

МЕТОДИ ЗА ОБОБЩЕНИЕ В ЙЕРАРХИЧНИ РАЗПОЗНАВАЩИ СИСТЕМИ ОТ СТРУКТУРНО-ЛИНГВИСТИЧЕН ТИП

Гочо Гочев

***Резюме.** Разглежда се йерархична разпознаваща система на три нива. Тя реализира структурно-лингвистичен подход на разпознаване. На базата на структурни признаци на различни нива, които се използват като елементи на терминален речник, се строят фразоструктурни описания за обектите в сцената и на сцената като цяло. Те се анализират с цел разпознаване на обектите в сцената и изграждане на модел за сцената. Дават се примери за анализ на двумерни и тримерни визуални сцени и методи за обобщение и разпознаване на символна информация.*

METHODS FOR GENERALIZATION IN HIERARCHICAL RECOGNITION SYSTEMS OF STRUCTURAL LINGUISTIC TYPE

Gotcho Gotchev

***Abstract.** The paper considers three level hierarchical recognition system. The system implements structural-linguistic recognition approach. Structural features of different levels are used as elements of terminal dictionary. Phrasal-structural descriptions of the objects in the scene and the whole scene are built on the basis of the structural features. The descriptions are then analysed in order to recognise the objects in the scene and to develop a model of the scene. Examples are given for analysis of two- and three-dimensional visual scenes. Methods for generalization and recognition of symbol information are proposed.*

1. Въведение

Структурно-лингвистичният подход на разпознаване се базира на информация за структурата на обектите. Анализирайки дадено структурно-лингвистично описание

Годишник на ТУ - София т. 52/2003 г.
Proceedings of the Technical University-Sofia, v. 52/2003e,

се разкриват подструктури за обекта и като цяло неговата структура.

Процесът е йерархичен като с всяко ниво се асоциира представяне на изображението на обекта, характеризирани с основен примитив като пиксел, структурен признак, ръб, контур и на най-високо ниво разпознати обекти от сцената. Управлението на обработката на информацията може да се реализира, използвайки известните методи: управление от данните (отдолу-нагоре) и управление от моделите (отгоре-надолу). Комбинирането на тези два метода (интерактивно управление) дава ефективна обработка на информацията, тъй като описанията на долните нива се съобразяват с моделите на горните нива, а моделите на горните нива се получават от описанията на долните нива, използвайки обратна връзка. От познавателна гледна точка процесите в тази йерархия за обработка в човешкото зрение най-добре са изучени на най-ниското ниво като физически процеси на възприемане и на най-високо ниво, обосновани с формалната логика. По средата между тези крайни нива лежи концептуална бездна [1]. Някои автори считат че тази бездна може да се преодолее с помощта на символна логика. Но термите и манипулационните правила на тази логика е необходимо да бъдат свързани с конкретна семантика. Така че изводите (описанията), които се получават на различните нива, семантично да свързват най-ниското с най-високото ниво. Друг подход, който застъпва авторът, може да бъде използването на формална езикова система за описание, като термите на различните нива имат конкретни концептуални същности. Необходима е възможността езиковите конструкции от по-ниското ниво да се обобщават до семантични терми за следващото ниво. В резултат на обобщенията на най-високото ниво се получава конкретно понятие (символ). Следвайки този подход, авторът развива структурно-лингвистичен подход за анализ и разпознаване, който се базира на предложен от него фразо-структурен език за описание [2].

2. Методи за обобщение и синтез на структура на йерархична разпознаваща система

Методите за обобщение се базират на фразо-структурния език. Използвайки този език, се конструират структурно-лингвистични описания на изображенията на двумерни (2D) и тримерни (3D) визуални сцени. Терминалният речник на формалната граматика за описание се формира

от структурни (топологични) признаци за изображенията на обектите в сцената. *Топологичният признак представлява свързана конфигурация от пиксели от скелета на обекта, на която може да се присвои семантична стойност.* Описанията за обектите от сцената, формирани на базата на топологичните признаци, се наричат първични. Анализирайки елементите на описанието, може да се разкрива структурата на даден обект и да се прави неговото разпознаване. От друга страна анализирайки елементите на първичното описание, може да се формира терминален речник с по-обобщени признаци и съответно да се формират по-обобщени описания. Броят на нивата зависи от решаваната задача. Обобщението на признаците може да следва схемата: *топологичен признак → линеен признак → 2D фигура → 3D тяло → модел на наблюдаваната визуална сцена.* Този процес може да се представи като сложен S-израз на базата на структурно-лингвистичното описание на дадена визуална сцена (в общия случай 3D):

(<структурно-лингвистичното описание> ((<КСГТ1>
 (<КСГТ2>
 ...
 (<КСГТr>)))
 (<КСГТ> ((<Т1>) (<Т2>)...(<Тn>))) {<Т-тяло>}
 (<Т> ((<стена 1>) (<стена 2>) ... (<стена n>)))
 (<стена> ((<фраза 1>) (<фраза 2>) ... (<фраза m>)))
 (<фраза> ((<P_i>) (<P_j>) (<L_r>)))
 (<P_i> (<атрибутен вектор>))
 (<L_r> (<атрибутен вектор>))

кадето

КСГТ – контурно-свързана група от тела; P-топологичен признак;
 L- линеен сегмент.

Анализирайки нивата на представянето на елементите на структурно-лингвистичното описание в S-израза, подходяща структура би била на три нива – фиг. 1. Тя се осигурява с процедурно знание за обработка и анализ на 2D и 3D визуални сцени. Колко нива на системата ще се използват, зависи от конкретната задача. В случай, че сцената съдържа двумерни обекти, моделът за нея се получава от второто ниво. Ако сцената е тримерна, то класификаторът на второто ниво разпознава тримерните тела, които образуват терминалния речник за описание на тримерната сцена.



Фиг. 1. Йерархична система с три нива

По-надолу се дават примери за анализ, обобщение и разпознаване на обектите на 2D и 3D визуални сцени. Да разгледаме изграждането на модел на сцена съдържаща плоски геометрични обекти – фиг. 2.

На първото ниво се формират първичните описания:

P_1P_2, P_2P_3, P_3P_1 .

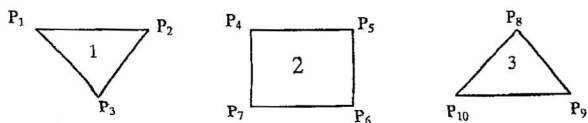
$P_4P_5, P_5P_6, P_6P_7, P_7P_4$.

$P_8P_9, P_9P_{10}, P_{10}P_8$.

На базата на тези описания се прави класификацията на първото ниво. Обучаващите правила са следните:

*Ако в описанието има последователно две отворени фрази и трета, която се затваря в първия признак на първата фраза, то описаният обект е **триъгълник**.*

*Ако в описанието има последователно три отворени фрази и четвърта, която се затваря в първия признак на първата фраза, то описаният обект е **четириъгълник**.*



Фиг. 2. Двумерни геометрични обекти

Разпознатите геометрични фигури се използват като терминален речник за второто ниво.

Резултатът на второто ниво е описанието на модела на сцената на базата на разпознатите обекти и тяхните отношения:

*триъгълник (Об1), четириъгълник (Об2, Д), триъгълник (Об3, Д).**

Съкращенията означават: Об – обект; Д - дясно. Моделът би могъл да се изгради по отношение на всеки обект.

Да разгледаме йерархичното описание, разпознаване и изграждане на модела за тримерна сцена - фиг. 3. Използват се трите нива на системата. На първото ниво се строят първичните описания и се прави разпознаването на видимите стени като двумерни геометрични обекти. На второто ниво се изгражда описанието на телата и сцената като цяло на базата на разпознатите стени, техните отношения и се прави разпознаването на телата съгласно обучението, в зависимост от това каква геометрична фигура е горната стена (респективно основата):

четириъгълник(1,ГС), четириъгълник(2,ДС), четириъгълник(3,ПС).

триъгълник(4,ГС), четириъгълник(5,ЛС), четириъгълник(6,ДС).

петъгълник(7,ГС), четириъгълник(8,ДС), четириъгълник(9,ПС),

*четириъгълник(10,ЛС).**

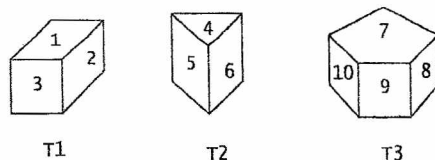
Съкращенията означават: ГС – горна стена; ДС – дясна стена; ПС – предна стена; ЛС – лява стена. Разпознатите тримерни тела образуват терминален речник за третото ниво. Резултатът от третото ниво в йерархията е моделът на възприеманата сцена. Моделът се изгражда по отношение на всяко тяло от сцената:

четириъгълна призма($T1$), триъгълна призма ($T2,Д$), петоъгълна призма($T3,Д$).*

триъгълна призма ($T2$), четириъгълна призма($T1,Л$), петоъгълна призма($T3,Д$).*

петоъгълна призма($T3$), триъгълна призма ($T2,Л$), четириъгълна призма($T1,Л$).*

Съкращенията означават: T – тяло; $Л$ – ляво. На второто и третото ниво се използва същият фразо-структурния език за описание [3]. Реализира се идеята, че езикът за описание е един. Елементите на терминалните речници за горните нива се получават от описанията на по-долното ниво като се използват процедури на разпознаване и обобщение. Възможно е за решението на някои задачи да се предифинират и някои елементи на нетерминалния речник. Връзките за елементите в телата и сцената следват в описанието като отворени фрази. Описанието на всяко тяло се представя като изречение. Сцената се описва като страница.



Фиг. 3. Изображение на визуална сцена с примерни геометрични обекти

Терминалните елементи на второ и трето ниво представляват понятия (имена на класове) с техните атрибути. В процеса на анализа и разпознаването се прави обобщение на възприеманата информация до понятията за обектите в сцената. На тяхна база се прави качествен модел на възприеманата околна среда. За обектите в модела могат да се изчислят количествени характеристики на базата на техните точни атрибути.

Интересна е задачата на обобщение при разпознаването на символна информация. Приемаме, че изображенията на фиг. 4 са представени на едно входно поле. Топологичните признаци се отделят и

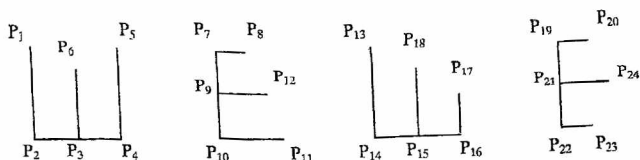
номерирано както е показано на фигурата. Първичните фразо-структурни описания са следните:

$$P_1 P_2; P_2 P_3 P_4; P_4 P_5; P_3 P_6; .$$

$$P_7 P_8; P_7 P_9 P_{10}; P_{10} P_{11}; P_9 P_{12}; .$$

$$P_{13} P_{14}; P_{14} P_{15} P_{16}; P_{16} P_{17}; P_{15} P_{18}; .$$

$$P_{19} P_{20}; P_{19} P_{21} P_{22}; P_{22} P_{23}; P_{21} P_{24}; . *$$



Фиг. 4. Изображения на символи

След класификацията на топологичните признаци и тяхното преиндексиране се получават описанията за обектите в сцената:

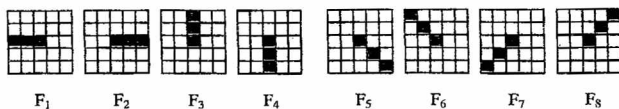
$$P_1 P_2; P_2 P_3 P_4; P_4 P_1; P_3 P_1; .$$

$$P_5 P_6; P_5 P_7 P_2; P_2 P_6; P_7 P_6; .$$

$$P_1 P_2; P_2 P_3 P_4; P_4 P_1; P_3 P_1; .$$

$$P_5 P_6; P_5 P_7 P_2; P_2 P_6; P_7 P_6; . *$$

От поелементното сравнение се вижда, че напълно съвпадат първото и третото, второто и четвъртото описание. Към един и същи клас се отнасят и обектите от изображението. По този начин може да се извърши класификацията на описанията. Топологичните признаци, които влизат в описанията, могат да се представят като съединени комбинации от синтезиращи признаци - фиг. 5 Синтезиращите признаци по същество представляват полулинейни елементи. Като се анализира съставът на тези елементи, обединени в даден характерен признак, може да се направи обобщение за състава на линейните (обобщените) признаци в структурата на изображението.



Фиг. 5. Синтезиращи признаци

На конкретни съединения на полулинейните елементи се поставят в съответствие конкретни обобщени признаци. Търсенето на обобщените линейни признаци и техните връзки се извършва на двоично решаващо дърво. Възлите на това дърво са полулинейните елементи от фиг. 5, а дъгите показват наличие или не на даден полулинейен елемент. Изходите на дървото означават съответните обобщени признаци и връзките между тях. В различните фрази на първичното описание за даден обект се срещат едни и същи признаци. Тези признаци генерират многократно едни и същи обобщени линейни признаци и се получават описания с излишък на информация. Ето защо от еднаквите топологични признаци се оставя по един. Получават се междинни описания. Структурните признаци на тези описания се подават на входа на решаващото дърво. Обобщените признаци, съответстващи на един полулинейен елемент, потвърждават наличието на даден линейен признак. Те водят до излишък на информация и затова изходите на решаващото дърво, съответстващи на един полулинейен елемент могат да бъдат забранени. След като се отхвърли излишъка от информация се получават следните обобщени описания:

ВЛХД ВСХД ВДХД.

ВЛХГ ВЛХС ВЛХД.

ВЛХД ВСХД ВДХД.

*ВЛХГ ВЛХС ВЛХД.**

Въведени са следните означения: *В*–вертикала (вертикална линия); *ВЛ*–вертикала лява; *ВС*–вертикала средна; *ВД*–вертикала дясна; *Х*–хоризонтала (хоризонтална линия); *ХГ*–хоризонтала горна; *ХС*–хоризонтала средна; *ХД*–хоризонтала долна. Последните обобщени описания са освободени и от излишна пунктуация. Тези обобщени описания са компактни и не съдържат излишна информация от гледна точка на класификацията на описанията. Те се получават близки до

описанията на естествен език. Например първото описание означава, че буквата “Ш” се състои от

“*хоризонтала долна, свързана с вертикала лява, с
вертикала средна,
с верикала дясна*”

Класификацията на обобщените описания се прави на базата на поелементно сравнение. Обучението се извършва без учител. За представително множество символи за съответните класове се конструират описания. Те се вземат да бъдат като еталони и по нататък се използват за класификация в условни класове. Човек съобщава имената за класовете. Обучението може да се извърши и с учител.

3. Заключение

Получените описания представляват интерес в теоретично отношение с цел конструиране на описания, близки до естествения език и разработка на йерархична система, способна да прави обобщение. Интерпретацията на описанията може да се разглежда като компютърно разбиране на структурата на обектите. Това разбиране може да се засили, ако към линейните елементи се добавяват и техните лингвистични стойности – голям, среден или малък, изчислявани по формули дадени в [3]. Но това ще ограничи инвариантността на описанията по отношение на някои трансформации като мащаб, разтяжение или свиване на линейните елементи на обектите и др.

ЛИТЕРАТУРА

- [1]. Smolensky, P. *Information Processing Dynamical Systems: Foundations of Harmony Theory*. Parallel Distributed Processing. Cambridge, 1988.
- [2]. Gotchev, G. *Computer Linguistic Analysis of Line Drawings*. Pattern Recognition, vol. 17, No. 4, 1984, 433+440.
- [3]. Гочев, Г. *Компютърно зрение и невронни мрежи*. ТУ, София, 1998.

Автор: Гочо В. Гочев, доц. д-р кат. “Програмиране и компютърни технологии”.

Постъпила на 16.09.2003г. Рецензент доц. д-р Иван Момчев