

## **К вопросу об использовании преобразования “неоднородная сетчатка – однородная зрительная кора” в системе технического зрения**

**С.А.Гладилин, Д.Г.Лебедев**

*Институт проблем передачи информации, Российская академия наук, Москва, Россия*

Поступила в редколлегию 10.11.2005

**Аннотация**—В статье предлагается методика построения системы технического зрения, использующей принцип “неоднородная сетчатка – однородная зрительная кора”. Обсуждаются обстоятельства, затрудняющие создание такой системы и указываются пути их преодоления.

### **1. ВВЕДЕНИЕ**

В настоящее время в мире наблюдается возрастающий интерес к интеллектуальной робототехнике. Это, в первую очередь, связано с разработкой антропоморфных роботов, предназначенных для существования в среде обитания человека и выполняющих, в основном ту же работу, которую до этого вынужден был выполнять человек. Иногда эти роботы называются гуманоидными или андроидными.

Зарубежные фирмы, такие как Honda, Sony и другие, вкладывают большие средства в разработку антропоморфных роботов. В Японии существует многолетняя правительственная программа поддержки и координации работы фирм в этом направлении. Основное применение интеллектуальных роботов носит социальный характер: это помощь престарелым, больным или одиноким людям. Учитывая большую востребованность подобной продукции рынком, в целях обеспечения коммерческого успеха, информация о разработках антропоморфных роботов, как правило, закрыта. Демонстрируется только результат (образец робота) на различных выставках.

Всё же, положение можно охарактеризовать следующим образом: основные усилия направлены на разработку управления походкой и манипулированием. В этом направлении достигнуты впечатляющие успехи. Что же касается зрительной системы роботов, то она пока тяготеет к стандартным решениям, пришедшим из телевидения.

В то же время, антропоморфными роботы названы не только за свое внешнее сходство с человеком (хотя это также имеет значение при работе робота среди людей, например, роботы-помощники, роботы-слуги и т.п.) но, главным образом, это должны быть роботы, использующие специфические методы обработки информационных потоков, заимствованные у человека. Это относится и к зрительной системе роботов.

Зрительная система человека весьма совершенна. Она в состоянии выделять яркие контрасты, соизмеримые с уровнем шума и успешно решать основную задачу – распознавание объектов на изображении. Зрительная система человека обладает уникальным свойством – обеспечением высокой остроты зрения при широком угле обзора. Считается, что прогрессивная эволюция приматов в значительной степени связана с формированием единого морфофункционального комплекса, в который входит существенно неоднородная сетчатка, а также

соответствующий аппарат глазодвигательных центров мозга, обеспечивающий возможность использования сетчатки в связи с задачами распознавания образов и афферентного синтеза.

Поэтому, применение бионических принципов обработки информации в технической зрительной системе, направленное на резкое повышение эффективности ее работы, является актуальной задачей. Настоящее исследование проводилось именно в этом направлении.

В данной работе обсуждаются обстоятельства, затрудняющие создание системы технического зрения с неоднородной сетчаткой и указаны пути устранения этих затруднений. Разработана соответствующая методика на существующей технической базе и проведено компьютерное моделирование работы СТЗ, использующей преобразование “неоднородная сетчатка – однородная кора” (НС-ОК).

В [1] разработана бионическая модель системы “неоднородная сетчатка – однородная кора” (НС-ОК), включающая компьютерную модель сетчатки человека для колбочкового зрения, математическое описание формы поверхности зрительной коры и алгоритм проектирования точки сетчаточной проекции на зрительную кору.

Проведен компьютерный эксперимент, позволяющий сделать вывод о том, что преобразование НС-ОК придает зрительной системе следующие свойства:

Корковые проекции имеют одинаково резкие контрасты при любом эксцентриситете;

Корковые проекции маленьких объектов в центре поля зрения и более крупных по мере смещения к периферии эффективно анализируются одними и теми же операторами;

Неоднородная сетчатка в зрительной системе человека позволяет использовать высокую локальную остроту зрения при большом ( $\sim 100^\circ$ ) угле обзора. Применение однородной сетчатки в этом случае привело бы к увеличению площади зрительной коры на несколько порядков. Хотя высокое разрешение в преобразовании НС-ОК сохраняется только в центре поля зрения, практика человеческой деятельности показывает высокую эффективность зрительной системы, построенной подобным образом.

Первым препятствием на пути применения системы НС-ОК в техническом зрении является отсутствие неоднородных фотоматриц как таковых. Предлагается использовать однородную фотоматрицу достаточно большого формата, сохранить ее высокое разрешение в центре и программным путем “загубить” его в направлении периферии. Таким образом, уже сейчас возможно имитировать работу неоднородной фотоматрицы с разрешающей способностью в центре 1 угл. мин. и углом обзора  $50^\circ$ .

На другое препятствие указывает разработанная бионическая модель. Это принципиальная невозможность преобразования НС-ОК, реализующего плавное снижение остроты зрения, не на поверхность объемного тела зрительной коры, как это сделано у человека, а на плоский растр, с которым принято работать (проводить анализ) в системах технического зрения. Это обстоятельство заставляет отказаться от использования плавной кривой падения разрешающей способности при смещении к периферии, заменив ее аппроксимирующей ступенчатой функцией.

## 2. МЕТОДИКА

Рассмотрим методику построения СТЗ, реализующую принцип НС-ОК.

Монохромное цифровое изображение, полученное от фотокамеры, будем называть исходным. Исходное изображение есть массив значений яркостей  $R_{lm}$ , где  $l, m$  – целые числа  $0 \leq l, m < n$ , где  $n$  – целое число, называемое форматом исходного изображения.

Опишем преобразования, осуществляемые над исходным изображением в СТЗ, построенной по принципу “неоднородная сетчатка – однородная кора”.

На исходном (квадратном) изображении выделяется ряд квадратных областей. Геометрический центр каждой области совпадает с геометрическим центром исходного изображения. Формат  $i$ -ой области (длину стороны квадрата) обозначим за  $n_i$ . Перенумеруем эти области так, чтобы  $n_i$  увеличивался с ростом  $i$ . Обозначим координаты левого нижнего угла  $i$ -ой области на исходном изображении за  $(r_i, r_i)$  – это возможно из-за квадратности исходного изображения и выделенных областей и совпадения их геометрических центров.

Для каждой выделенной области выбирается натуральный коэффициент  $k_i$ , возрастающий с ростом  $i$ , причем для  $i=1$  выбирается  $k_i = 1$ . Выделим в  $i$ -ой области из всех точек каждую  $k_i$  по горизонтали и вертикали и назовем их базовыми. Говоря математическим языком, базовыми точками  $i$ -ой области будем называть множество точек исходного изображения с координатами  $(r_i + xk_i + \left[ \frac{k_i}{2} \right], r_i + yk_i + \left[ \frac{k_i}{2} \right])$ , попадающих в  $i$ -ую область, где  $x, y$  – некоторые целые неотрицательные числа. Условие попадания базовой точки в  $i$ -ую область можно записать как

$$\begin{cases} xk_i + \left[ \frac{k_i}{2} \right] < n_i \\ yk_i + \left[ \frac{k_i}{2} \right] < n_i \end{cases}$$

Обозначим

$$p_t^i = tk_i + \left[ \frac{k_i}{2} \right]$$

Определим понятие конвергентной яркости базовой точки с координатами  $(p_x^i, p_y^i)$  как среднего арифметического по квадрату размером  $k_i \times k_i$  пикселей с геометрическим центром в базовой точке.

Для нечетных  $k_i$  базовая точка является геометрическим центром квадрата  $k_i \times k_i$ , границы которого проходят по границам пикселей. Отсюда

$$S_{xy}^i = \sum_{l=p_x^i - \left[ \frac{k_i}{2} \right]}^{p_x^i + \left[ \frac{k_i}{2} \right]} \sum_{m=p_y^i - \left[ \frac{k_i}{2} \right]}^{p_y^i + \left[ \frac{k_i}{2} \right]} R_{lm}$$

Для четных  $k_i$  рассмотрим перекрывающиеся квадраты размером  $(k_i + 1) \times (k_i + 1)$ , но пиксели, относящиеся к нескольким квадратам, будем учитывать с пропорционально меньшим весом:

$$S_{xy}^i = \sum_{l=p_x^i - \left[ \frac{k_i+1}{2} \right]}^{p_x^i + \left[ \frac{k_i+1}{2} \right]} \sum_{m=p_y^i - \left[ \frac{k_i+1}{2} \right]}^{p_y^i + \left[ \frac{k_i+1}{2} \right]} c_{lm}^i R_{lm}$$

где

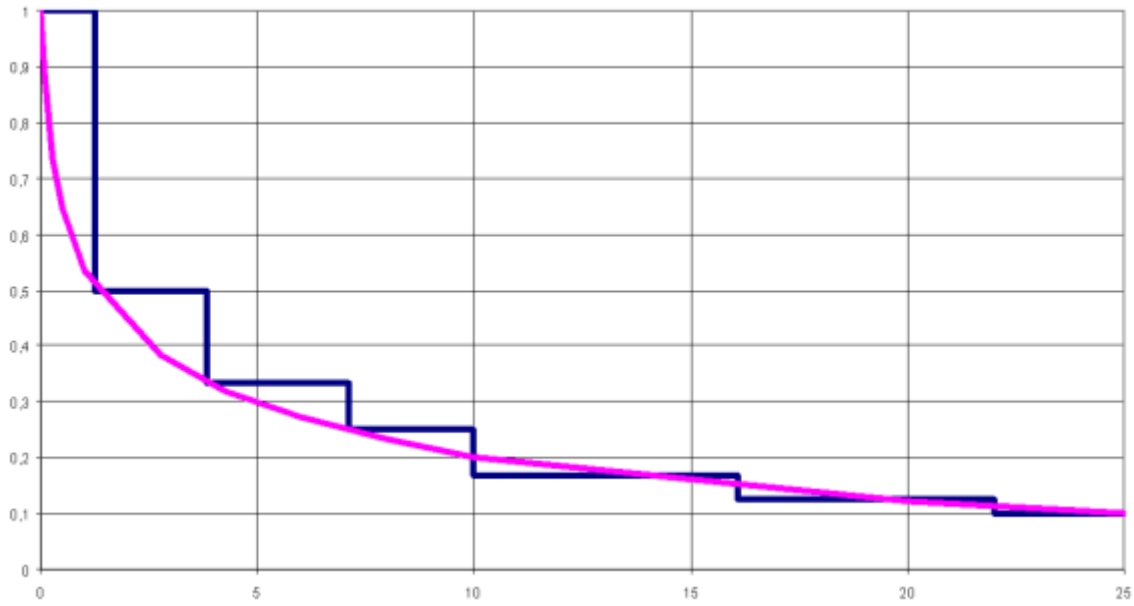


Рис. 1.  
 Ступенчатая функция, аппроксимирующая кривую остроты зрения человека

$$c_{lm}^i = \begin{cases} 1, & p_x^i - \left[ \frac{k_i+1}{2} \right] < l < p_x^i + \left[ \frac{k_i+1}{2} \right] \quad p_y^i - \left[ \frac{k_i+1}{2} \right] < m < p_y^i + \left[ \frac{k_i+1}{2} \right] \\ \frac{1}{2}, & l = p_x^i \pm \left[ \frac{k_i+1}{2} \right] \quad p_y^i - \left[ \frac{k_i+1}{2} \right] < m < p_y^i + \left[ \frac{k_i+1}{2} \right] \\ \frac{1}{2}, & p_x^i - \left[ \frac{k_i+1}{2} \right] < l < p_x^i + \left[ \frac{k_i+1}{2} \right] \quad m = p_y^i \pm \left[ \frac{k_i+1}{2} \right] \\ \frac{1}{4}, & l = p_x^i \pm \left[ \frac{k_i+1}{2} \right] \quad m = p_y^i \pm \left[ \frac{k_i+1}{2} \right] \end{cases}$$

Введем понятие зонального изображения, соответствующего  $i$ -ой области, как изображения, составленного только из базовых точек области со значениями в них, равными конвергентным яркостям. Из определения базовых точек и конвергентной яркости следует, что зональное изображение есть, по сути, уменьшенная в  $k_i$  раз копия соответствующей ему области. Размеры областей на исходном изображении ( $n_i$ ), а также соответствующие им коэффициенты уменьшения ( $k_i$ ) выбираются исходя из ступенчатого приближения функции относительной остроты зрения человека. Значение этой функции равно кортикальному усилению, т.е. обратно пропорционально коэффициенту уменьшения, введенному для СТЗ.

### 3. КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ

Моделировалась работа СТЗ, использующую преобразование НС-ОК.

Для разрешающей способности 1 угл. мин. и формата фотоматрицы 3000.×3000, обеспечивающей угол зрения 50°, была сформирована ступенчатая функция, аппроксимирующая функцию относительной остроты зрения человека, со ступеньками высотой обратной натуральным числам – 1,  $\frac{1}{2}$ ,  $\frac{1}{3}$ ,  $\frac{1}{4}$ ,  $\frac{1}{6}$ ,  $\frac{1}{8}$  (рис. 1). На основе полученного приближения рассчитаны форматы зональных изображений и степень охвата ими поля зрения (таблица).

Целью компьютерного моделирования было получение зональных изображений для визуального сравнения с исходным. Исходным изображением (монохромное, формата 3000.×3000)

Таблица

Коэффициент уменьшения	Охват поля зрения	Формат зонального изображения
1	2,5° .×2,5°	152.×152
2	7,7° .×7,7°	231.×231
3	14,2° .×14,2°	284.×284
4	20,0° .×20,0°	300.×300
6	32,1,9° .×32,1°	322.×322
8	43,9° .×43,9°	330.×330
10	50,0° .×50,0°	300.×300

*Форматы зональных изображений и степень охвата ими поля зрения*

был горный пейзаж с 256 градациями яркости с большим количеством мелких деталей. Описанный выше алгоритм формирования зональных изображений прост и эффективен за счет принятого решения о целочисленности коэффициентов уменьшения и использования только целочисленной арифметики. Исходное изображения и полученные зональные изображения представлены на рис. 2. На основе эксперимента сделаны следующие выводы:

- Все зональные изображения имеют приблизительно одинаковую четкость (крутизну фронтов)
- Объекты на зональных изображениях сохраняют углы и пропорции

Таким образом, в системе технического зрения использующей преобразование НС-ОК и обладающей возможностью работы с высоким локальным разрешением при большом угле обзора, анализ сцены производится только на небольшом количестве зональных изображений малого формата. При использовании в СТЗ фотоматрицы формата 9 мегапикселей и локальном разрешении 1 угл. мин. объем обрабатываемого цифрового массива сокращается в 16 раз.

#### 4. ЗАКЛЮЧЕНИЕ

Практическому внедрению преобразования НС-ОК должны предшествовать следующие исследования:

- Разработка алгоритма анализа зональных изображений (выделения контрастов, ориентированных линий, углов и т.п.)
- Разработка алгоритмов последовательного анализа сцены путем смены точки фиксации взора, распознавания образов в процессе анализа зональных изображений и последовательного анализа сцены. Конечным продуктом является распознавание сцены, достаточное для решения поставленной перед СТЗ задачей (поиск заданного объекта, оценка возможности продвижения и т.п.)

Перечисленные этапы дальнейшей работы безусловно связаны с практической реализацией СТЗ, потому что проверка эффективности того или иного предложенного алгоритма должна производиться на зональных изображениях реальной трехмерной сцены. В сущности, речь идет о создании лабораторного макета СТЗ, предоставляющего материал для отработки указанных алгоритмов. В связи с исследовательским назначением можно выработать ряд требований к подобной СТЗ:

- Системы следует реализовывать на базе цифровой фотокамеры формата не менее 9 мегапикселей
- Система может быть монохромной и монокулярной
- Фотокамера должна иметь возможность поворотов в вертикальной и горизонтальной плоскостях на заданные углы, а также управления фокусировкой объектива

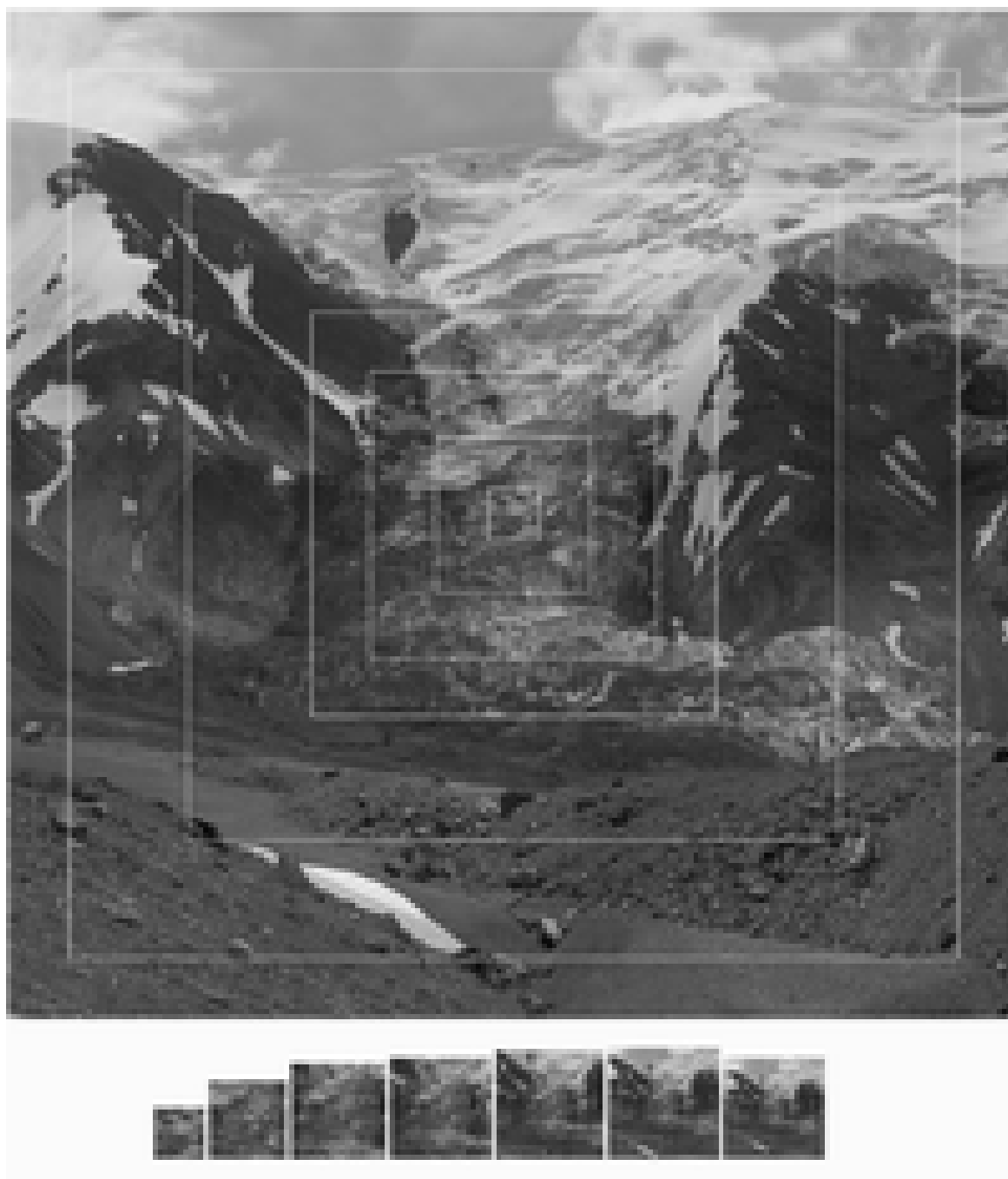


Рис. 2.

## СПИСОК ЛИТЕРАТУРЫ

1. Гладилин С. А. Компьютерная модель зрительной системы с кортикальным усилением *Информационные процессы*, 2005, т. 5, № 5, стр. 414–425.

*Статью представил к публикации член редколлегии В.И. Венец*