

# СИСТЕМА РАСПОЗНАВАНИЯ ДВИЖУЩИХСЯ ОБЪЕКТОВ НА БАЗЕ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ

С. Яковлев, *Mg.sc.comp.* e-mail: : [sergeyk@fis.lv](mailto:sergeyk@fis.lv), [tac@inbox.lv](mailto:tac@inbox.lv)

**Аннотация.** Разработаны методы с более высокой скоростью работы по сравнению с методами «перцептрон» и «угловая классификация». Реализованы основы системы распознавания образов в динамической последовательности изображений (РОДПИ). Для этого проведены серии экспериментов, выдвинут и экспериментально подтвержден ряд гипотез, следствия которых применимы в архитектуре реализованной системы

## 1. Введение

В данной статье с одной стороны рассматривается модель зрения – перцептрон Розенблата [1] с некоторыми усовершенствованиями [2]. С другой стороны, модель памяти – угловая классификация CC4 (Cornel Classification) [3] - дополняет этот подход, показывая особенности интеллекта, не заметные в модели перцептрона.

Метод CC4 положил начало методу под названием CCM (Cornel Classification In Model) [4], в котором основной упор был сделан на увеличение скорости работы. В статье приведены сравнительные характеристики упомянутых методов и их вариаций. Но наибольшего внимания заслуживает возможность совмещения метода CCM и перцептрона, а также формирование единицы восприятия «нейрон». С помощью модели восприятия под условным названием «нейрон» решается задача восстановления и анализа изображения. Эти задачи решаются в условиях, когда распознаваемый объект находится в движении.

Работа алгоритма CCM, и последующее совмещение алгоритмов CC4 с перцептроном, основывается на следующих принципах:

1. Расчет активности среднего слоя нейронной сети математическим способом, без полного перебора элементов сети;
2. Применение  $\alpha\gamma$  – системы подкрепления, которая не позволяет бесконечно увеличиваться весам AR связей, что ускоряет сходимость при обучении в несколько раз;
3. Разделение нейронной сети на блоки, каждый из которых ответственен за свою ограниченную область. Такое разделение существенно уменьшает время распознавания по сравнению с неразделенной нейронной сетью;
4. Уменьшение  $A$  – элементов в нейронной сети почти вдвое по сравнению с минимально необходимым числом (на каждый стимул по одному  $A$  - элементу).

Данные принципы позволяют существенно увеличить скорость работы нейронной сети без существенной потери качества и отказа распознавать.

## 2. Сравнение различных версий методов, основанных на методе ССМ

Характерной чертой классических методов является низкая скорость работы. Поэтому первой задачей стала разработка таких методов, которые работали бы с высокой скоростью и при этом, по возможности, не теряли бы качества по сравнению с классическими методами. Было разработано несколько различных методов, но основой их стал метод ССМ, который, в свою очередь, основывался на принципах угловой классификации. В таблице 1 отражены характеристики работы предложенных методов по сравнению с методами перцептрон и угловая классификация.

Таблица 1

Характеристики разработанных методов

Метод	Скорость после обучения	Качество (кол-во ошибок)	Память	Обучение	Вероятность отказа
<i>СС4</i>	-	-	N - эталонов; Ув. Эталонов $\neq$ Ув. Качества	1 итерация	0%
<i>Перцептрон</i>	-	-	N - эталонов	N итераций	0%
<i>ССМ/ССМ+</i>	↑ в 2 раза	↑10%	см. СС4	1 итерация	0%
<i>ССМ2/ССМ2+</i>	↑ в 2 раза	↑ 15%-25%	см. СС4	1 итерация	0%
<i>ССМ3</i>	-	↑ 55%	см. СС4	1 итерация	0%
<i>Перцептрон + ССМ</i>	↑ в 7 раз	↑ 10%	N - эталонов	N итераций	5%
<i>Модель "нейрон" [i]</i>	↑ в 7 раз	↑ 10%	60% * N - эталонов	<2000*i	5%

Данные результаты получены на основании экспериментов по решению задачи восстановления изображений. Следует отметить, что метод ССМ3 дает наилучшее качество, но низкую скорость, а совмещение методов «перцептрон+ССМ», дает высокую скорость без падения качества. Поэтому он и был положен в основу модели восприятия «нейрон».

## 3. Основы системы, которая восстанавливает и анализирует несколько движущихся объектов

### 3.1. Постановка задачи

**Модельная ситуация:** Человек видит передвижение объекта (как минимум три статических изображения). Данный объект человек отличает из общего фона

местности за счет того, что он движется. Этот объект сохраняется в памяти человека, но не целиком, а частично, так как память человека ограничена по объему. Таким образом, человек видит и запоминает несколько объектов, причем объекты могут повторяться, но на разном фоне. Затем человеку показывают изображение одного из двух запомненных объектов, и он должен воссоздать в зрительной памяти как можно точнее соответствующий объект и сказать, что это за объект.

Чтобы *решить данную задачу*, ее следует разбить на следующие шесть технических задач:

1. Отбор изображения объекта из общего фона на основании трех статических изображений, где объект движется.
2. Выбор случайной последовательности и запоминание случайных точек (определенных выбранной случайной последовательностью) в нейронной сети, моделирующей память. Количество точек определяется доступным размером памяти, но желательно не менее 10% от общего количества точек.
3. Выбор конкретных точек (например, точки расположены линиями, образующими сетку на изображении) и запоминание цвета данных точек в нейронной сети, с целью на такую стимуляцию получить реакцию в виде случайной последовательности, выбранной в п. 2.
4. На основании показанных изображений, а точнее, цвета точек на изображении (например, точки расположены линиями, образующими сетку на изображении), определить какой это объект – первый или второй.
5. На основании определенной в п. 4. случайной последовательности воссоздать изображение, запомненное в нейронной сети.

Рассмотрим подробнее данные задачи. *Первая задача* решается с помощью детерминированного алгоритма – детектора движения (см. раздел 3.3.). *Задачи 2 и 5*, а также *задачи 3 и 4*, представляют собой задачу восстановления и анализа объектов, и их описание вынесено в дальнейшие соответствующие разделы.

### 3.2. Модель восприятия – «Нейрон»

«Нейрон» представляет собой абстрактную единицу некоторой системы («Внешнего мира»), которая имеет некоторые способности, основанные на самообучении, и выполняет определенную часть предметной задачи. Структурно «нейрон» является целой искусственной нейронной сетью, но так как он является обобщенной единицей предметной задачи, то для простоты он назван просто «нейрон».

Сложность задачи, которую может решать отдельный «нейрон», определяется в основном количеством элементов для запоминания (А-элементов среднего слоя нейронной сети). Экспериментально было определено, что при количестве 2000 элементов «нейрон» показывает хорошие результаты в плане

производительности. При большем количестве элементов, хоть и решается более сложная задача, но скорость экспоненциально падает.

Основные принципы работы «нейрона» такие же, как в методе ССМ, но кроме него применяется перцептронное обучение  $\alpha \gamma$  - системы подкрепления для AR связей. Это позволяет использовать «нейрон» в трех режимах:

1. Точное сравнение новых объектов с запомненными объектами;
2. Нечеткое сравнение (определенное матрицей М) новых объектов с запомненными объектами;
3. Нечеткое сравнение, сжатием информации, посредством обобщения данных.

Первые два режима обеспечиваются методом ССМ, а третий с помощью перцептронного обучения. Регулируются они коэффициентом обобщения  $K$ . Уровень активности А-элементов при  $k=1$  такой, что возбуждаются только те А-элементы, которые точно совпадают со сравниваемым объектом. При  $k=2$  возбуждаются те А-элементы, которые близки по значению со сравниваемым объектом. При  $k=3$  уровень активности А-элементов практически случайный, что позволяет для перцептрона создать определенную непротиворечивую G-матрицу. Это, в свою очередь, позволяет применить перцептронное обучение.

Обучение позволяет сократить количество А-элементов. Теоретически, чтобы процесс обучения сходил, нужно иметь как минимум столько же А-элементов, сколько имеется объектов для запоминания, т.е. на каждый элемент запоминания по одному А-элементу. Иначе процесс сходимости не гарантируется. Но практически получается сократить число А-элементов на 80%, при этом в 95% случаев процесс обучения сходится без проблем.

Таким образом, «нейрон» обладает 1000 А-элементов, может запомнить 1800 объектов, и имеет три низкоуровневые функции сравнения.

### 3.3. Алгоритм детектора движения (ДД)

Задача решается с помощью детерминированного алгоритма. На вход подаются три статических изображения размером 640x480 и 256 цветовой гаммой (рис.1), полученные с помощью видеокамеры. Детектор движения находит движущийся объект и подает его на вход системы. Как видно, объект может быть частично поврежден из-за помех в статических изображениях (например, тенями, отбрасываемыми объектом во время передвижения). Более хорошего изображения можно добиться, если увеличить число анализируемых статических изображений, т.е. увеличить время наблюдения за объектом.

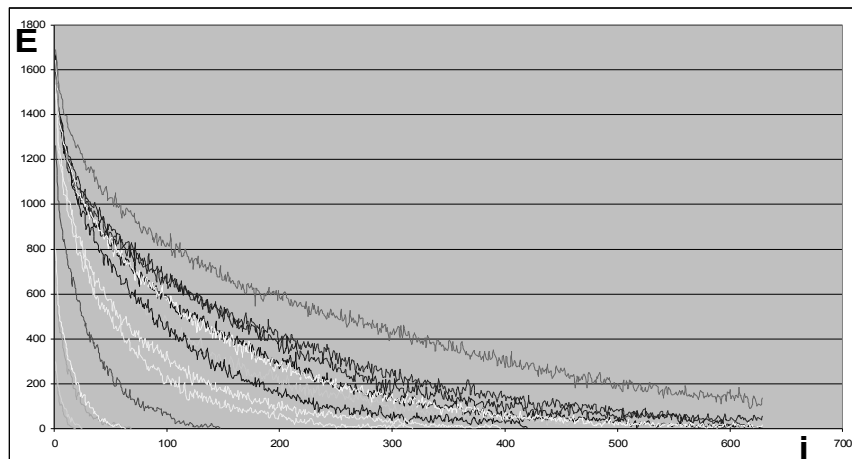
### 3.4. Задача восстановления

*Задачи 2 и 5* являются, соответственно, обучением нейронной сети и ее экзаменацией, где на вход подается случайная последовательность координат точек  $(X_i, Y_i)$ , а на выходе имеем воссозданный образ объекта.

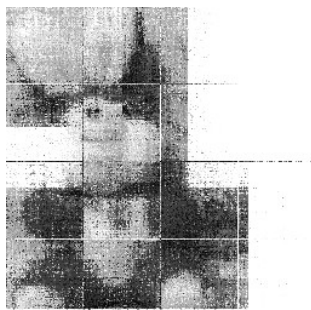


Различия в сходимости 16-ти «нейронов» возникают из-за разной насыщенности и сложности соответствующего участка изображения. Это видно на рис.1 блок – 16 «Neuron», где сходимости «нейронов» изображены соответственно области изображения, которую они контролируют. Таким образом, характеристика сходимости позволяет судить о сложности соответствующих областей объекта и его контурах. Это может быть использовано в задаче анализа, когда нужно найти одинаковые объекты.

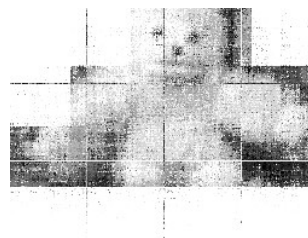
В процессе экзаменационной проверки были получены результаты, представленные на рис. 3–6. Таким образом, задача восстановления решена, причем следует отметить, что скорость восстановления  $\approx 30$  сек., а качество восстановления – удовлетворительное, то есть не хуже, чем дают другие классические методы.



**Рис. 2: Экспоненциальный вид сходимости при обучении для всех «нейронов».**  
E – количество ошибок, i – номер итерации



**Рис. 3: Результат эксперимента № 1**



**Рис. 4: Результат эксперимента № 2**

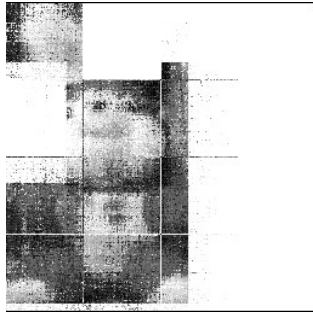


Рис. 5: Результат эксперимента № 3

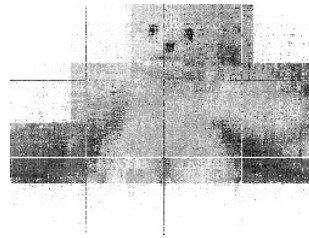


Рис. 6: Результат эксперимента № 4

### 3.5. Задача анализа

*Задачи 3 и 4* являются также соответственно обучением нейронной сети и ее экзаменацией, но здесь на вход подается цветовая сетка изображения, а на выходе получаем номер объекта (что это за объект) и соответствующую случайную последовательность точек, а точнее, коэффициент (размером в 32 бита), подавая который в функцию генерации случайных чисел, получаем случайную последовательность точек. Объединяя задачи 2, 3, 4, 5 графически, получим систему - рис.7.

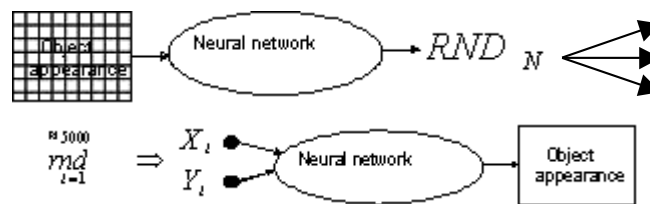


Рис. 7. Графическое представление задач восстановления и анализа

Кроме значения цвета, на вход подается еще расположение данной точки (x,y). Без этого нейронная сеть обладала бы только информацией об общей цветовой гамме изображения, а для различения объектов такой информации хватило бы только при небольшом количестве объектов в памяти. Качество распознавания объектов также зависит от количества точек, поданных с показываемого изображения. Максимальным является размер сетки 16x16 – для области 64x64 (контролируемой одним нейроном), т.е. каждая четвертая точка.

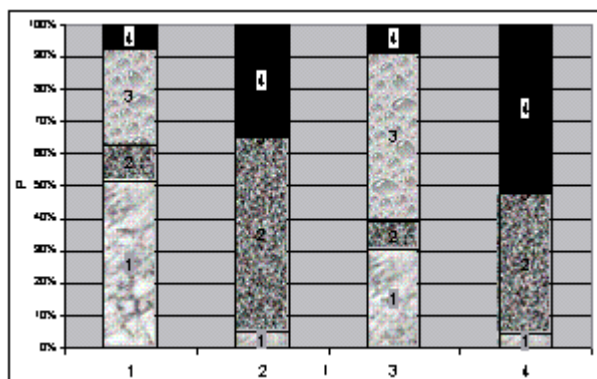


Рис. 8. Отнесение "нейроном" показываемого изображения к запомненным (Цветовая сетка 4x4)

Вероятность отнесения какого-либо объекта к одному из запомненных рассчитывается, исходя из того, сколько из 16 нейронов высказались в пользу того или иного объекта. Каждый нейрон исходит из того числа совпадений, которое получает при сравнении в своей области. На рис.8 видно, что при сетке 4x4 линий при показе первого изображения система из 16 нейронов приходит к выводу, что это изображение равноценно первому с  $\approx 50\%$  вероятностью по сравнению с другими, и третьему  $\approx 30\%$  вероятности, а 2-ому и 4-ому менее 5%. Это означает, что 1-е и 3-е изображения сходны, что и было в эксперименте – на 1-ом и 3-ем изображении был один и тот же объект “заяц”, но на разном фоне. Аналогичные результаты и для 2-ого и 4-ого изображения, в которых был объект “обезьяна”, но тоже на разных фонах.

#### 4. Заключение

Подводя итоги, можно сказать, что получена высокая скорость распознавания без заметного ухудшения качества изображения при восстановлении. Предложенный подход позволяет находить и выделять из фона движущиеся объекты. Заложены основы системы, которая выполняет низкоуровневые операции при идентификации движущихся объектов. Это позволяет в дальнейшем проектировать более сложные системы, в которых система РОДПИ является базовой системой, при решении элементарных операций по восстановлению и анализу объектов, находящихся под видео наблюдением.

#### 5. Список литературы

1. Розенблатт Ф. (1965). *Принципы нейродинамики (Перцептроны и теория механизмов мозга)*, Мир, Москва.



2. Jakovlev S. (2000). Efficiency analysis of a perceptron model in multiparameter classification tasks. *Scientific proceedings of Riga Technical University*, Issue 5, Vol.2, RTU, Riga, P. 93-99.
3. Subhash C. Kak. (1998). On generalization by neural networks, *Information Sciences*, 111, P.293-302.
4. Яковлев С. (2002). Способ математического расчета уровня активации нейронной сети – ССМ. *Proceedings of the International Conference "Traditions and Innovations in Sustainable Development of Society. Information Technologies"*, Rezekne, February 28 – March 2, P. 187-194.■

© Сергей Яковлев, (2004). Система распознавания движущихся объектов на базе искусственных нейронных сетей. ИТК НАНБ, Минск.