

УДК 681.3

МОДЕЛИРОВАНИЕ ПРИНЯТИЯ РЕШЕНИЯ В ЗРИТЕЛЬНОЙ СИСТЕМЕ НА ПРИМЕРЕ МЕХАНИЗМА ЗРИТЕЛЬНОГО ВОСПРИЯТИЯ ОТРЕЗКОВ ПРЯМЫХ

В.Г. Калмыков

Институт проблем математических машин и систем

e-mail: vl.kalmykov@gmail.com

В начале 70-х годов прошлого века была предложена, разработана и практически осуществлена система управления, в которой решения принимались с использованием известных на то время принципов функционирования нервной системы человека [Бир, 10].

В современных информационно-технологических системах различного назначения – научно-исследовательских, управляющих – в процессе принятия решения предполагается совместная работа зрительного восприятия человека – оператора и средств машинного зрения.

Современное состояние развития компьютерной науки и техники позволяет рассматривать возможность создания информационных технологий, которые предусматривают использование произвольной, искаженной помехами мультимедийной, в частности, визуальной информации об объектах существенно изменчивой формы. Для оптимального взаимодействия человека – оператора и средств машинного зрения в рамках одной системы должны быть учтены физиологические особенности зрительного восприятия.

Структурная модель полутонового изображения [Калмыков, Вишневский, Власова, 1] дает возможность однообразного по форме представления произвольных изображений. Такие представления хорошо согласовываются со стандартом MPEG-7 [Стандарт, 2], и могут быть достаточно легко к нему адаптированы.

Наибольшую трудность представляют собой задачи приведения к структурной модели произвольных изображений, заданных в растровом виде, искаженных помехами. В общем случае эта задача еще не решена. В числе многих свойств зрительного восприятия особенно удивительной является возможность в течение долей секунды выполнять сегментацию поля зрения на объекты и фон даже при изменениях фона, сравнимых по величине с динамическим диапазоном изображения. Задача сегментации, при попытке решить ее математически обоснованными методами обработки изображений приводит к NP-полным решениям, не реализуемым современными компьютерными средствами [Шлезингер, Главач, 3].

В настоящее время в рамках структурной модели успешно промоделировано выделение контуров полутоновых и бинарных изображений и представление их как последовательностей отрезков цифровых прямых [Вишневский, Калмыков, 4]. Основным элементом контура является отрезок цифровой прямой, поскольку дуги цифровых кривых образованы последовательностями отрезков цифровых прямых. В основу алгоритмов выделения отрезков цифровых прямых положена гипотеза о механизме выделения отрезков произвольных прямых линий в первичной зрительной коре головного мозга.

Предлагаемая гипотеза разработана на основе следующих известных сведений о механизме зрительного восприятия [Хьюбел, 5].

1. Изображение в поле зрения глаза проецируется на сетчатку, образованную зрительными рецепторами, которую можно приближенно рассматривать как некоторое двумерное дискретное пространство.

2. Каждый нейрон поля зрения глаза (*ганглиозная клетка*) собирает информацию от некоторого множества *зрительных рецепторов*, называемого *рецептивным полем*. Среднее расстояние τ между ганглиозными клетками в центральной части поля зрения глаза определяет дискретность поля зрения. Каждой ганглиозной клетке соответствует нейрон первичной зрительной (*стриарной*) коры головного мозга. Будем называть такие нейроны *первичными нейронами*.

3. В стритарной коре обнаружены нейроны, которые возбуждаются (активируются) при предъявлении в определенном месте поля зрения *стимула* – отрезка определенной длины и определенного направления. Длины отрезков, на которые реагируют такие нейроны, соответ-

ствують углу не более $2'$ поля зрення. Будем называть такие нейроны *линейными*. Количество различаемых направлений отрезков, выходящих из одной точки, ограничено, зависит, по-видимому, от количества первичных нейронов, которые могут входить в рецептивные поля линейных нейронов, конфигурации этих рецептивных полей и не превышает 2 – 3 десятков. Естественно предположить, что каждой ганглиозной клетке в центральной части поля зрення соответствует некоторое количество линейных нейронов.

Положим для определенности, что на однотонном фоне в поле зрення расположены стимулы – изображения однотонных пятен, отличающихся по тону от фона. Такие объекты, как известно, полностью определяются их контурами, которые могут быть представлены отрезками прямых линий. Известно, что в стриарной коре изображения пятен определяются сигналами нейронов, соответствующих контурам пятен.

По сигналам линейных нейронов можно судить о том, что в поле зрення глаза имеются стимулы, соответствующие отрезкам прямых некоторых фиксированных направлений. Чем длиннее отрезок прямой, тем больше линейных нейронов данного направления, рецептивные поля которых возбуждаются стимулами, выдают сигналы. Существуют также сложные нейроны, рецептивные поля которых образованы линейными нейронами, которые и способны выдать сигнал о наличии в поле зрення отрезков произвольной длины фиксированных направлений.

Более подробно рассмотрим вопрос о механизме различения произвольного направления отрезков прямых.

Общеизвестно, что зрение человека различает отрезки прямых очень большого количества направлений. Отличия между направлениями двух произвольных отрезков достаточно большой длины могут измеряться минутами и, даже, секундами, особенно если в поле зрення находятся несколько отрезков. Таким образом, количество возможных направлений реально различаемых прямых линий, выходящих из одной точки, гораздо больше того, которое соответствует экспериментально обнаруженным для некоторой точки поля зрення линейным нейронам, соответствующим фиксированным направлениям. С другой стороны, трудно представить какое огромное количество нейронов могло бы понадобиться, для отображения любого возможного реально различаемого направления в каждой точке поля зрення. В таком случае должен существовать специальный зрительный механизм, который реализован в виде связей между нейронами, и который обеспечивает выделение и различение отрезков прямых линий произвольной длины и произвольного направления при разумных затратах нейронов.

Естественно предположить, что сигналы линейных нейронов об отрезках фиксированных направлений используются для построения образа отрезка прямой произвольной длины и произвольного направления. Самым простым случаем обнаружения отрезка является такой, что направление исходного отрезка в поле зрення совпадает с одним из фиксированных направлений линейных нейронов. Если направление некоторого отрезка прямой произвольной длины в поле зрення не совпадает полностью ни с одним из тех фиксированных направлений, которые соответствуют линейным нейронам (то есть это направление является промежуточным между двумя соседними фиксированными направлениями), то такой отрезок не может быть отображен сигналами линейных нейронов, которые соответствуют продолжающим друг друга отрезкам ограниченной длины одного фиксированного направления. В то же время отдельные линейные

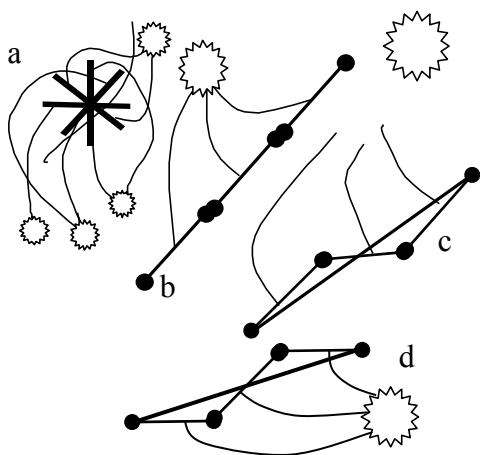


Рис.1. Отображение отрезков прямых произвольного направления (b,c,d) посредством нейронов, возбуждающихся отрезками фиксированных направлений и длины (a).

нейроны, имеющие направление ближайшее к направлению рассматриваемого в поле зрения отрезка выдают соответствующие сигналы. Значит, есть возможность продолжить последовательность отрезком другого ближайшего фиксированного направления (Рис.1). Таким образом, для создания образа отрезка произвольных направления и длины необходимо, чтобы принимались решения об отрезке прямой, если возникает цепочка возбужденных нейронов, которые соответствуют продолжающим друг друга отрезкам фиксированной длины “соседних” направлений (Рис.2,b,c,d). То есть всякий наблюдаемый отрезок прямой соответствует дискретной последовательности фиксированных направлений.

Предлагаемая гипотеза демонстрирует один из возможных механизмов отображения и распознавания отрезков прямых произвольных длин и направлений средствами, характерными для нейрофизиологии и может быть использована в качестве рабочей гипотезы при экспериментальных исследованиях зрительного восприятия. Однако данная гипотеза не дает ответа на следующие вопросы. Каким должен быть порядок расположения отрезков фиксированной длины в последовательности, чтобы эта последовательность воспринималась как отрезок прямой? Будет ли восприниматься линия как прямая, если последовательность будет образована, например, двумя отрезками направления «1», затем двумя отрезками “соседнего” направления «2», затем снова двумя отрезками направления «1»?

Рассмотрим ту же задачу выделения отрезков прямых в процессе структурного анализа контуров объектов бинарного изображения с использованием основных положений рассмотренной гипотезы.

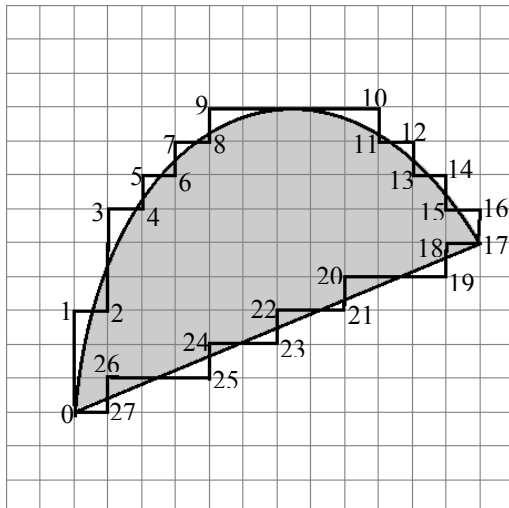


Рис. 2. Пример дискретизации контура

Отрезки цифровых прямых образуются при дискретизации изображений, содержащих контуры, состоящие из отрезков прямых и дуг кривых. Пиксели изображения имеют форму квадрата, образованного соседними парами вертикальных и горизонтальных линий решетки, использованной при дискретизации изображений.

Будем рассматривать дискретное изображение, как двумерный клеточный комплекс [Ковалевский,6]. В данном случае двумерные элементы – это пиксели. Помимо пикселей имеются *креки* (*crac*) и точки.

Креки – это стороны пикселей, являющиеся одномерными элементами. Точки являются конечными точками креков и угловыми точками пикселей. Контур объекта в таком случае

– это связанная замкнутая последовательность *контурных* креков, граничных между пикселями объекта и фоном. Из каждой точки можно провести 4 крека. Креки соответствуют отрезкам фиксированных направлений зрительной системы. На рис. 2 приведен пример исходного контура объекта, а также его цифровой эквивалент как последовательности креков. Связные части последовательности креков можно объединить в *L-элементы*.

Под L-элементом будем понимать связанную последовательность креков вертикальной (горизонтальной) ориентации, которая содержит не более одного крека горизонтальной (вертикальной) ориентации. Смысл использования L-элементов при обработке визуальной информации, в частности, в процессе структурного анализа контуров заключается в том, что L-элементы являются минимально возможными структурными объектами, еще сохраняющими после дискретизации пространственную ориентацию исходного отрезка прямой.

Определить отрезок цифровой прямой в последовательности L-элементов контура – значит указать координаты его начальной и конечной точек. Пусть (x_1, y_1) , (x_2, y_2) – целочисленные координаты начальной и конечной точек отрезка прямой. Угол наклона отрезка определяется отношением разностей его координат $n = \Delta x = |x_1 - x_2|$ и $m = \Delta y = |y_1 - y_2|$, которое в общем

случае не является целым числом. Положим для определенности, что $n > m$. Очевидно, что отображение отрезка прямой произвольного угла наклона посредством L-элементов, имеющих целочисленные значения длины l , осуществляется путем чередования L-элементов, длины которых отличаются на 1, например, равны соответственно $l, l+1$, причем $l \leq n/m \leq l+1$. Порядок их чередования определяет структуру отрезка цифровой прямой и определяется значениями членов цепной дроби $[l; k_1, k_2, \dots, k_t]$ или

$$\frac{n}{m} = l + \frac{r}{m} = l + \frac{1}{k_1 + \frac{r_1}{r}} = \dots = l + \frac{1}{k_1 + \frac{1}{k_2 + \frac{r_2}{r_1}}} = \dots = l + \frac{1}{k_1 + \frac{1}{k_2 + \frac{1}{k_3 + \dots + \frac{1}{k_t}}}}, \quad (1)$$

получаемой при делении целых чисел n и m . Для определенности и без нарушения общности будем полагать, что $r < m - r$, $r_1 < r - r_1$, $r_2 < r_1 - r_2$, ... и т.д.

k_1 подряд идущих L-элементов длины l и один L-элемент длины $l+1$, если он имеется в последовательности, образуют K_1 -элемент, состоящий из k_1 L-элементов, иначе говоря, K_1 -элемент "длины" k_1 . Аналогично, k_2 подряд идущих K_1 -элементов длины k_1 и один K_1 -элемент длины k_1+1 , если он имеется в последовательности, образуют K_2 -элемент длины k_2 и так далее до исчерпания членов цепной дроби. Числитель r определяет количество L-элементов

длины $l+1$ в данном отрезке цифровой прямой, а также и количество K_1 -элементов, каждый из которых содержит k_1 или k_1+1 L-элементов, один из которых имеет длину $l+1$. В свою очередь числитель r_1 определяет количество K_2 -элементов

длины k_1+1 в данном отрезке цифро-

вой прямой, а также и количество K_1 -элементов, каждый из которых содержит k_1 или k_1+1 L-элементов, один из которых имеет длину $l+1$, а остальные – длину l . Вообще числитель r_{t-1} определяет количество K_{t-1} -элементов длины $k_{t-1}+1$ в данном отрезке цифровой прямой, а также и количество K_t -элементов, каждый из которых содержит k_{t-1} или $k_{t-1}+1$ K_{t-1} -элементов, один из которых имеет длину $k_{t-1}+1$, а остальные – длину k_{t-1} .

K_t -элемент, содержащий в своем составе среди K_{t-1} -элементов длины k_{t-1} (L-элементов длины l) один K_{t-1} -элемент длины $k_{t-1}+1$ (один L-элемент длины $l+1$), называется завершенным.

Под отрезком цифровой прямой с координатами начальной и конечной точек (x_1, y_1) , (x_2, y_2) будем понимать последовательность L-элементов, имеющих одинаковые направления g, q , целочисленные длины которых равны соответственно $l, l+1$, а $l \leq n/m \leq$

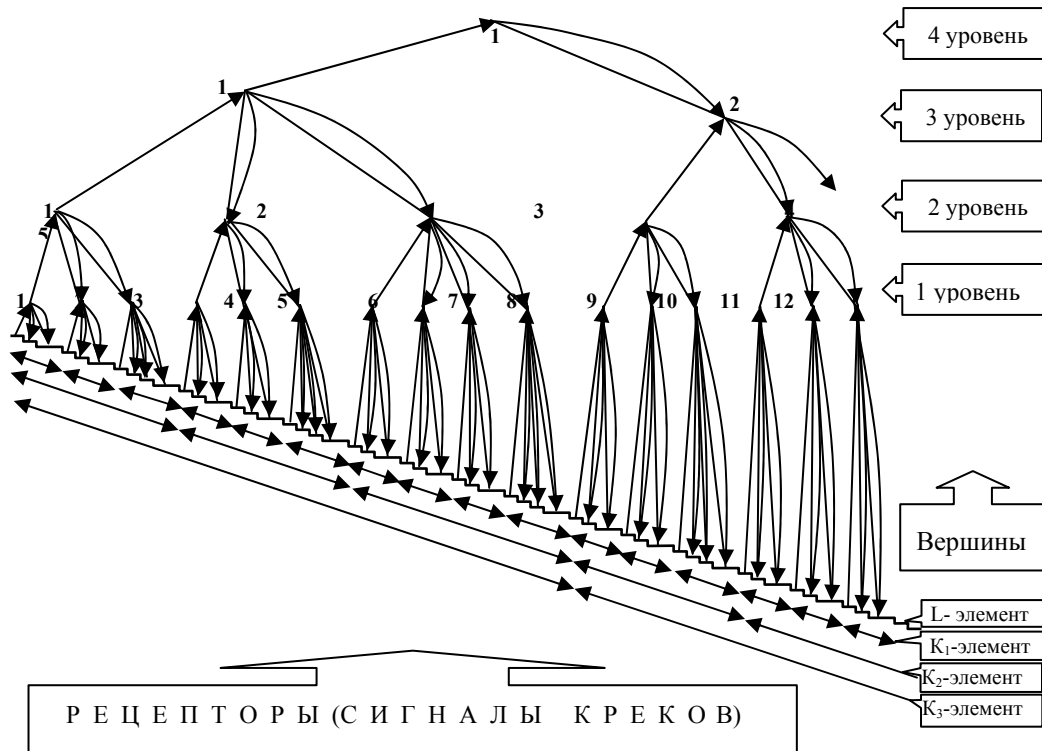


Рис. 3. Формирование образа отрезка цифровой прямой

$l+1$, где $n = |x_1-x_2|$ и $m = |y_1-y_2|$, причем порядок чередования L -элементов разных длин определяется цепной дробью n/m .

Процесс формирования образа отрезка цифровой прямой упомянутым алгоритмом можно представить в терминах растущих пирамидальных сетей [Гладун,7] (рис.3) с учетом работы [Рабинович,8]. Рассматривается частный случай растущей пирамидальной сети, при котором пирамиды вершин не имеют пересечений. В качестве рецепторов используются сигналы креков, которые формируют L -элементы [Власова, Калмыков,9]. Множество вершин сети связаны концепторами и обратными связями. Концепторы представлены восходящими прямыми стрелками и линиями без стрелок. Обратные связи представлены нисходящими стрелками. Рассматривается последовательная процедура процесса формирования образа – слева направо.

Выводы

1. Рассмотрение полученных в нейрофизиологии данных о нейронах, возбуждающихся стимулами в виде отрезков фиксированной длины, позволило сформулировать гипотезу о принятии решения, то есть о формировании образа отрезков произвольной длины и направления.

2. Моделирование процесса принятия решения о выделении в последовательности креков отрезков произвольной длины позволило определить сложную структуру отрезка прямой и сложную, иерархическую процедуру принятия решения даже в простом случае выделения отрезка прямой. Возникает вопрос об обнаружении похожих структур и механизмов в зрительной системе.

3. Использование методов обработки изображений в информационных технологиях сходных с механизмами зрительного восприятия человека позволит повысить качество принятия решения.

Литература

1. Калмыков В.Г., Вишневский В.В., Власова Т.М. Структурная модель полутонного изображения и ее использование в задаче сегментации изображений. // International Journal "Information Technologies and Knowledge". – Sofia. – 2008. - v. 2, №5 – P. 153.
2. MPEG-7 Overview <http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm>
3. М.Шлезингер, В.Главач Десять лекций по статистическому и структурному распознаванию. – Київ: Наукова думка, 2004. – 535 с.
4. Вишневский В.В., Калмыков В.Г. Структурный анализ цифровых контуров изображений как последовательностей отрезков прямых и дуг кривых. // Искусственный интеллект 4'2004, Институт проблем искусственного интеллекта НАНУ – Донецк: "Наука и образование" 2004 С.450-457.
5. Хьюбел Д. "Глаз, мозг, зрение." – Москва: Мир, 1990 - 239 с.
6. Kovalevsky V.A. Applications of Digital Straight Segments to Economical Image Encoding, In Proceedings of the 7th International Workshop, DGCI'97, Montpellier, France, December 3-5, 1997, Springer 1997, P. 51-62.
7. Гладун В.П. Планирование решений. – Киев: Наукова думка, 1987. – 167с.
8. Рабинович З.Л. О естественных механизмах мышления и интеллектуальных ЭВМ // Кибернетика и системный анализ. – 2003. -№5 - С.82-88.
9. Власова Т.М., Калмыков В.Г. Алгоритм и программа распознавания контуров изображений как последовательности отрезков цифровых прямых // Математичні машини і системи. – 2005. -№4 - С.84-95.
10. Бир С. "Мозг фирмы." Москва: Едиториал УРСС, 2005 – 416.