1973 год ВЫЧИСЛИТЕЛЬНЫЕ СИСТЕМЫ Выпуск 55

УДК 51:153:681.3.06

АКТИВНОЕ ВОСПРИЯТИЕ ЗРИТЕЛЬНОЙ ИНФОРМАЦИИ

А.С. Нудельман

Настоящая работа является развитием подхода к изучению зрительного восприятия, описанного в [I].

Кратко изложим некоторые стороны этого подхода. Пусть зрительный анализатор A находится на расстоянии L от картины K. Формальной моделью данной ситуации является пара $\langle \mathcal{Y}, F \rangle$, где $\mathcal{Y}$ – язык описания, сопоставленный анализатору A и расстоянию L, а F – фигура, зависящая от картины K и языка $\mathcal{Y}$. Каждое предложение языка $\mathcal{Y}$ есть утверждение о существовании вполне определенных свойств у рассматриваемой фигуры. Свойствами фигуры могут быть только свойства точек, линий, направлений и углов, то есть таких объектов, выделяемых в фигуре, которые представляются наиболее примитивными. В связи с этим говорится, что изучается именно непосредственное восприятие. Класс $F(S)$ всех фигур, на которых истинно некоторое предложение S языка $\mathcal{Y}$, можно рассматривать как образ и тогда алгоритм проверки истинности S на любой фигуре будет решающим правилом. Выражение "анализатор A воспринял картину K на расстоянии L " имеет следующую формальную интерпретацию: в языке $\mathcal{Y}$ выделено множество всех предложений, истинных на фигуре F. Впредь восприятие анализатором A картины K будем называть п а с с и в н ы м, поскольку в формальной модели истинность любого предложения языка $\mathcal{Y}$ на фигуре F определяется только действительным наличием у этой фигуры утверждаемых предложением свойств.

Рассмотрим предложение S_r [I , стр. 56], которое на обычном языке формулируется следующим образом: "существует ‹вертикальная› и ‹прямая› линия $\mathcal{L}_1(t_1,t_2)$, длина которой ‹равна> высоте фигуры, причем начальная точка t_1 этой линии лежит выше конечной точки t_2 ; существует < горизонтальная > и < прямая > линия $\mathcal{L}_2(t_1, t_5)$ (начальные точки линий $\mathcal{L}_1$ и $\mathcal{L}_2$ совпадают), длина которой < равна > ширине фигуры, причем точка t_5 лежит правее точки t_1 ; длина линии $\mathcal{L}_2$ не < больше > длины линии $\mathcal{L}_1$, но и не < существенно меньше > ; если существует линия $\mathcal{L}_3$, < параллельная > линии $\mathcal{L}_1$, лежащая < правее > линии $\mathcal{L}_1$ и имеющая < значительную > длину, то линия $\mathcal{L}_3$ не заходит < достаточно далеко > вниз". Здесь выражения "<P>" указывают формальные предикаты, определенные в языке $\mathcal{Y}$, ориентировочный смысл которых передан выражениями "P". Нетрудно убедиться, что предложение S_Γ истинно на фигурах $F_1 - F_7$ и ложно на фигурах F_8, F_9, F_{10} (рис. I).

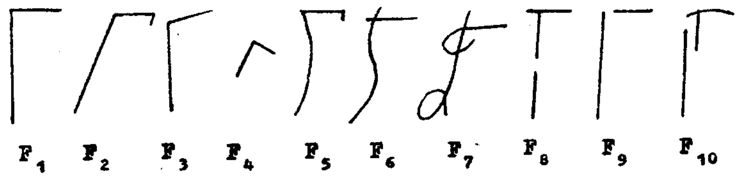

Рис. I

Из практических соображений часто требуется, чтобы все фигуры $F_1 - F_{10}$ принадлежали одному классу фигур, определенным образом порожденному предложением S_Γ , т.е. являлись бы реализациями одного образа (в расширенном смысле). Этот класс будем называть F^α - классом и обозначать через $F^\alpha(S_\Gamma)$. Ниже дается определение F^α - класса для произвольного предложения языка $\mathcal{Y}$. Но сначала - несколько понятий.

Пусть F - фигура [I , § I] и t_1, t_2 - точки, принадлежащие этой фигуре. Тогда отрезок прямой линии $[t_1, t_2]$ назовем д о п у с т и м ы м в F , если и только если для любой точки t фигуры F , отличной от t_1, t_2: $\tau(t, t_1) > \tau(t_1, t_2)$ или $\tau(t, t_2) > \tau(t_1, t_2)$. Здесь $\tau(t', t'')$ - расстояние между точками t' и t'' .

На рис. 2 пунктирными линиями показаны допустимые отрезки к фигурам F_{11} , F_{12} и F_{13} (сплошные линии).

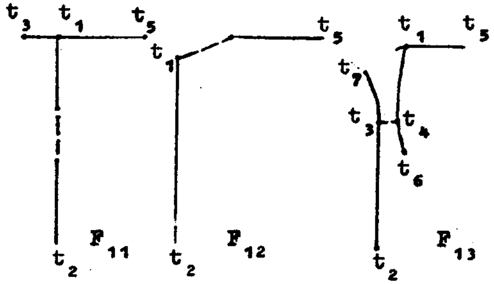

Рис. 2

Фигуру, полученную из F добавлением конечного числа допустимых в F отрезков, назовем д о п о л н е н н о й фигурой F .

Будем говорить, что линия $\mathcal{L}(t_1, t_2)$ имеет направление $\bar{E}$, если направление в любой точке t, $t \neq t_2$ этой линии - $\bar{E}(t)$, $t \in \mathcal{L}(t_1, t_2)$ [I, стр. 4I] - < совпадает > с $\bar{E}$ (как и прежде, <P> - предикат, выразимый в языке $\mathcal{Y}$). Пусть в линии $\mathcal{L}(t_1, \ldots, t_k, t_{k+1}, \ldots, t_n)$, $1 < k < n-1$, подлинии $\mathcal{L}_1(t_1, t_k)$, $\mathcal{L}_2(t_k, t_{k+1})$ и $\mathcal{L}_3(t_{k+1}, t_n)$ имеют направления $\bar{E}, \bar{E}_2$ и $\bar{E}$, соответственно. Тогда если $\bar{E}_2$ < отличается > от $\bar{E}$, то линию $\mathcal{L}_2(t_k, t_{k+1})$ будем называть и з л о м о м , а о линии $\mathcal{L}(t_1, t_n)$ будем говорить, что она является линией н а п р а в л е н и я $\bar{E}$ с и з л о м о м. В общем случае линия может иметь конечное число попарно непересекающихся изломов.

Рассмотрим фигуру F и выделенную в этой фигуре линию $\mathcal{L}(t_1, \ldots, t_k, t_{k+1}, \ldots, t_n)$ направления $\bar{E}$ с изломом (t_k, t_{k+1}) . Пусть линия $\mathcal{L}'(t_k, t')[\mathcal{L}'(t', t_{k+1})]$ такова, что:

I) $\mathcal{L}' \subset F$;

2) $\mathcal{L}' \cap \mathcal{L} = \{t_k\}[\{t_{k+1}\}]$;

3) линия $\mathcal{L}'$ имеет направление $\bar{E}$;

4) расстояние между произвольной точкой из $\mathcal{L}'$ и ближайшей к ней точкой из $\mathcal{L}_1(t_{k+1}, t_n)[\mathcal{L}_1(t_1, t_k)]$, $\mathcal{L}_1 \subset \mathcal{L}$, не < превышает > величины $\tau(t_k, t_{k+1})$.

В этом случае линию $\mathcal{L}'$ назовем линией, с в я з а н н о й с изломом (t_k, t_{k+1}) .

В соответствии с вышеприведенными определениями в дополненной фигуре F_{13} (рис. 2) можно выделить < вертикальную > линию $\mathcal{L}(t_1, t_4, t_3, t_2)$ с изломом (t_4, t_3) . Линии $\mathcal{L}_1(t_7, t_3)$ и $\mathcal{L}_2(t_4, t_6)$ являются связанными с изломом (t_4, t_3) .

Предложение S назовем у с л о в н о и с т и н н ы м на фигуре F , если в качестве хотя бы одной линии, существование которой обеспечило истинность S , выделена линия из F с изломом. Впредь через $M^0(F,S)$, $M^1(F,S)$, $M^2(F,S)$ будем обозначать множество линий из F, дополненной F, дополненной F, существование которых обеспечивает истинность S на F, истинность S на дополненной F, условную истинность S на дополненной F, соответственно. Сформулируем теперь определение F^{α} – класса.

Для любого предложения S языка $\mathscr{Y}$ класс $F^{\alpha}(S)$ есть множество фигур такое, что:

I) если $F \in F(S)$, то $F \in F^{\alpha}(S)$;

2) если дополненная $F \in F(S)$, то $F \in F^{\alpha}(S)$;

3) если S условно истинна на дополненной F и при этом все изломы линий из $M^2(F,S)$ являются допустимыми в F отрезками, то $F \in F^{\alpha}(S)$.

С введением класса $F^{\alpha}(S)$ понятие образа существенно расширяется за счет увеличения числа допустимых нетопологических преобразований. Фигуры класса $F(S)$ тоже имеют допустимые нетопологические преобразования, но все они заключаются только в "стирании" лишних линий (при установлении истинности предложения S_{Γ} на дополненной фигуре F_{11} (рис. 2) линия $\mathscr{L}(t_1,t_3)$ "стирается").

Принцип перестройки фигуры, состоящей в проведении в фигуре новых линий, является первым аспектом (из двух) формальной модели активного восприятия. Второй аспект заключается в сопоставлении каждой фигуре F из класса $F^{\alpha}(S)$ неотрицательного действительного числа $\mathscr{I}(F,S)$, характеризующего "величину" перестройки фигуры F до дополненной F, на которой предложение S истинно или условно истинно. Это число будем называть интенсивностью перестройки.

Обозначим через $\mathscr{D}(M^k)$, $k = 0,1,2$, сумму длин всех линий из $M^k(F,S)$; через $d(M^k)$, $k = 1,2$, – сумму длин всех допустимых в F отрезков, являющихся подлиниями линий из $M^k(F,S)$; через $N(M^2)$ – множество попарно непересекающихся линий, связанных с изломами линий из $M^2(F,S)$, и такое, что линия из $N(M^2)$ является подмножеством F и пересечение линии из $N(M^2)$ с линией из $M^2(F,S)$ либо пусто, либо содержит только одну точку; через $\ell(N(M^2))$ – сумму длин всех линий из $N(M^2)$.

Пусть $F \in F^{\alpha}(S)$. Определим $\mathscr{I}(F,S)$.

Если $F \in F(S)$, то $\mathscr{I}(F,S) = 0$, иначе

$$\mathscr{I}(F,S) = \min\left[\min_{i,j} \frac{d(M_i^2)}{\mathscr{D}(M_i^2) - d(M_i^2) + \ell(N_j(M_i^2))},\ \min_{k} \frac{d(M_k^1)}{\mathscr{D}(M_k^1) - d(M_k^1)}\right].$$

Здесь "числа" вида $\frac{x}{0}$ принимаются равными $+\infty$.

Так, для фигуры F_{13} (рис. 2) интенсивность перестройки $\mathscr{I}(F_{13}, S_{\Gamma})$ будет равна величине

$$\frac{\tau(t_4,t_3)}{\bar{\mathscr{L}}(t_1,t_5) + \bar{\mathscr{L}}(t_1,t_4) + \bar{\mathscr{L}}(t_3,t_2) + \bar{\mathscr{L}}(t_7,t_3) + \bar{\mathscr{L}}(t_4,t_6)},$$

где $\bar{\mathscr{L}}$ – длина линии $\mathscr{L}$.

В заключение рассмотрим пример использования формальной модели активного восприятия в алгоритме распознавания. Исходная ситуация: имеются описания образов O_1 и O_2 – предложения S_1 и S_2, соответственно; имеется фигура F, которую необходимо отнести либо к O_1, либо к O_2. Алгоритм состоит в следующем:

шаг I: если $F \in F^{\alpha}(S_1)$, то шаг 4;
шаг 2: если $F \in F^{\alpha}(S_2)$, то шаг 9;
шаг 3: отказ I; шаг I0;
шаг 4: если $F \notin F^{\alpha}(S_2)$, то шаг 8;
шаг 5: если $\mathscr{I}(F,S_1) < \mathscr{I}(F,S_2)$, то шаг 8;
шаг 6: если $\mathscr{I}(F,S_2) < \mathscr{I}(F,S_1)$, то шаг 9;
шаг 7: отказ 2; шаг I0;
шаг 8: $F \in O_1$; шаг I0;
шаг 9: $F \in O_2$;
шаг I0: конец.

На основе этого алгоритма была составлена программа, имитирующая восприятие цифр 0 и 6, и проведен эксперимент, описанный в [2].

Литература

I. НУДЕЛЬМАН А.С. Об одном формализованном подходе к изучению зрительного восприятия. Настоящий сборник, с.36–78.

2. НУДЕЛЬМАН А.С. Эксперимент по сравнению зрительного "восприятия" машины с восприятием человека. Настоящий сборник, с.84–89.

Поступила в ред.-изд.отд.
I8 апреля I973 года