

УДК 004932+621.396

ПРИНЯТИЕ РЕШЕНИЙ ОБ АКТУАЛЬНЫХ ЗАДАЧАХ ТЕХНИЧЕСКОГО ЗРЕНИЯ ПУТЕМ СРАВНИТЕЛЬНОГО АНАЛИЗА МЕХАНИЗМОВ ЗРИТЕЛЬНОГО ВОСПРИЯТИЯ И АЛГОРИТМОВ ОБРАБОТКИ ВИЗУАЛЬНОЙ ИНФОРМАЦИИ

В.Г. Калмыков

Институт проблем математических машин и систем НАН Украины

e-mail: vl.kalmykov@gmail.com

Совместный анализ результатов исследований зрительного восприятия в области нейрофизиологии и обработки визуальной информации в области информатики позволяет:

– определить перспективные для практического использования пути исследований зрительного восприятия;

– определить перспективные направления развития систем технического зрения.

Рассмотрены некоторые известные механизмы зрительного восприятия и соответствующие им методы обработки визуальной информации.

1. Механизм зрительного восприятия границ объектов и метод Канни

В нейрофизиологии исследован механизм определения элементов границ объектов [1]. Основой этого механизма является наличие нейронов, которые возбуждаются, если на их рецептивное поле воздействуют сигналы изображения, которые содержат элемент границы объекта определенного направления рис. 1, 2.

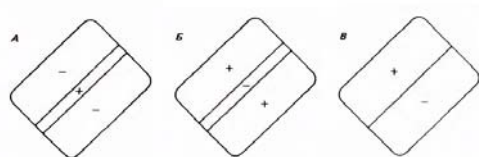


Рис. 1. Типичные рецептивные поля нейронов стриарной коры: зона возбуждения обозначена +, зона торможения –; А – белая линия на черном фоне; Б – черная линия на белом фоне; В – граница черного и белого – контур

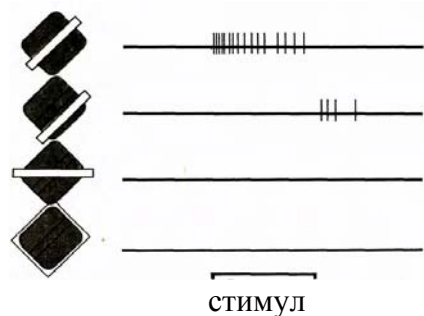


Рис. 2. Реакция нейрона с рецептивным полем типа А, предъявление стимулов сверху вниз:

1 – оптимальный стимул; 2 – стимул совпал с зоной торможения – слабый отклик после окончания действия стимула; 3 – несовпадение направления – действие на зоны возбуждения и торможения одинакова – импульсов нет; 4 – нет стимула – импульсов нет

Фильтр Собела			
Направление	Маска 3x3		
вертикальное	-1	0	1
	-2	0	2
	-1	0	1
горизонтальное	1	2	1
	0	0	0
	-1	-2	-1

размытия изображения, используя фильтр Гаусса для снижения уровня помех, исключения нежелательных деталей и элементов текстуры изображения

$$g(m,n)=G_{\sigma} * v(m,n),$$

Механизму зрительного восприятия элементов границ объектов в информатике соответствует метод Канни [2] – метод выделения граничных элементов объектов изображения по скачкообразному изменению градиента яркости.

Метод Канни заключается в следующем.

На исходном изображении $V=\{v(m,n)|m=1,M;n=1,N\}$ выполняют операцию

где G_σ – фильтр Гаусса для значения дисперсии σ , $g(m,n)$ – элемент "размытого" изображения V_g .

Вычисляют частные значения градиентов для горизонтального – $g_m(m,n)$ и вертикального – $g_n(m,n)$ направлений на размытом изображении $g(m,n)$, используя, например, оператор Собела, чтобы получить значения общего градиента $M(m,n)$ и его направления $\theta(m,n)$ как

$$M(m,n) = \sqrt{g_m^2(m,n) + g_n^2(m,n)} \quad \theta(m,n) = \arctg(g_n(m,n)/g_m(m,n))$$

По значениям $M(m,n)$ с использованием порога T , который должен быть подобран так, чтобы все контурные элементы были выделены, в то время как большинство помех были бы подавлены, получают значения $M_T(m,n)$:

$$M_T(m,n) = \begin{cases} M(m,n), & \text{если } M(m,n) > T \\ 0, & \text{иначе} \end{cases}$$

Все ненулевые элементы объединяют, чтобы получить замкнутый контур объекта, для чего используют специальный алгоритм. Пример градиентного фильтра приведен в таблице. Пример работы фильтра Канни приведен на рис. 3, 4.



Рис. 3. Исходное изображение

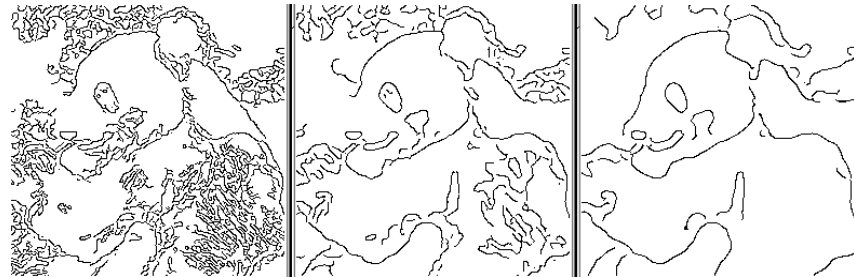


Рис. 4. Выделение границ методом Канни при значениях $\sigma = 1, 2, 3$

Сравнительный анализ метода Канни с механизмом зрительного восприятия элементов контура объекта позволяет установить следующее.

В механизме зрительного восприятия элементы контура объекта изображения определяются путем возбуждения нейрона, соответствующему элементу контура объекта. Этот элемент контура представляет собой отрезок линии, не имеющей толщины. Для метода Канни в качестве граничного элемента контура используется пиксел – элемент изображения, имеющий реальные размеры. При этом контур – граница объекта по методу Канни является линией, имеющей толщину, что может вносить серьезные искажения при обработке изображения. Использование двумерных клеточных комплексов [3] при описании изображения позволяет устранить этот недостаток.

Таким образом, совместное рассмотрение механизма зрительного восприятия и метода обработки изображения позволяет, в данном случае, усовершенствовать метод обработки изображения.

2. Динамическое изменение разрешающей способности зрительной системы

Зрительная система человека работает в режиме следующих друг за другом зрительных актов. Зрительный акт длится примерно 150 мс. После чего происходит нистагм – подергивание глазодвигательных мышц, в результате чего изображение на сетчатке смещается и начинается следующий зрительный акт. В результате исследований размеров зон возбуждения рецептивных полей было установлено [4], что размер их не остается постоянным в течение зрительного акта.

Исключительную важность использования переменной разрешающей способности в системах технического зрения можно показать на следующем простом примере распознавания текстов (рис. 1) с использованием программы Finereader. Выбрана не имеющая смысла последовательность знаков A65H7, чтобы исключить влияние использования словарей на

результат розпознавання. Результат розпознавання исходных изображений: рис. 5а – успешно, рис. 5б – отказ от распознавания по причине невозможности найти (определить) объект в поле изображения.



Рис. 5. Изображения текста: дискретизация выполнена при а) низком и б) высоком разрешении

Причиной отказа от распознавания для текста 5б является наличие большого количества текстурных деталей, что не предусмотрено алгоритмами, используемыми в программе. Вообще распознавание текстур является существенно более сложной задачей, чем распознавание текстов. При низком разрешении детали, мешающие распознаванию, утрачиваются, чем и объясняется успешная работа программы.

Можно сделать вывод, что зрительное восприятие визуальной информации происходит при разных значениях разрешающей способности. В частности, при малом разрешении обрабатываются крупные детали, в то время как при максимальном разрешении различаются все подробности объектов интереса. Таким образом, использование переменной разрешающей способности в системах технического зрения должно привести к получению качественно новых возможностей этих систем.

3. Модель функционирования нейрона в процессе генерации потенциалов действия

Модель работы нейрона, которая объясняет и учитывает изменения размеров зон возбуждения рецептивных полей в течение зрительного акта предложена автором в процессе системологического анализа известных представлений о функционировании нейрона. При разработке модели учитывалось влияние хлорной проводимости в процессе генерации потенциала действия в течение всего зрительного акта.

Однако, оставался открытым вопрос о механизме изменения размеров зон возбуждения рецептивных полей. Модель работы нейрона, которая объясняет и учитывает изменения размеров зон возбуждения рецептивных полей в течение зрительного акта разработана в процессе системологического анализа известных представлений о функционировании нейрона. При разработке модели учитывалось влияние хлорной проводимости в процессе генерации потенциала действия в течение всего зрительного акта.

Под воздействием постсинаптического потенциала возбуждающих рецепторов, (а также при отсутствии воздействия тормозных рецепторов) начинают открываться потенциалзависимые натриевые каналы, нарастает натриевый ток и увеличивается деполяризация, в результате чего открывается еще больше натриевых каналов и внутриклеточный потенциал мембраны возрастает вплоть до величины натриевого равновесного потенциала.

Во время нарастания натриевого тока и увеличивающейся деполяризации, ионы хлора поступают через поры по концентрационному градиенту и при уменьшении противодействующего электрического потенциала. Кроме того, наступает момент (возможно, это момент максимального значения потенциала), когда не могут не открыться потенциалзависимые хлорные каналы, которые также пропускают внутрь клетки ионы хлора. После достижения максимального значения уровня деполяризации начинается процесс выхода ионов калия и натрия, которые возвращают потенциал мембраны к исходному состоянию. Однако за счет поступления ионов хлора потенциал покоя при генерации каждого следующего импульса смещается на некоторую величину в сторону гиперполяризации (рис.6).

Порог открывания, в частности, потенциалзависимых натриевых каналов для различных участков мембраны неодинаков [4]. В области аксонного холмика нервной клетки

порог их открывания составляет 10 мВ, в удаленных от аксонного холмика – 20 –30 мВ).

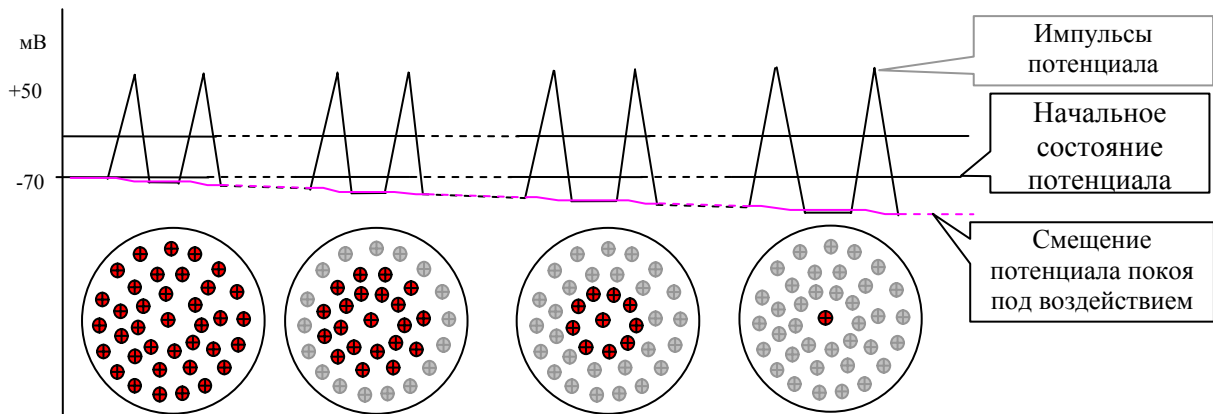


Рис. 6. Изменение зоны возбуждения рецептивного поля

Итак, после каждого очередного (n-го) импульса потенциала действия происходит смещение потенциала покоя в сторону гиперполяризации. В результате в начале следующего (n+1-го) импульса потенциалзависимые натриевые каналы, которые имеют наибольший порог открывания (наиболее удалены от аксонного холмика), перестают открываться под воздействием постсинаптического потенциала возбуждающих рецепторов, находящихся в зоне расположения этих натриевых каналов. Это значит, что потенциал порога открывания упомянутых натриевых каналов больше, чем исходный потенциал покоя плюс смещение. То есть некоторая часть потенциалзависимых натриевых каналов уже не могут участвовать в накоплении заряда для генерации n+1-го импульса.

Можно также предположить, что расстояние возбуждающего рецептора от аксонного холмика соответствует расстоянию в поле зрения от соответствующей данному рецептору точки до центра рецептивного поля. Так что «неучастие» некоторого количества возбуждающих рецепторов в процессе генерации соответствует уменьшению зоны возбуждения рецептивного поля. Таким образом, в процессе зрительного восприятия, точнее, одного зрительного акта, в зрительной системе существует последовательно информация о наблюдаемом изображении с различной разрешающей способностью, вначале низкой, затем все возрастающей и достигающей максимального значения к концу зрительного акта.

Выводы. Сопоставительный системный анализ результатов исследований зрительного восприятия в области нейрофизиологии и обработки визуальной информации в области информатики позволил:

- определить новое направление развития систем технического зрения с переменной разрешающей способностью;
- предложить новую модель функционирования нейрона, учитывающую изменение размеров зон возбуждения рецептивных полей;
- усовершенствовать существующие методы обработки изображений, например, метод Канни.

Список литературы

1. David H. Hubel Eye, brain, and vision New York : Scientific American Library : Distributed by W.H. Freeman, 1988. – 240 p. : ill.
2. Павлидис Т. Алгоритмы машинной графики и обработки изображений. М Радио и связь 1986г. 400с
3. 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.
4. Н.Ф.Подвигин Динамические свойства нейронных структур зрительной системы Ленинград: Наука, 1979. – 158 с. : илл.