

СПИСОК ЛИТЕРАТУРЫ

1. Буч Г. Объектно-ориентированный анализ и проектирование с примерами приложений на C++. – М.: Бином; СПб.: Невский диалект, 2001. – 560 с.
2. Касьянов В.Н., Поттосин И.В. Системы конкретизации: подход и основные понятия // Препринт ВЦ СО АН СССР, № 349. – Новосибирск, 1982. – 22 с.
3. Burstall R.M., Darlington J. A Transformation system for developing Recursive Programs. // Journal of the ACM. – 1977. – V. 24. – P. 44–67.
4. Цейтин Г.С. Является ли математика частью информатики? // Компьютерные инструменты в образовании. – 1999. – № 5. – С. 3–7.
5. Непейвода Н.Н. Математик и прикладник: о взаимно(не)понимании // Труды Удмуртского государственного университета. – 2005. – № 7. – С. 23–31.
6. Новосельцев В.Б. Теория структурных функциональных моделей // Сибирский математический журнал. – 2006. – Т. 47. – № 5. – С. 1014–1030.
7. Тейз А., Грибомон П. и др. Логический подход к искусственному интеллекту. Т. 1. От классической логики к логическому программированию. – М.: Мир, 1990. – 322 с.
8. Вирт Н. Структурное программирование. – М.: Мир, 1975. – 189 с.
9. Клини С. Математическая логика. – М.: Мир, 1973. – 480 с.
10. Чень Ч., Ли Р. Математическая логика и автоматическое доказательство теорем. – М.: Наука, 1983. – 360 с.

Поступила 14.04.2008 г.

Ключевые слова:

Логический вывод, полнота теории, пропозициональное исчисление, разрешающий алгоритм, метод резолюций, обратный метод.

УДК 004.89

МЕХАНИЗМЫ ВЫДЕЛЕНИЯ КОНТРАСТНЫХ ФРАГМЕНТОВ СЦЕНЫ СИСТЕМОЙ ВОСХОДЯЩЕГО ВНИМАНИЯ

С.В. Аксёнов

Томский политехнический университет
E-mail: asv@osu.cctpu.edu.ru

Система восходящего внимания анализирует фрагменты зрительных сцен, выделяя наиболее контрастные и информативные области. В работе предлагается подход, расширяющий модель восходящего внимания Л. Итти с целью учета более эффективного противостояния между цветовыми компонентами и снижения влияния к трансформациям и ширине локальных характеристик сцены.

Когда человеческий глаз рассматривает окружающую среду, далеко не все её компоненты воспринимаются одинаково. Некоторые фрагменты автоматически выделяются из своего окружения, и зрительная система настраивается на их интерпретацию. В этом случае говорят, что эти области были локализованы фокусом внимания. Подчеркнем, что данный процесс идет не параллельно, а предшествует распознаванию образов.

В самом общем случае выделение фрагментов в сцене происходит в результате взаимодействия двух потоков внимания. Первый, называемый восходящим, представляет распространение информации от низших отделов зрительной системы (рецепторы) к высшим (корковые структуры). Его задача состоит в выявлении наиболее контрастных областей сцены. Второй, или нисходящий, поток анализа протекает в обратном направлении (от высших до низших корковых отделов) и он служит для локализации фрагментов, синтезируемых источником. В работе рассматривается модель, описывающая работу первого потока, выявляющего наиболее информативные и отличающиеся от общего фона области, попавшие на зрительную сцену.

Основные принципы функционирования современного понимания системы восходящего внимания были заложены К. Кохом и С. Ульманом [1]. Согласно их представлению, весь объем зрительной информации анализируется на определенных цветовых представлениях сцены. Выявление значимых характеристик фрагментов сцены (как, например, цвет или интенсивность) производится двумерными областями нейронов, настроенными на определенный сигнал. Эти области также называются картами признаков. После нахождения активностей всех карт, результаты всех вычислений комбинируются в одном двумерном массиве клеток, называемом картой особенностей. Расположение наиболее активных клеток карты особенностей, найденное согласно процедуре WTA (Winner-Takes-All), позволяет локализовать наиболее информативные области. Другими словами, положение наиболее активной клетки указывает на фрагмент сцены, который попадет в высшие области (центральное представление) для последующей обработки в первую очередь. Далее, выделяются области сцены, ассоциированные со вторым по активности нейроном карты особенностей и т. д. Таким образом, фокус внимания будет переходить от наиболее контрастных сегментов сцены к менее «интересным». Модель такого взаимодействия дана на рис. 1.

Развитие этой парадигмы продолжилось в работах Л. Итти [2–4], предположившего о существовании карт признаков на нескольких уровнях иерархии. К значимым характеристикам сцены относились 4 цветовых оттенка: красный, жёлтый, зелёный, синий, значение интенсивности цвета и локальные ориентации (0, 45, 90 и 135°). Несмотря на результаты, полученные при использовании этой модели, при решении ряда теоретических и практических задач [5] были выявлены и определенные недостатки (о чем упоминал сам разработчик). При значительном упрощении процесса выделения контрастных фрагментов, в расчет не брались взаимодействия, происходящие между простыми и сложными клетками [6, 7] зрительной системы. Также цветовая схема модели не отражала те возможности восприятия, которыми обладают биологические аналоги.

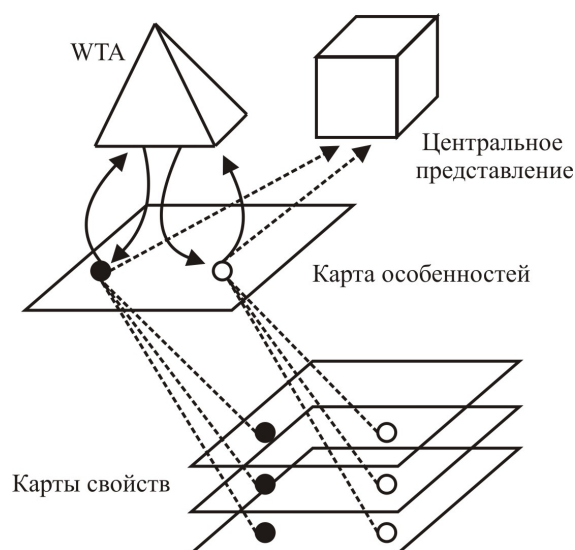


Рис. 1. Общая структура системы восходящего внимания

Известно что, зрительная система воспринимает все цвета неодинаково. Для одних цветов зрение может отличать больше оттенков, чем для других. Для решения проблемы восприятия цветов предложен вариант перехода к цветовому пространству Lu^*v^* [8]. Модель Lu^*v^* оценивает расстояние между цветами согласно процедуре, наиболее близкой к восприятию человеческого глаза. За счет этого перехода локализация областей, значительно отличающихся от фона, происходит в биологически правдоподобной форме.

Если для анализа цветового потока достаточно изменения цветового пространства, то уменьшение чувствительности к трансформациям локальных ориентаций протекает более сложным образом.

Исследования, проводившиеся над нижневисочной корой головного мозга у макак, показали наличие особых клеток зрительного пути, способных настраиваться на определенные сложные объекты [9–11]. Опираясь на эти работы, М. Райзенхубер и Т. Поггио создали иерархическую искусственную нейронную сеть с операцией максимум

(оригинальное название — HMAX) и расширили её возможности в серии работ [12, 13]. Структура модели отражает предположение, что определение свойств, с одной стороны, и инвариантность к положению и масштабу, с другой стороны, может быть выполнена путем процессов разделения объекта на части с последующим усложнением характеристик на более высоких уровнях иерархии. Для сокращения влияния трансформаций локальных ориентаций используем взаимодействие между простыми и сложными клетками в сети HMAX.

Простые клетки вычисляют взвешенную сумму своих входов и служат для выделения свойств объекта (аналог линейных фильтров). Задача сложных клеток состоит в группировании простых клеток, разделяющих несколько параметров (таких как, например, наклон линий — границ объекта), но различных по одному или большему числу параметров (например, ширина линий, длина). Сложные нейроны производят поиск максимальной активности простых клеток, входящих в их рецептивную область. В результате сложные клетки инвариантны ко всем параметрам, анализируемым ассоциированными простыми клетками. На рис. 2 показаны рецептивные поля простых клеток, моделирующие дифференциальный гауссовый фильтр разной ширины, но одинакового наклона и ассоциированные с ними сложные нейроны, указывающие на максимальную активность заданной характеристики.

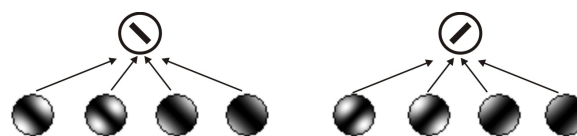


Рис. 2. Операция MAX

Почему предлагается именно этот механизм? В случае модели нейрона, осуществляющего взвешенное суммирование, выход сложной клетки будет инвариантным в течение всего времени, пока стимул попадает в рецептивное поле этой клетки. Однако значение выхода не позволяет судить о существовании линии требуемой ориентации на рецептивном поле, так как выход клетки представляет собой взвешенную сумму выходов нейронов рецептивного поля. Это говорит о потере специфичности свойств объекта. В случае операции «максимум» выход сложной клетки есть максимальный сигнал от его афферентов и позволяет судить о лучшем соответствии части стимула с предпочтительными свойствами входных нейронов. Теоретически показана возможность существования MAX-подобной операции в сложных клетках области V1 зрительного пути [14, 15].

Покажем, каким образом будет локализованы наиболее контрастные области сцены. Общий вид модели приведен на рис. 3.

В результате линейной фильтрации получается 3 обрабатываемых потока: для цвета, интенсивности и ориентаций. Исследование контраста требует

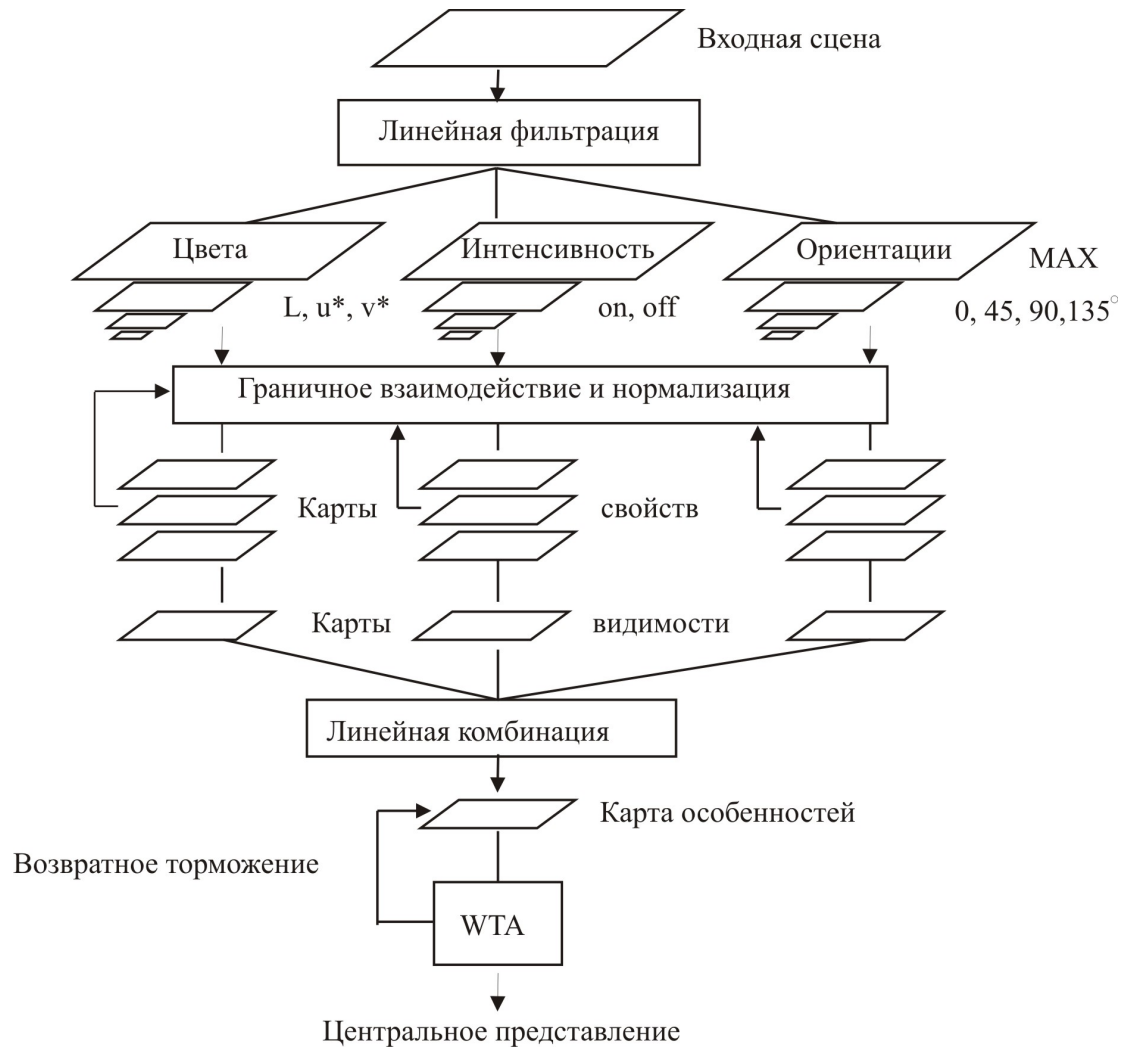


Рис. 3. Предлагаемая модель восходящего внимания



Рис. 4. Пирамидальное представление входной сцены

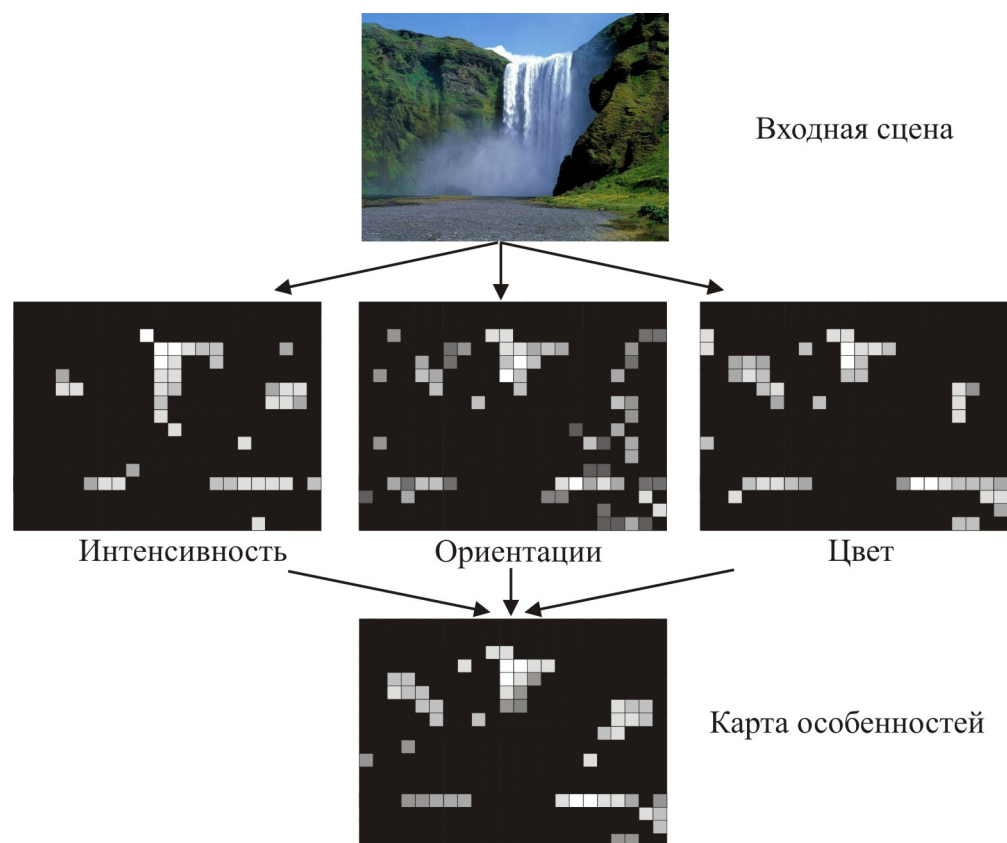


Рис. 5. Демонстрация функционирования модели

пирамидального представления входной сцены [16]. Другими словами, каждый поток обрабатывает входную сцену на нескольких масштабах.

Для всех последующих представлений в каждом потоке происходит двукратное уменьшение размеров, рис. 4. Это позволяет оценить степень отличности прямоугольных фрагментов сцены от их ближайших соседей, путем центрального взаимодействия граничных областей [17].

Для цветного представления используем переход в пространство Lu^*v^* . Обработка локальных ориентаций выполняется на каждом уровне с помощью дифференциальных гауссовых фильтров, отличающихся только шириной характеристики. После расчета активности простых клеток определяется наиболее подходящая локальная характеристика с помощью операции MAX.

Скорость такой обработки можно увеличить, если воспользоваться обновленной версией алгоритма [13]. Полученные карты свойств активируют для каждого потока карту видимости. Это, во-первых, сводит все представления к одному масштабу, и, во-вторых, позволяет судить о наличии той или иной характеристики внимания для фрагментов сцены.

Более сложным процессом является объединение полученных результатов. Как, например, можно сложить интенсивность и наличие локальных ориентаций? Несколько возможных решений этой

задачи продемонстрированы в [18]. Для построения нашей модели карты особенностей был взят алгоритм из [18], производящий настройку влияния каждого потока на конечный результат путем обучения с учителем на базе примеров.

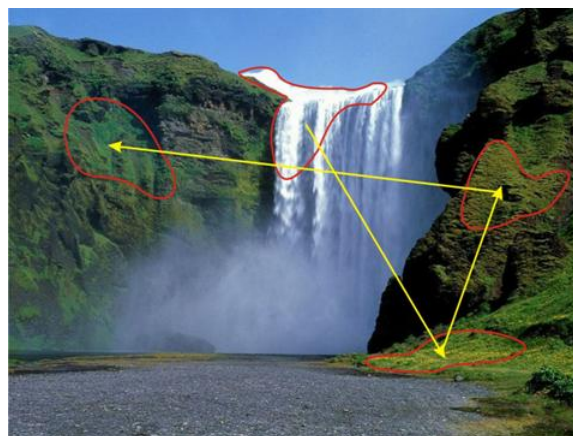


Рис. 6. Переход фокуса внимания

На рис. 5 показаны входная сцена активности карт видимости и особенностей. Более светлые области соответствуют более контрастным областям сцены.

Активность клетки на карте особенностей позволяет судить о контрасте ассоциированной с ней области входной сцены. После выявления самой активной области на карте особенностей, её

необходимо удалить из последующего анализа для перехода к следующей самой активной области путем возвратного торможения [19, 20].

В результате учета активностей соседних пикселей на карте особенностей, можно проследить перемещение фокуса внимания по сцене. Фокус внимания первоначально указывает на самую контрастную область, далее он перемещается (показано стрелками) на области, ассоциированные с менее активными клетками на карте, рис. 6.

СПИСОК ЛИТЕРАТУРЫ

1. Koch C., Ullman S. Shift in selective visual attention: towards the underlying neural circuitry // *Human Neurobiology*. — 1985. — № 4 (4). — P. 219–227.
2. Itti L., Koch C. A saliency-based search mechanism for overt and covert shifts of visual attention // *Vision Research*. — 2000. — № 40 (10–12). — P. 1489–1506.
3. Itti L., Koch C. Computational modeling of visual attention // *Nature Reviews Neuroscience*. — 2001. — № 2 (3). — P. 194–203.
4. Peters R.J., Iyer A., Itti L., Koch C. Components of bottom-up gaze allocation in natural image // *Nature*. — 2005. — № 45. — P. 2397–2416.
5. Walther D., Riesenhuber M., Poggio T., Itti L., Koch C. Towards an integrated model of saliency-based attention and object recognition in a primate's visual system // *Journal of Cognitive Neuroscience*. — 2002. — № 14. — P. 46–47.
6. Hubel D., Wiesel T. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex // *Journal of Physiology*. — 1962. — № 160. — P. 106–154.
7. Hubel D., Wiesel T. Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat // *Journal of Neurophysiology*. — 1965. — № 28. — P. 229–289.
8. Malacara D. Color Vision and Colorimetry: Theory and Applications. — NY: SPIE Press, 2002. — P. 215.
9. Bruce C., Desimone R., Gross C. Visual properties of neurons in a polysensory area in the superior temporal sulcus of the macaque // *Journal of Neurophysiology*. — 1981. — № 46. — P. 369–384.
10. Moran J., Desimone R. Selective attention gates visual processing in the extrastriate cortex // *Science*. — 1985. — № 229. — P. 782–784.
11. Connor C., Preddie D., Gallant J., Van Essen D. Spatial attention effects in macaque area V4 // *Journal of Neuroscience*. — 1997. — № 17. — P. 3201–3214.
12. Riesenhuber M., Poggio T. Hierarchical models of object recognition in cortex // *Nature Neuroscience*. — 1999. — № 2 (11). — P. 1019–1025.
13. Riesenhuber M., Poggio T., Riesenhuber M., Poggio T. Are cortical models really bound by the «binding problem»? // *Neuron*. — 1999. — № 24. — P. 87–93.
14. Wang G., Tanifuji M., Tanaka K. Functional architecture in monkey inferotemporal cortex revealed by in vivo optical imaging // *Neuroscience Research*. — 1998. — № 32. — P. 33–46.
15. Logothetis N. Object vision and visual awareness // *Neurobiology*. — 1998. — № 8. — P. 236–544.
16. Burt P.J., Adelson E.H. The Laplacian Pyramid as a Compact Image Code // *IEEE Transactions on Communications*. — 1983. — № 31. — P. 532–540.
17. Leventhal A.G. The Neural Basis of Visual Function // *Vision and Visual Dysfunction*. — Boca Raton, FL: CRC Press. — 1991. — № 4. — P. 76–95.
18. Itti L. Models of bottom-up and top-down visual attention: Doctor of Philosophy thesis. — Pasadena, 2000. — P. 70–115.
19. Cannon M.W., Fullenkamp S.C. Spatial interactions in apparent contrast: inhibitory effects among grating patterns of different spatial frequencies, spatial positions and orientations // *Vision Research*. — 1991. — № 31 (11). — P. 1985–1998.
20. Cannon M.W., Fullenkamp S.C. A model for inhibitory lateral interaction effects in perceived contrast // *Vision Research*. — 1996. — № 36 (8). — P. 1115–1125.

Поступила 14.04.2008 г.

Ключевые слова:

Восходящее внимание, восприятие образов, нейронные сети.