

э л е к т р о н н ы й ж у р н а л

МОЛОДЕЖНЫЙ НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК

Издатель: ФГБОУ ВПО «Московский государственный технический университет им. Н.Э. Баумана»

Реализация программного модуля распознавания изображений на основе самоорганизующихся карт Кохонена средствами системы matlab

Студент

Кафедра «Системы автоматического управления»: К.В. Парфентьев

Научный руководитель: А.И. Гаврилов,

к. т. н., доцент кафедры «Системы автоматического управления»

УДК 00.004.8

05.13.00

МГТУ им. Н.Э. Баумана
rudefellow@gmail.com

Для успешного решения задачи распознавания объектов зачастую применяются нейросетевые подходы, имитирующие работу мозга человека. Один из таких подходов использует самоорганизующиеся карты Кохонена (SOM)[1]. Идея сети предложена финским учёным Т. Кохоненом.

Стратегия метода построения SOM заключается в представлении множества экспериментальных данных в виде топологической карты – решетки размерностью $n \times m$, состоящей из нейронов. Каждый нейрон представляет собой вектор, определяющий в d – мерном пространстве признаков некоторую характерную точку. В результате реализации алгоритма обучения вектора-нейроны распределяются в пространстве признаков, аппроксимируя статистические взаимосвязи между данными обучающего множества.

Рассмотрим общую формулу алгоритма обучения:

Пусть задан набор векторов $M = \{m_i\}$ $i = \overline{1 \dots n}$, где m_i — “нейроны” карты (синаптические весовые вектора), он задан на некоторой регулярной сетке, задающей топологию, на которую будут отображаться данные выборки.

Введем в пространстве векторов меру близости $d(x, y)$. Выбирая из обучающего множества вектора данных, найдем вектор-нейрон карты m_c , для которого выполняется условие $d(x_j, m_c) = \min (d(x_j, m_i))$. Этот нейрон и будет, пользуясь терминологией нейронных сетей, нейроном-“победителем” (в литературе по SOM, как правило, используется специальный термин “Best matching unit” — BMU).

Произведем настройку нейронов карты по рекурсивной формуле:
$$m_i(t+1) = m_i + \alpha(t)h_{ci}(t)[x(t) - m_i(t)],$$
 где $h_{ci}(t)$ — симметричная невозрастающая функция, определенная вокруг m_c , называемая “функцией близости” (она определяет степень влияния нейрона — “победителя” на соседние нейроны), $\alpha(t)$ — коэффициент обучения (убывающая функция времени), причем $0 < \alpha(t) < 1$.

Таким образом, на каждом шаге алгоритма обучения определяется ближайший (в силу заданной метрики) к поступившим на вход сети данным нейрон, который

«подтягивается» к входным данным вместе со своими соседями, определяемыми функцией близости $h_{ci}(t)$. Графическая интерпретация функционирования алгоритма обучения представлена на рис.1.

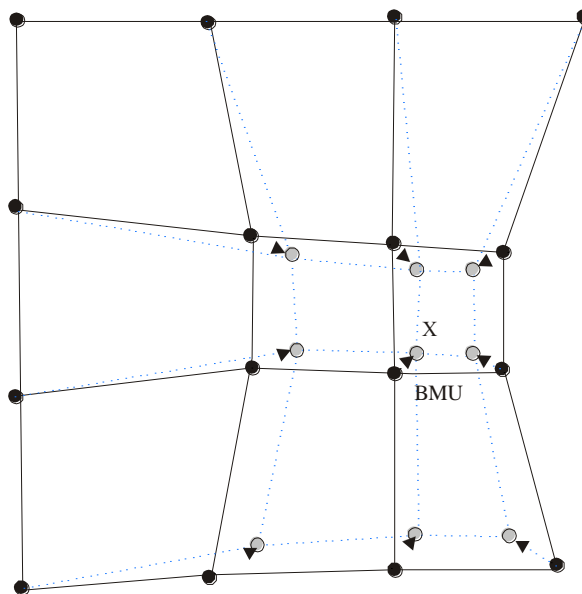


Рис.1. Графическая интерпретация основной идеи алгоритма обучения SOM

Сети Кохонена позволяют эффективно проводить кластеризацию и визуализировать полученные результаты, что дает возможность проводить достаточно глубокий анализ как системы в целом, так и отдельных ее параметров. Кроме того, в отличие от многослойных сетей перцептронного типа, SOM избавлены от попадания в локальные минимумы при обучении [1], не требуют априорной информации о системе, позволяют отслеживать появления ошибок и нежелательных режимов в системе. Способность самоорганизующихся сетей выявлять статистические зависимости на пространстве входных векторов в результате обучения на множестве примеров позволяет эффективно решать данную проблему.

Метод построения модели системы на основе SOM заключается в представлении множества экспериментальных данных в виде топологической карты – решетки размерностью $n \times m$, состоящей из нейронов. Каждый нейрон представляет собой вектор, определяющий в d – мерном пространстве признаков некоторую характерную точку. В результате реализации алгоритма обучения вектора-нейроны распределяются в пространстве признаков, аппроксимируя статистические взаимосвязи между данными обучающего множества.

При использовании SOM для анализа (мониторинга) различных систем и процессов необходимо реализовать следующие процедуры:

- сбор данных (измеряемых входных, выходных сигналов и характеристик процесса);
 - предварительная обработка данных (фильтрация);
 - выделение признаков с целью наилучшего определения проблемной среды с позиций анализа;
 - нормализация (взвешивание);
 - обучение и “разметка” карты;
 - подтверждение работоспособности (принятие решения об адекватности модели);
- визуализация.

Причем выполнение этапов 1,5,6,7 является обязательным, остальных – опциональным.

В качестве программного средства реализации использован пакет Matlab R2010a и тулбокс SOM, доступный на сайте Хельсинского технического университета[2]. В первом

приближении в качестве векторов рассмотрена последовательность строк пикселей. Для решения нашей задачи использована база данных COIL-20[3]. Все фотографии в этой базе уже изначально полутоновые, сняты при однородном освещении, имеют высокий контраст. Поэтому кроме сжатия до 32 на 32 пикселей мы обрабатывать фотографии не будем. Таким образом, каждый вектор будет содержать 1024 координаты.

База данных содержит фотографии каждого из предметов (по 72 на объект) в разных ракурсах (рис.3). Это позволяет сети «запомнить» объект. Поскольку помимо получения статистики необходимо сравнить несколько подходов, задействуется только каждая четвертая фотография при обучении, остальные ракурсы не используются. Система обучается пяти объектам, по 18 фотографий каждый.

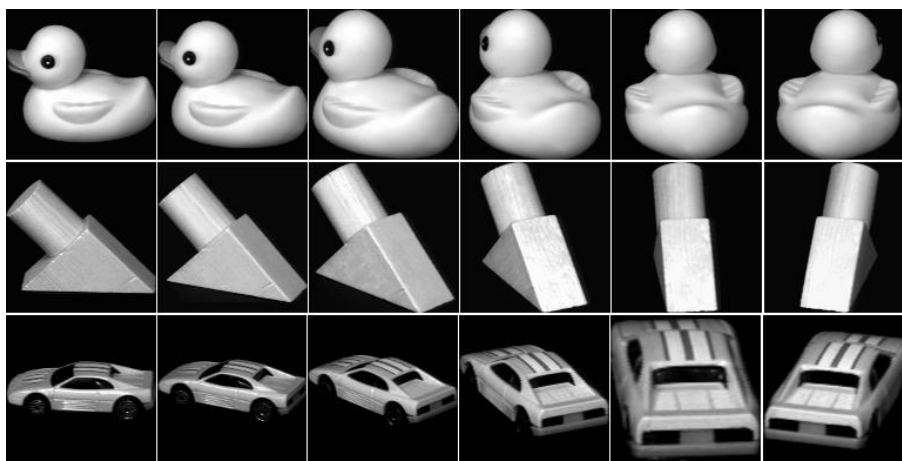


Рис.3. Примеры первых 6 фотографий для каждого предмета

Для повышения качества распознавание применяются преобразование Фурье, синусное и косинусное преобразование. Они позволяют перевести полученные координаты яркостей в частотную область. Также эти преобразования могут сократить количество существенной информации. По этим причинам подобные преобразования нашли применение в сжатии файлов.

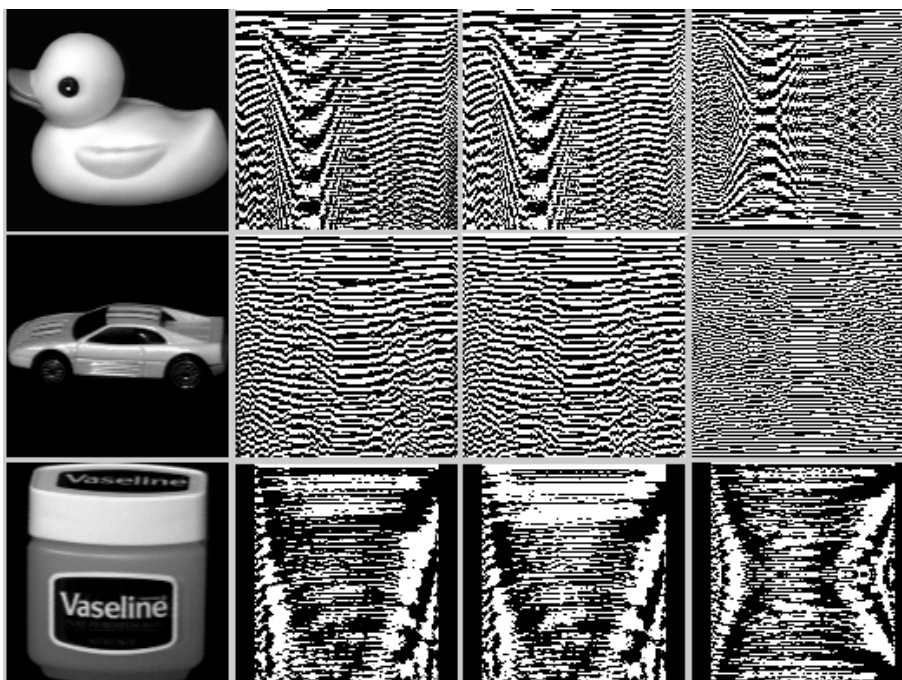
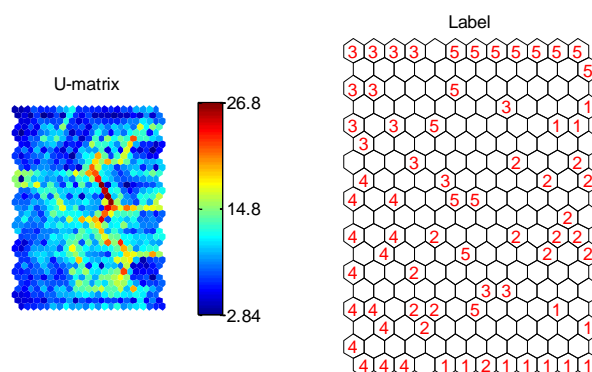


Рис.4. Слева направо: оригинал, косинусное преобразование, синусное преобразование, преобразование Фурье

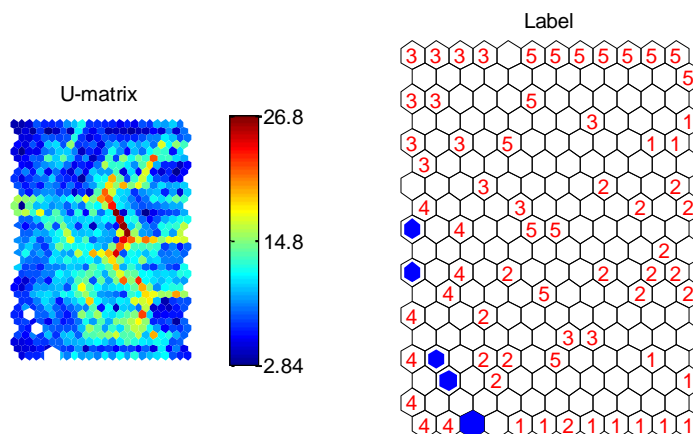
Создание карты размером 18 на 12:



SOM 16-Nov-2011

Рис.5. Карта размером 18 на 12

Следующим этапом является тестирование полученной карты. Для этого были задействованы по 6 фотографий, не входящих в обучающую выборку, для каждого из 5 объектов. Также были задействованы 6 фотографий 5 объектов, не входящих в обучающую выборку. Пример попадания на карту 18 на 12 тестовых фотографий объекта 4 представлен на рис. 6.



SOM 16-Nov-2011

Рис.6. Пример тестирования карты размером 18 на 12

Получены следующие результаты: ложные срабатывания - 0 из 30. Результаты распознавания объектов, входивших в обучение – 26 из 30 или 87%.

Для улучшения результатов используется модифицированная технология обучения. Помимо вектора пикселей размерностью 1024 добавляется индекс объекта (1,2,3 ...). Определять принадлежность фотографии тому или иному объекту будем с помощью bmi

(best matching unit) – нейрона, находящегося наиболее близко к данному вектору. Введем ограничение на расстояние от тестируемого вектора до соответствующего bmu. Такая технология позволяет создавать карты небольшого размера без ошибок при обучении.

При тестировании карты получены следующие результаты:

Карта 27 на 3: 27 правильных срабатываний из 30 (90%) и 0 ложных срабатываний из 30 (0%).

Карта 24 на 4: 28 правильных срабатываний из 30 (93%) и 1 ложное срабатывание из 30 (3%).

Проведем предыдущий эксперимент(с использованием bmu) с применением преобразования Фурье и синусных, косинусных преобразований.

Косинусное преобразование:

Карта размером 19 на 9 нейронов: 28 правильных срабатываний из 30 (93%) и 0 ложных срабатываний из 30 (0%).

Карта размером 13 на 11 нейронов: 30 правильных срабатываний из 30 (100%) и 1 ложное срабатывание из 30 (0%).

Синусное преобразование:

Карта размером 16 на 8 нейронов: 29 правильных срабатываний из 30 (97%) и 1 ложное срабатывание из 30 (3%).

Преобразование Фурье:

Карта размером 13 на 5 нейронов: 13 правильных срабатываний из 30 (43%) и 1 ложное срабатывание из 30 (3%).

Как видно из результатов, задача распознавания была решена успешно. При моделировании использовались два метода распознавания: с разметкой и без разметки. Тестирование с разметкой карты показало более высокие результаты (ошибка распознавания уменьшилась примерно на 3%). Синусное и косинусное преобразование проявили себя в этой задаче хорошо. Во всех экспериментах они показали более высокие (примерно на 3-6%) результаты, чем тестирование без преобразований. Касаясь их непосредственного сравнения, косинусное проявило себя немного лучше, чем синусное (в среднем на 3%). Именно с его использованием удалось достичь наиболее высоких результатов и сократить ошибки до минимума. В данной постановке задачи преобразование Фурье привело к снижению вероятности успешного распознавания до 43 %.

Основные результаты работы

- Рассмотрены математическое описание, алгоритмы обучения и принцип работы самоорганизующихся карт Кохонена.
- Решена задача распознавания предметов при помощи самоорганизующихся карт Кохонена.
- Повышена эффективность распознавания за счет применения синусного и косинусного преобразований.

Литература

1. Kohonen T. Self-Organization and Associative Memory. - New-York: Springer - Verlag, 1988. – 620 p.
2. Kohonen T., Hynninen J., Kangas J., Laaksonen J. [SOM_PAK: The Self-Organizing Map Program Package](#). Technical Report A31, Helsinki University of Technology, Laboratory of Computer and Information Science, FIN-02150 Espoo, Finland, 1996.
3. Nene S. A., Nayar S. K., Murase H., - Columbia Object Image Library (COIL-20), Technical Report CUCS-005-96, February 1996.