t2}//$ jeżeli $d/\beta, L/G_{t2}// < d/\beta, L/G_{t1}//$.

Stosując /1/ do określenia podobieństwa pomiędzy syntaktycznymi obrazami i zbiorami syntaktycznych obrazów możemy wykorzystać metodę grupowania do rozpoznawania obrazów. Procedura grupowania dla celów syntaktycznego rozpoznawania obrazów jest następująca:

Mamy zbiór syntaktycznych obrazów $X = \{x_1, x_2, \dots, x_n\}$ gdzie x_i jest drzewem. Klasyfikujemy x_i , $i = 1 \dots n$ do m grup i wyznaczamy gramatykę G_k , $k = 1, \dots, m$, charakteryzującą każdą grupę.

1/Dla obrazu x_1 tworzymy gramatykę G_{t1} na podstawie x_1 ,

2/Tworzymy parser A_{t1} dla G_{t1} ,

3/Dla obrazu x_2 wykorzystujemy A_{t1} do określenia czy x_2 jest podobne do x_1 .

Porównujemy odległości pomiędzy x_1 i x_2 , tj. $d/x_1, x_2//$ lub $d/L/G_{t1}//, x_2//$ z progiem t . Jeżeli $d/L/G_{t1}//, x_2// < t$ to x_1 i x_2 są w tej samej grupie. Tworzymy gramatykę G_{t2} na podstawie $\{x_1, x_2\}$. Jeżeli $d/L/G_{t1}//, x_2// \geq t$ to tworzymy nową grupę zawierającą x_2 i nową gramatykę G_{t2} na podstawie x_2 . Mamy dwie grupy charakteryzowane przez G_{t1} i G_{t2} .

4/Powtarzamy p.2 tworząc G_{t1} i G_{t2}

5/Powtarzamy p.3 dla nowego obrazu. Otrzymujemy m grup $G_{n1}, G_{n2}, \dots, G_{nm}$.

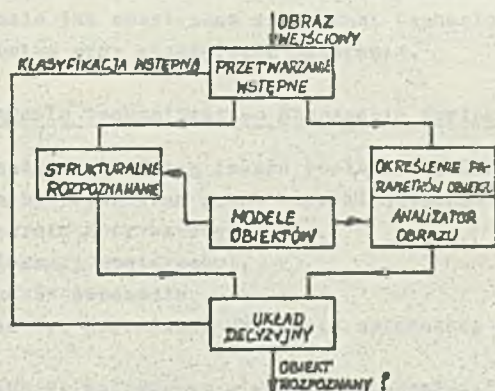
4. System wizyjny robota - strukturalne rozpoznawanie i analiza obrazów

System wizyjny robota wykorzystujący strukturalną metodę rozpoznawania i analizy obrazów przedstawiono na rys.4. W systemie tym, obraz konturowy zakodowano przy wykorzystaniu kodu łańcuchowego. Zbiór prymitywów obrazowych jest analogiczny jak na rys.3. Trzy elementy rastru zastępowano jednym prymitywem obrazowym zgodnie z zasadą przedstawioną na rys.5. Pozwala to uproszczyć analizę syntaktyczną. Eksperyment przeprowadzono dla zbioru części maszyn do pisania. Analizę rozpoczynało od górnego lewego wierzchołka obrazu. W przypadku systemu wizyjnego robota służącego do rozpoznawania części maszyn do pisania można uzyskać dobre wyniki przyjmując wartość $t = 3 + 5$.

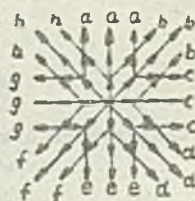
Strukturalne metody rozpoznawania i analizy obrazów w systemie wizyjnym robota pozwalają uzyskać dokładne wyniki, względnie krótki czas przetwarzania, jednak wymagają optymalizacji rozwiązania hardware'owego.

LITERATURA

- [1]. Choraś R.S.: Strukturalne /syntaktyczne/ metody rozpoznawania i analizy obrazów. Mat.I Krajowej Konf. Przetwarzanie sygnałów w telekomunikacji sterowaniu i kontroli, Bydgoszcz, 1984, str.83-87.
- [2]. Choraś R.S.: A syntactic approach to image analysis. Proc.EUSIPCO 83, str.127-129.
- [3]. Choraś R.S.: A syntactic approach to automatic recognition of industrial parts. Proc.3th Scandinavian Conf. on Image Analysis.Copenhagen,1983.
- [4]. Choraś R.S.: Syntaktyczna metoda rozpoznawania części maszynowych, Mat.I Krajowej Konf. Przetwarzanie sygnałów w telekomunikacji, sterowaniu i kontroli, Bydgoszcz, 1984, str.88-92.
- [5]. Fu K.S.: Tree languages and syntactic pattern recognition, Pattern recognition and artificial intelligence, ed.by Chen C.H., Academic Press, 1976.



Rys.4. System wizyjny robota



Rys.5. Zasada kodowania obrazu konturowego

Recenzent: Prof.dr hab.inż.Antoni Woźniak

Wpłynęło do Redakcji do 30.03.1984r.

СИНТАКСИЧЕСКИЕ МЕТОДЫ ОБРАБОТКИ ОБРАЗОВ В СИСТЕМЕ ЗРИТЕЛЬНОГО ВОСПРИЯТИЯ ПРОМЫШЛЕННОГО РОБОТА

Резюме

В статье представлены синтаксические методы обработки образов в системе зрительного восприятия промышленного робота. Основным отличием синтаксического распознавания образов от всех описанных ранее, является непосредственное использование структуры образов в процессе распознавания.

STRUCTURAL METHODS TO PATTERN ANALYSIS IN A MACHINE VISION SYSTEM

Summary

This paper describes structural /syntactic/ approach to image analysis in a machine vision system. By "image analysis" we mean that an image is described by simpler constituents and their relations to each other. Simple constituents depend on the particular problem to be solved. The idea of describing a complex pattern in terms of a /hierarchical/ composition of simpler subpatterns is promising.