

СЕМАНТИКА РАСПОЗНАВАНИЯ ЗРИТЕЛЬНЫХ ОБРАЗОВ

© 2011 г. **В. В. Александров***, доктор техн. наук; **В. В. Александрова****, канд. пед. наук;
А. А. Зайцева*, канд. техн. наук

* Учреждение Российской академии наук Санкт-Петербургский институт информатики и автоматизации РАН, Санкт-Петербург

** Российский государственный педагогический университет им. А.И. Герцена, Санкт-Петербург

E-mail: alexandr@iias.spb.su

В статье сопоставляются внешние реальные образы (изображения, видео) с внутренним миром их семантического представления. Предложен алгоритм семантического узнавания и распознавания зрительных образов. Приведено определение понятия “интеллектуальной оптики”.

Ключевые слова: зрительный образ, семантический поиск, проблема Кука, узнавание, распознавание, интеллектуальная оптика.

Коды OCIS: 100.5010, 330.5020.

Поступила в редакцию 09.03.2011.

Мир не есть все, что может быть (не *Infinatum*, а *Indeterminatum*). Продвижение науки как в бесконечно малое, так и бесконечно большое, требует “интеллектуальной оптики” и новых метафизических постулатов.

Г. Лейбниц

Интеллектуальная оптика

Познание через зрительный образ [1, 2] – это внутренний мир воображения, видение нашего разума, компилирующего и интерпретирующего не только окружающий мир, но и себя в нем.

Разум человека, его творческое рациональное *о-сознание* и иррациональное, интуитивное *под-сознание* по своей природе превосходят ту конечность, которая характерна для всего внешнего природного физического окружения. Значит, биологическому разуму предоставлена бесконечная воспринимающая способность мышления посредством языка предмета восприятия. Внешний мир образов и внутренний мир воображений – это виртуальная пространственная и временная свертка событий. Проблема “интеллектуальной оптики” в неосвоенности технологий производства **биотронов** – биологических сенсоров, транслирующих сигналы разной природы окружающей среды (тактильные, зрительные, слуховые и др.) непосредственно

в “бульон” физиологических нейронных сетей. Это подобно **оптронам**, открывшим эру широкополосной, скоростной передачи данных и преобразующим свет в электронный сигнал (оптоволокно) [3].

Компьютерный электронный разум (искусственный интеллект) в большей степени опирается на последовательный, вычислительный анализ некой априорно выбранной математической модели, которую разум программиста превращает в конечный набор инструкций. Последующие решения, каждое из которых зависит от решения, принятого на предыдущей фазе, строятся по схеме дерева посредством серии бинарных дизъюнкций, предикатов и т. д. Это принципиально конфликтная схема, при которой игрокам известны правила игры и каждый принимает решения **поступательно, по одному за раз**, как в шахматной партии. Вычислительная сложность таких NP-схем – это сборка мозаики вслепую. Решение возможно методом полного перебора, но это не раскрывает информационное содержание контента (семантики).

Семантическая идентификация неразличимости

“Параллельный” электронный разум, напротив, передоверяет отдельным фрагментам иерархической структуры список понятий и распределения “семантических компонентов”, составляющих исходное изображение. Разработанная компьютерная программа модифицируется в поисках оптимального решения, семантического распознавания при минимальном битовом представлении [4]. Итерационная конвертация программа-данные выявляет локальные аттракторы смены информационной неопределенности семантического контента.

Семантическая селекция основана на оценке двух процессов: среднеквадратического отклонения (СКО) между исходным изображением ($\varepsilon = 0$) и последующими слоями ($\varepsilon > 0$)

и семантической оценки – СКО между соседними слоями.

На рис. 1 и 2 приведены примеры, иллюстрирующие итерационный процесс семантической сегментации и показывающие возможность реализации автоматизированного выбора ε -слоя информационной достаточности для семантического узнавания (распознавания).

Этот эмпирически найденный критерий семантической идентифицируемости позволяет реализовать процесс автоматизированного выбора ε -слоя, достаточного для семантического распознавания, что, в свою очередь, делает возможным организовать ассоциативный визуальный поиск и построение специализированных кодеков.

Приведенным в эпитафье афоризмом “не бесконечность, а неопределенность” Г. Лейбниц отметил аналогию, существующую между

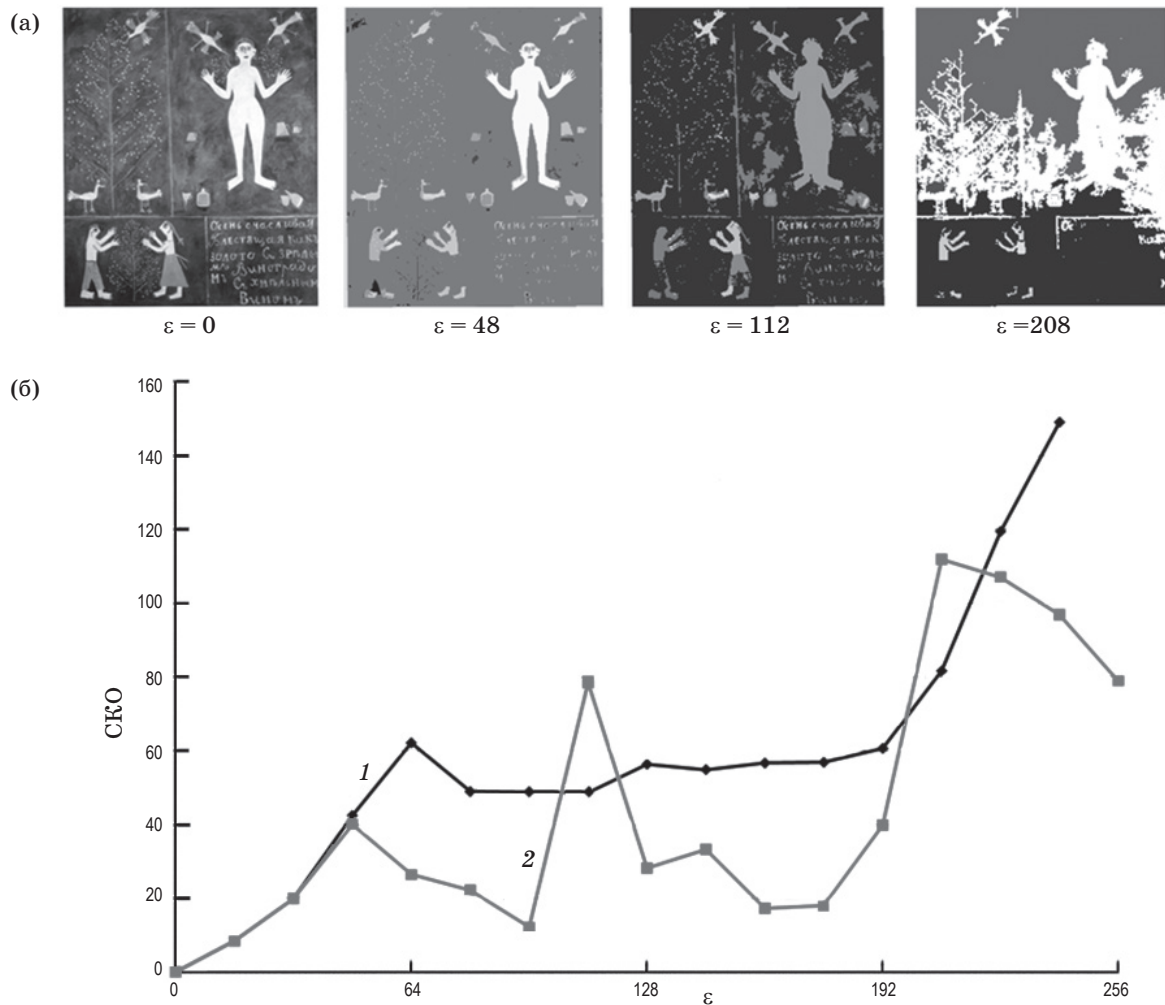


Рис. 1. Репродукция картины М. Ларионова (а) и результаты ее обработки компьютерной программой семантической сегментации (б). 1 – СКО между исходными изображениями ($\varepsilon = 0$) и последующими слоями, 2 – семантическая оценка СКО между слоями.

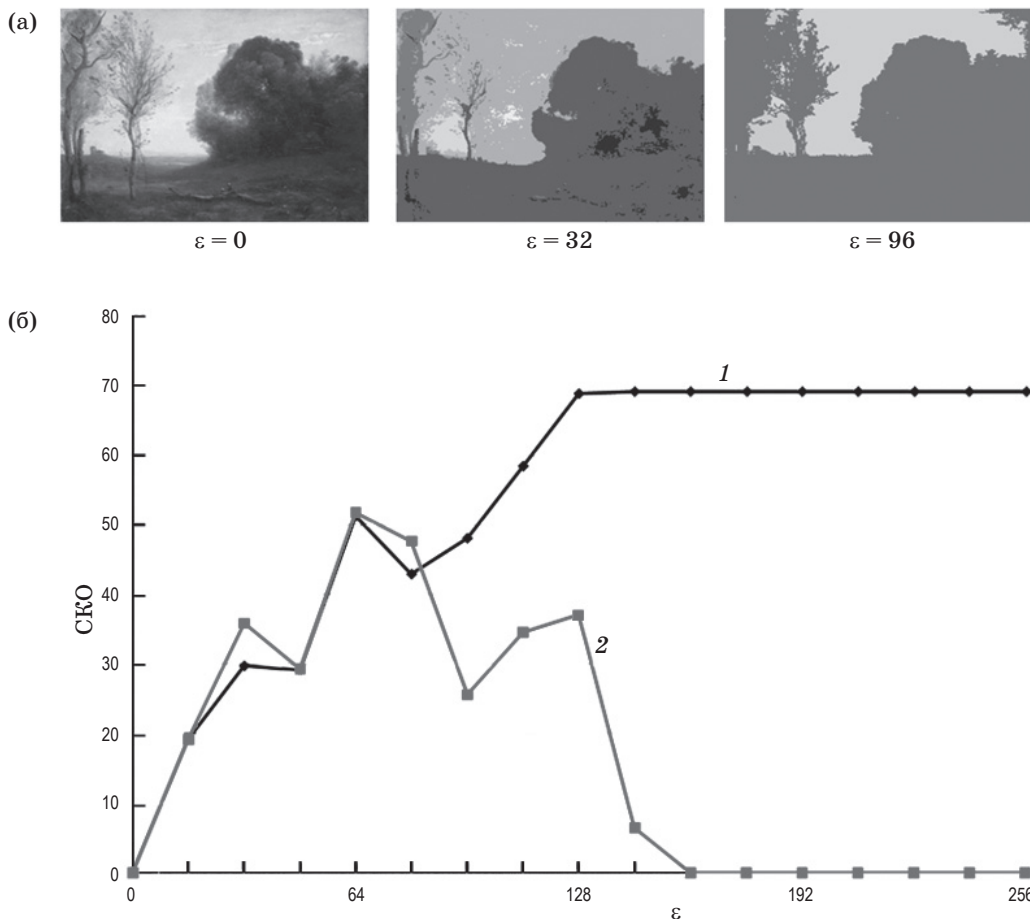


Рис. 2. Репродукция картины Камиля Коро “Утро” (а) и результаты ее обработки компьютерной программой семантической сегментации (б). 1 и 2 – те же, что на рис. 1.

проблемой энтропии (как мерой неопределенности) и проблемой структуры континуума (вложенной иерархии): “обе имеют общий логический корень, связанный с актуальной бесконечностью, что требует “интеллектуальной оптики” и новых метафизических *постулатов* (не *Infinitem*, а *Indeterminatum*)” [5].

Познание, когнитивность обычно связывают с воображением как искусством (интуитивным, иллюзорным и трудно формализуемым) и наукой (аксиоматической формализацией). Такое противопоставление искусства и науки проистекает из культа рационализма XVIII столетия. С тех пор рационалистический, научный уклон современного образования придает правдоподобность только двум сторонам эмпиризма, на которых базируются естественные науки, – причине и следствию.

Отсюда и удивительные спекуляции приписывания, рефлексия оптических свойств и фи-

зических процессов – фурье-анализа, адекватности свойств биологической среды глаза и зрительного анализатора человека. Больше 50 лет, начиная с У. Маккаллока и В. Питтса [6], продолжают работы по созданию искусственных (компьютерных) нейронных сетей, моделирующих еще до конца не изученные функциональные свойства биологического разума. Тест А. Тьюринга “Может ли машина мыслить?” подменяется конгруэнтностью, “идентификацией неразличимости” Г. Лейбница. Такие компьютерные игры, как “шахматы” и “эрудиты”, в действительности реализуют полный перебор (а он для данных игр принципиально конечен по памяти и глубине поиска), и только в этом случае современные компьютерные технологии конкурентоспособны с биологическим интеллектом и разумом. Но достаточно незначительно изменить правила игры, перепутать входные данные, как адаптация биологическо-

го разума выигрывает. Разум человека легко различает иллюзии и осуществляет семантическую селекцию, а объемное пространственное цветовое видение (3D-видео) – это чаще всего искусственные порождения мира воображения и фантазий.

Возникает проблема существования “ментальной каузальности”, вытекающей из причинно-следственной симметричности и замкнутости физических теорий. Суть проблемы в следующем вопросе: можно ли всерьез относиться к свободе человеческого выбора (сознательного и/или бессознательного), если физический мир каузально замкнут, т. е. у всякого физического события всегда имеется физическая же причина?

Другими словами, если мы принимаем решение, то вовсе не потому, что такова наша воля, а потому, что в нашем головном мозге произошли соответствующие физико-химические реакции, которые привели к возбуждению нервной проводимости и затем – к сокращению мышечных волокон. А физико-химические реакции в коре головного мозга также имеют свои физические причины в виде предшествующих им реакций. Поэтому не остается никакого места для свободной воли, а есть лишь следование физических объектов физическим же законам. Отсюда – отрицание некоей, отдельной от физической “ментальной причинности” (например “лента Мебиуса”); восприятие фантазмов, “ментальных состояний” как соответствующих конфигураций (не до конца изученных, однако изучаемых) электрохимических полей головного мозга [7].

Мир образов и мир воображения

Мир образов представляется идеей (или формой), которая ограничивает бесконечное, осуществляя “вписывание” образа в упорядоченный семантический, смысловой ряд. Однако биологический разум, все еще с непознанной целью, порождает образы несуществующих химер, такие невозможные в физическом пространстве фигуры, как “лента Мебиуса”, “лестница Эшера” и другие [8], строит абстракции субстанциональности точки, числа, бесконечности и т. д. Любая запись, изображение или программа, в свою очередь, может рассматриваться как большое бинарное число. В современной юриспруденции появилось понятие незаконного изображения (из-за непристойности или секретного статуса). Соответственно, связанное с ним число также может рассматри-

ваться как незаконное. Все эти знания, “врожденные идеи” – эксклюзивные качества биологического разума человека.

Для электронного разума это лишь увеличение разрешения (до 3840×2160 пикселей), скорости обработки, передачи и архивации без возможностей интеллектуального и семантического поиска.

Для биологического разума видение и понимание – единый процесс многофункциональности “интеллектуальной оптики”, проиллюстрированный фильмом “Аватар”. Пять инноваций, составляющих единое целое, привели к качественной модернизации видео- и звуковых студий и компьютерного 3D-видео.

- **Трехмерная камера Fusion 3D** – “интеллектуальный глаз”. Объективы камеры (подобно глазу) синхронно фокусируются на близких и дальних объектах.

- **Пять суперкомпьютеров** – общий объем цифровых данных превысил 1 Пб (10^6 Гб), и это не технологический, а интеллектуальный прорыв.

- **Система захвата движений** – 120 камер, осуществляющих “оцифровку” актеров, с точностью до миллиметра. Качественно иной уровень мультимедийных технологий.

- **“Виртуальная камера”** – программно реализует аппликацию и симуляцию несуществующих ландшафтов и сцен.

- **Язык “на’ви”, придуманный лингвистом**, – обеспечивает эмоциональное звуковое сопровождение, усиливающее воздействие на вербальное мышление (“слова развивают разум”), подчеркивая эффект соучастия через интерфейсное взаимодействие слухового и зрительного восприятия. При этом учитывается психофизиологический закон восприятия – темп изменения языка не должен превышать способности к его освоению, адаптации и памяти.

Заключение

Экспериментальные исследования процессов зрительного восприятия биологическим разумом показали наличие потребности сопоставления внешних образов с внутренним миром знаний (человек, говоря словами И. Гете, “видит то, что знает”).

Например, в [9] испытуемым предъявляли репродукции пейзажей и предлагали их кратко описать, указав время суток и погоду. В качестве тестового материала служили разные ре-

продукции. На рис. 2 приведена одна из них – репродукция картины К. Коро “Утро”. Выбор этой репродукции обусловлен утонченными цветовыми тональностями (без резких контрастов), соответствующими расплывчатости форм предметов, в то же время они прекрасно передают непосредственный, мгновенный взгляд и выявляют семантическую значимость.

Послойная обработка (рис. 2) иллюстрирует узнавание как процесс семантической иден-

тификации неразличимости. Биологический разум интерполирует известные ему оттенки. Именно поэтому проблема семантического узнавания значительно проще поиска и распознавания неизвестного. Это аналогично проблеме Кука: “решение какой-либо задачи часто занимает больше времени, чем проверка правильности решения” [10].

В этом и проявляется роль обучения, воображения и знания.

* * * * *

ЛИТЕРАТУРА

1. *Alexandrov V.V., Gorsky N.D.* From Humans to Computers: Cognition through Visual Perception // Singapore, New Jersey, London, Hong Kong: Welt Scientific, 1991. 203 p.
2. *Александров В.В.* Глаз и визуальное восприятие // Оптический журнал. 1999. Т. 66. № 9. С. 54–63.
3. *Александров В.В., Полонников Р.И., Трофимов Е.И.* Оптоэлектронные устройства и перспективы их использования в РЭА // Труды конф. по электронной технике / ЦНИИ “Электроника”, 1971. В. 5(31).
4. *Аксенов А.Ю., Зайцева А.А., Кулешов С.В.* Критерий ε -идентифицируемости в обработке аудио- и видеоданных // Труды VIII Всерос. научно-практич. конфер. “Современные информационные технологии в науке, образовании и практике”. Оренбург. 2009. С. 348–350.
5. *Лейбниц Г.В.* О свободе // Сочинения в четырех томах. Т. 1. М.: 1982. С. 312–317.
6. *Мак-Каллок У.С., Питтс В.* Логическое исчисление идей, относящихся к нервной активности // В сб.: “Автоматы” под ред. К.Э. Шеннона и Дж. Маккарти. М.: Изд-во иностр. лит., 1956. С. 363–384.
7. *Джадан Игорь.* Арифметика нанотехнологической революции // “Агенство политических новостей”. Лицензия минпечати. Эл. № 77-2792. <<http://www.apn.ru/publication/article21126.htm>>
8. *Hofstadter D.A.* Godel, Esher, Bach: An eternally golden braid. N. Y.: Harvester Press, 1979. 843 p.
9. *Николаенко Н.Н., Черниговская Т.В.* Опознание сложных цветовых образов и функциональная асимметрия мозга // Вопросы психологии. 1989. № 1. С. 107–112.
10. *Александров В.В., Кулешов С.В., Цветков О.В.* Цифровая технология инфокоммуникации. Передача, хранение и семантический анализ текста, звука, видео. СПб.: Наука, 2008. 244 с.