

УДК 681.5

Ю.В. Паржин¹, Д.В. Гринев², В.В. Онищенко²

¹Национальный технический университет «ХПИ», Харьков

²Харьковский университет Воздушных Сил им. И. Кожедуба, Харьков

ПОСТРОЕНИЕ ПОРОЖДАЮЩИХ ГРАММАТИК И КОНЕЧНЫХ АВТОМАТОВ ДЛЯ СЕМАНТИЧЕСКОЙ СЕТИ СТРУКТУРНО-ЛИНГВИСТИЧЕСКИХ КОНЦЕПТОВ В ПРОЦЕССЕ ОБУЧЕНИЯ СИСТЕМЫ РАСПОЗНАВАНИЯ

Предлагается методика построения порождающих грамматик и конечных автоматов для семантической сети структурно-лингвистических концептов в процессе обучения системы распознавания трехмерных объектов, учитывающая различные углы отклонения камеры от положения в надир.

распознавание образов, семантическая сеть, структурно-лингвистический концепт, порождающая грамматика, конечный автомат, классификация, надир, направление съемки

Введение

Типичная задача автоматического распознавания образов представляет собой процесс анализа ис-

следуемых изображений по выбранным признакам и выработки решающего правила, позволяющего на основании заданного критерия отнести каждое изображение к одному из заранее сформированных

классов объектов распознавания [1 – 3]. Для решения такого рода задач необходимо иметь развитый аппарат, ориентированный на иерархическую структуру объектов, опирающийся на систему непроектируемых элементов и отношений между ними. Развитие такого аппарата основывается на аналогии между структурой изображения, подверженной различного рода проективным искажениям и деформационным преобразованиям контура, и структурой фразы в естественном или искусственном языке.

Для уменьшения влияния проективных искажений на результат распознавания, необходимо построить и обучить семантическую сеть эталонных структурно-лингвистических концептов, инвариантных как аффинным, так и деформационным преобразованиям контуров изображений трехмерных объектов [4].

При построении семантической сети структурно-лингвистических концептов необходимо исследовать свойства структурных элементов в процессе моделирования, сформировать пространство признаков, минимизировав объем исходной информации путем выделения минимального числа наиболее представительных для каждого класса эталонов.

Каждый из эталонных фрагментов семантической сети, представляющий собой множество допустимых цепочек непроектируемых элементов (структурно-лингвистических концептов), должен быть представлен своей порождающей грамматикой. На основании сформированных грамматик необходимо синтезировать конечные автоматы, обеспечивающие разделение объектов на классы (подклассы).

Целью данной статьи является разработка методики построения порождающих грамматик и конечных автоматов для семантической сети структурно-лингвистических концептов в процессе обучения при распознавании трехмерных объектов, учитывающая различные углы отклонения камеры от положения в надир и различные направления съемки.

Результаты исследований

Основой для построения языковых моделей сцен и объектов служит аппарат математической лингвистики [5].

Конечное множество правил, определяющих способ образования цепочек языка, называется грамматикой. Порождающей грамматикой будем называть выражение вида:

$$G = (V_N, V_T, P, S),$$

где V_N – множество нетерминальных символов (переменных); V_T – множество терминальных символов (констант); P – конечное множество правил подстановки; S – специальный символ, называемый начальным или аксиомой.

При этом общий алфавит порождающей грамматики $V = V_N \cup V_T$ и $V_N \cap V_T = \emptyset$.

Предположим, что у нас имеются два класса C_1 и C_2 изображений трехмерных объектов. Пусть структурно-лингвистические концепты этих классов могут быть построены из признаков, принадлежащих некоторому конечному множеству терминальных символом V_T . Допустим, что существует такая грамматика G , что порождаемый ею язык состоит из предложений, принадлежащих исключительно одному из классов, скажем C_2 . Очевидно, что эта грамматика может быть использована в целях классификации изображений, так как заданная структура может быть отнесена путем грамматического разбора к C_1 , если она является предложением языка $L(G)$. В противном случае изображение приписывается классу C_2 .

По приведенной выше схеме классификации изображение попадает в класс C_2 исключительно потому, что оно не принадлежит классу C_1 . Если обнаружится, что изображение не является правильным предложением для G , предполагается, что оно должно принадлежать C_2 . Тем не менее, не исключено, что изображение не принадлежит и классу C_2 . Для того, чтобы обеспечить возможность изъятия цепочки, необходимо задать две грамматики G_1 и G_2 , порождающие языки $L(G_1)$ и $L(G_2)$ соответственно. Изображение I зачисляется в класс, язык которого позволяет считать это изображение грамматически правильным предложением. Если обнаружится, что изображение может принадлежать обоим классам, его произвольно зачисляют в любой из этих классов. Если оно не является предложением ни $L(G_1)$, ни $L(G_2)$, изображение изымается.

В случае m классов необходимо рассматривать m порождающих грамматик и связанных с ними языков $L(G_i)$, $i = 1, 2, \dots, m$. Распознаваемый объект относится к классу C_i в том и только том случае, если он является предложением языка $L(G_i)$. Если объект является предложением более чем одного языка или не принадлежит ни одному из них, он может быть изъят из рассмотрения или произвольно отнесен к одному из классов неопределенных объектов.

Выполним моделирование с использованием пакета 3DS Max 7 для самолета Boeing.

Чтобы отследить поведение структурных элементов при вращении самолета, необходимо для каждой из полученных в процессе моделирования проекций построить структурные концепты.

Результаты моделирования были получены для углов α_n и γ_k с шагом дискретизации $\Delta = 10^\circ$. На основании проведенных исследований был сделан вывод о том, что поведение структурных элементов концептов наиболее характерно для 8 основных направлений: справа; спереди_справа; спереди; спереди_слева; слева; сзади_слева; сзади; сзади_справа.

Кроме того, за эталонную проекцию следует взять изображение, полученное при съемке в надир.

Таким образом, для построения семантической сети W структурно-лингвистических концептов [4]

необходимо выполнить моделирование поведения структурных элементов $v_{i,j}$ (где i – порядковый номер структурного элемента; j – направление ориентации структурного элемента) при различных углах α_n отклонения камеры от положения в надир и различных направлениях γ_k камеры по отношению к объекту съемки. За базовое направление γ_1 камеры примем угол $\gamma_1 = 0^\circ$. Значения остальных углов $\gamma_2, \dots, \gamma_8$ будут распределены согласно рис. 1. Направление γ_0 , представляющее собой съемку в надир, на рисунке не показано.



Рис. 1. Соответствие углов и направлений γ_k съемки

Численные значения углов α_n отклонения камеры от положения в надир представлены на рис. 2.

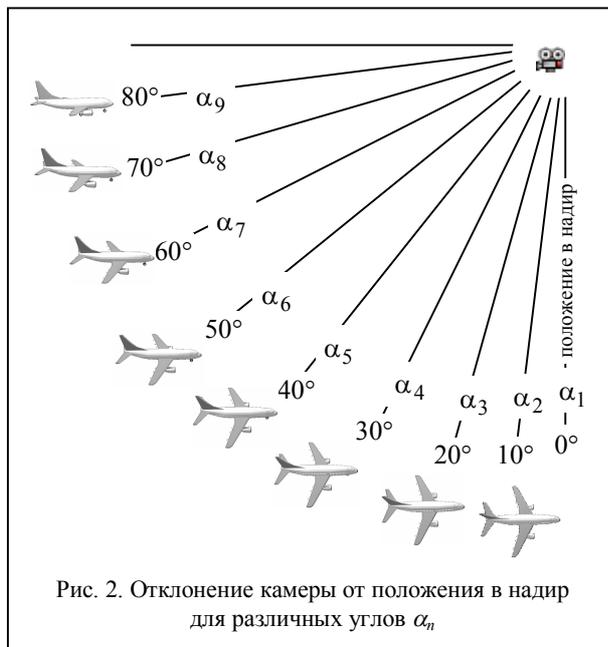


Рис. 2. Отклонение камеры от положения в надир для различных углов α_n

Рассмотрим процесс построения грамматик для направления "справа". Изображения самолета, а также их структурные концепты для данного направления ($\gamma_1 = 0^\circ$) и углов α_n представлены в табл. 1.

Таблица 1

Изображения самолета и их структурные концепты

Угол α_n	Вид справа	
	изображение	концепт
10°		
20°		
30°		
40°		
50°		
60°		
70°		
80°		

Представим после процесса нормализации [6, 7] и выбора системы ориентации структурных эле-

$$R \xrightarrow{1} eA \xrightarrow{2} edB \xrightarrow{5} edgC \xrightarrow{9} edgeD \xrightarrow{10} edgeh.$$

Следовательно, рассмотренная цепочка может порождаться грамматикой $G_R^{\gamma_1}$. Для фрагмента семантической сети, представленного на рис. 5, порождающая грамматика $G_L^{\gamma_1}$ будет иметь вид:

- $V_T = \{b, c, d, e, f\}$; $V_N = \{L, A, B, C, D, E, F\}$; $S = L$;
 P: 1) $L \rightarrow eA$; 6) $C \rightarrow fD$; 11) $E \rightarrow eF$;
 2) $A \rightarrow fB$; 7) $C \rightarrow d$; 12) $F \rightarrow b$.
 3) $B \rightarrow dC$; 8) $D \rightarrow c$;
 4) $B \rightarrow cC$; 9) $D \rightarrow cE$;
 5) $B \rightarrow bC$; 10) $D \rightarrow bE$.

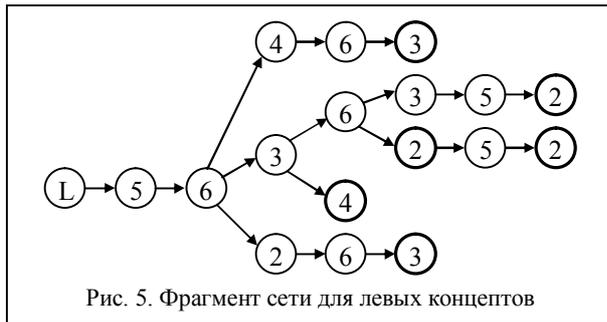


Рис. 5. Фрагмент сети для левых концептов

Подобным образом необходимо построить грамматики для оставшихся направлений γ_k .

После того как определены грамматики, необходимо разработать процедуру, устанавливающую, является или нет данный объект допустимым предложением. При неудачном исходе грамматического разбора заданное изображение отклоняется как представляющее неправильное предложение.

Усовершенствование процесса грамматического разбора связано с применением правил синтаксиса грамматики. Синтаксис определяется как соединение и конкатенация объектов. Синтаксическое правило устанавливает некоторые допустимые (или запрещенные) отношения между объектами.

В этой терминологии грамматика является не более чем множеством синтаксических правил, определяющих допустимые или желательные отношения между производными элементами. Синтаксически ориентированный грамматический анализатор, таким образом, включает в процесс грамматического разбора синтаксис грамматики.

Граматики $G_R^{\gamma_1}$ и $G_L^{\gamma_1}$, рассмотренные ранее, были в основном схемами, порождающими цепочки $LCpt_R(I^{\alpha_n, \gamma_1})$ и $LCpt_L(I^{\alpha_n, \gamma_1})$. Для осуществления дальнейших исследований, рассмотрим некоторые аспекты теории автоматов и введем понятие автомата как системы распознавания цепочек.

Конечный автомат A над алфавитом Σ определяется следующим образом:

$$A = (K, \Sigma, \delta, q_0, F),$$

где K – конечное непустое множество состояний; Σ – конечный входной алфавит; δ – отображение

$K \times \Sigma$ в K ; q_0 – начальное состояние, $q_0 \in K$; F – множество заключительных состояний, $F \subseteq K$.

Построим конечный автомат $A_R^{\gamma_1}$, способный распознавать цепочки символов $LCpt_R(I^{\alpha_n, \gamma_1})$, порождаемые грамматикой $G_R^{\gamma_1}$. При этом входной алфавит будет определяться множеством направлений ориентации производных элементов $\Sigma = \{b, c, d, e, f, g, h\}$. Множество состояний определим в виде $K = \{q_0, q_1, q_2, q_3, q_4, q_5, q_6, q_7, q_8, q_9, q_{10}, q_{11}, q_{12}, q_{13}\}$. При этом множество заключительных состояний $F = \{q_{10}, q_{12}, q_{13}\}$. Исходя из правил подстановки рассматриваемой грамматики $G_R^{\gamma_1}$, сформируем множество отображений δ следующим образом:

- $\delta(q_0, e) = q_1$; $\delta(q_4, e) = q_6$; $\delta(q_8, h) = q_{10}$;
 $\delta(q_1, d) = q_2$; $\delta(q_4, f) = q_7$; $\delta(q_9, h) = q_{10}$;
 $\delta(q_1, c) = q_3$; $\delta(q_5, e) = q_8$; $\delta(q_9, g) = q_{11}$;
 $\delta(q_1, b) = q_4$; $\delta(q_5, d) = q_9$; $\delta(q_{10}, e) = q_{12}$;
 $\delta(q_2, g) = q_5$; $\delta(q_6, d) = q_9$; $\delta(q_{11}, e) = q_{12}$;
 $\delta(q_3, e) = q_6$; $\delta(q_7, d) = q_9$; $\delta(q_{12}, h) = q_{13}$.

Диаграмма состояний этого автомата (рис. 6) состоит из вершин, соответствующих каждому из возможных состояний, и ориентированных ребер, соединяющих взаимно достижимые состояния.

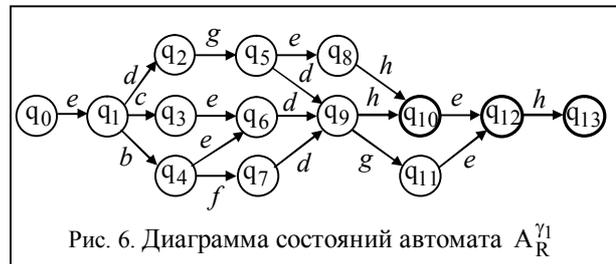


Рис. 6. Диаграмма состояний автомата $A_R^{\gamma_1}$

Каждое ребро диаграммы обозначается соответствующим символом из множества Σ , обуславливающим переход автомата в указанное состояние.

Предположим, что автомат находится в состоянии q_0 и на вход подается цепочка $LCpt_R(I^{\alpha_3, \gamma_1}) = v_{1,5} \wedge v_{2,4} \wedge v_{3,7} \wedge v_{4,5} \wedge v_{5,8}$. Учитывая содержимое табл. 3, входная цепочка будет иметь вид: *edgeh*. Автомат просматривает цепочку слева направо по одному символу за такт. Встретив первый символ *e*, автомат переходит в состояние q_1 . Следующий символ *d* заставляет его перейти в состояние q_2 . После прочтения цепочки $LCpt_R(I^{\alpha_3, \gamma_1})$ автомат будет находиться в одном из конечных состояний q_{10} . Полный цикл работы автомата будет состоять из состояний $q_0, q_1, q_2, q_5, q_8, q_{10}$.

Если в результате просмотра некоторой цепочки автомат будет находиться в одном из возможных заключительных состояний, то она допускается ав-

томатом A . Множество всех цепочек x , допускаемых автоматом A , обозначим $T(A)$, т. е.

$$T(A) = \{x \mid \delta(q, x) \text{ находится в } F\},$$

где $\delta(q, x)$ – состояние автомата после прочтения цепочки x .

Если все изображения представлены множеством цепочек, то конечный автомат можно рассматривать как устройство, обеспечивающее разделение их на классы.

Построим теперь конечный автомат для грамматики $G_L^{\gamma_1}$. Так конечный автомат $A_L^{\gamma_1}$, способный распознавать цепочки символов $LCpt_L(I^{\alpha_n, \gamma_1})$, порождаемые грамматикой $G_L^{\gamma_1}$, будет иметь вид:

$$\Sigma = \{b, c, d, e, f\}; \quad F = \{q_7, q_8, q_9, q_{11}\};$$

$$K = \{q_0, q_1, q_2, q_3, q_4, q_5, q_6, q_7, q_8, q_9, q_{10}, q_{11}\} /$$

.Множество отображений δ определим следующим образом:

$$\delta(q_0, e) = q_1; \quad \delta(q_3, f) = q_6; \quad \delta(q_6, b) = q_9;$$

$$\delta(q_1, f) = q_2; \quad \delta(q_5, f) = q_6; \quad \delta(q_8, e) = q_{10};$$

$$\delta(q_2, d) = q_3; \quad \delta(q_4, f) = q_6; \quad \delta(q_9, e) = q_{10};$$

$$\delta(q_2, b) = q_5; \quad \delta(q_4, d) = q_7; \quad \delta(q_{10}, b) = q_{11};$$

$$\delta(q_2, c) = q_4; \quad \delta(q_6, c) = q_8.$$

Диаграмма состояний автомата представлена на рис. 7.

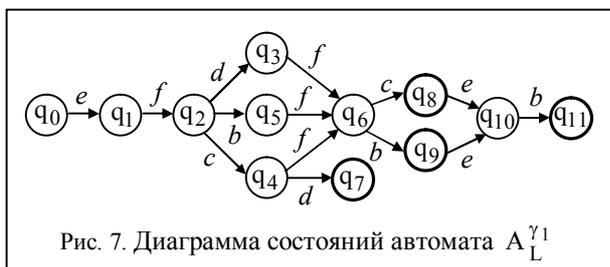


Рис. 7. Диаграмма состояний автомата $A_L^{\gamma_1}$

Полный лингвистический концепт $LCpt(I^{\gamma_1})$ для направления γ_1 будет распознан при условии, что конечные автоматы $A_R^{\gamma_1}$ и $A_L^{\gamma_1}$ будут одновременно находиться в одном из заключительных состояний (табл. 4) для соответствующих углов α_n .

Таблица 4

Таблица состояний

α_n	F для автомата $A_R^{\gamma_1}$	F для автомата $A_L^{\gamma_1}$
10°	q_{10}	q_{11}
20°	q_{10}	q_{11}
30°	q_{10}	q_{11}
40°	q_{10}	q_9
50°	q_{13}	q_8
60°	q_{13}	q_8
70°	q_{13}	q_7
80°	q_{12}	q_7

Заключение

Для решения задачи автоматического распознавания необходимо иметь развитый аппарат, ориентированный на иерархическую структуру объектов, опирающийся на систему производных элементов и отношений между ними. Развитие такого аппарата основывается на аналогии между структурой изображения и структурой фразы в естественном или искусственном языке. Существенно, что эти формальные средства могут быть использованы не только для описания предъявленного изображения, но и для порождения всего класса изображений определяемого выбранными языковыми средствами.

В рамках данной статьи разработана методика построения порождающих грамматик и конечных автоматов для семантической сети структурно-лингвистических концептов в процессе обучения системы при распознавании трехмерных объектов, учитывающая различные углы отклонения камеры от положения в надири и различные направления съемки.

На основании проведенных исследований был сделан вывод о том, что применение разработанной методики позволяет осуществить обучение системы распознавания и обеспечить соотнесение исследуемого изображения к одному из заранее сформированных классов.

Список литературы

1. Форсайт Д., Понс Ж. Компьютерное зрение. Современный подход. – М.: Вильямс, 2004. – 928 с.
2. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2006. – 1072 с.
3. Шапиро Л., Стокман Дж. Компьютерное зрение. – М.: БИНОМ. Лаборатория знаний, 2006. – 752 с.
4. Паржин Ю.В., Гринев Д.В., Онищенко В.В. Побудова семантичної мережі структурно-лінгвістичних концептів об'єктів розпізнавання // Системи управління, навігації та зв'язку – К.: ЦНДІ НіУ, 2007. – Вип. 2. – С. 28-30.
5. Гинзбург С. Математическая теория контекстно-свободных языков. – М.: Мир, 1970. – 326 с.
6. Паржин Ю.В., Гринев Д.В. Нормализация структур контурных изображений распознаваемых объектов // Радіоелектронні і комп'ютерні системи. – 2005. – Вип. 1 (9). – С. 58-62.
7. Паржин Ю.В., Гринев Д.В., Онищенко В.В. Определение оси нормализации в концептуальных структурах контурных изображений с проективными искажениями // Системи обробки інформації. – Х.: XV ПС, 2006. – Вип. 9 (58). – С. 109-111.
8. Гринев Д.В. Классификация и идентификация объектов с использованием структурно-лингвистического метода // Системи обробки інформації. – Х.: ХВУ, 2004. – Вип. 11 (39). – С. 44-48.

Поступила в редколлегию 17.07.2007

Рецензент: д-р техн. наук, проф. В.А. Краснобаев, Харьковский университет Воздушных Сил им. И. Кожедуба, Харьков.