

Тант Зин Пьо

Национальный исследовательский университет “МИЭТ”
Аспирант кафедры “Вычислительная техника”

Thant Zin Phyo

*National Research University of Electronic Technology
Postgraduate student of department of “Computer Engineering”*

E-Mail: thantzin.miet@gmail.com

Тин Чжо

Национальный исследовательский университет “МИЭТ”
Аспирант кафедры “Вычислительная техника”

Htin Kyaw

*National Research University of Electronic Technology
Postgraduate student of department of “Computer Engineering”*

E-Mail: thantzin.miet@gmail.com

Пья Сон Ко Ко

Санкт-Петербургский Государственный Морской Технический Университет
Аспирант кафедры вычислительной техники и информационных технологий

Pyae Sone Ko Ko

St.Petersburg State Marine Technical University

Postgraduate student of department of computer engineering and information technology

E-Mail: thantzin.miet@gmail.com

Пайе Тэйн Наинга

Национальный исследовательский университет “МИЭТ”
Аспирант кафедры “Система автоматического управления и контроля”

Pyae Thein Naing

*National Research University of Electronic Technology
Postgraduate student of department of the automatic control and monitoring*

E-Mail: thantzin.miet@gmail.com

05.13.01 – Системный анализ, управление и обработка информации

**Методика системы распознавания образов с помощью
самоорганизующихся карт Кохонена нейронных сетей на основе Matlab**

Methods of image recognition system using self-organizing maps Kohonen neural
networks on base of Matlab

Аннотация: В статье рассматриваются методики распознавания разнообразных образов на изображении с помощью нейронных сетей (самоорганизующиеся карты Кохонена), которые используют для проектирования многомерных данных, аппроксимации плотности и кластеризации. Параметры сети включают в себя размерность массива нейронов, число нейронов в каждом измерении, форму окрестности, закон сжатия окрестности и скорость обучения.

Abstract: The article discusses techniques for recognizing variety of images using by neural networks (Kohonen self-organizing maps), which are used for the design of multi-dimensional data density approximation and clustering. Network settings include the dimension of an array of neurons, the number of neurons in each dimension, shape neighborhood, the neighborhood of the law of compression and speed of learning.

Ключевые слова: Распознавание образов; нейронные сети; самоорганизующиеся карты; обучение образов; преобразование; обработка изображений.

Keywords: Pattern recognition; neural networks; self-organizing maps; training image; conversion; image processing.

Введение

Самоорганизующиеся карты (*Self Organizing Maps – SOM*) представляют собой нейронные сети, обучаемые без учителя. Они используются для классификации и распознавания образов, которые могут быть использованы на многих этапах обработки изображений [1- 3]. Алгоритм SOM был реализован в программе MATLAB с различными дополнительными параметрами, позволяющими настройку модели в соответствии с требованиями пользователя. Для более легкого применения SOM был разработан графический *интерфейс пользователя (GUI)*.

На рис. 1 предлагается блок-схема системы распознавания образов. На первом этапе, вычисляется 2D-DCT (двумерное дискретное косинусное преобразование) для каждого изображения, и из коэффициентов *DCT* (дискретное косинусное преобразование) образуются векторы признаков. На втором этапе, SOM (самоорганизующаяся карта) используется методом «Обучение без учителя» для классификации векторов на группы, чтобы распознавать объект, который во входном изображении присутствует или нет в базе данных изображений. Если объект классифицируется как присутствующий, то изображение наилучшего совпадения, найденного в базе данных обучения, отображается как результат, иначе результат показывает, что объект не найден в базе данных изображений [4,5].

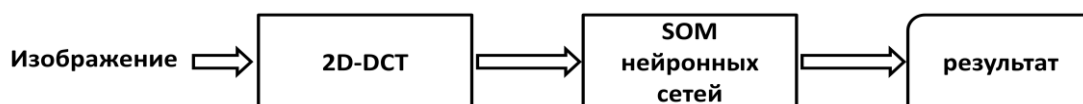


Рис. 1. Предложенный метод для системы распознавания образов

Дискретное косинусное преобразование

DCT представляет изображение в виде суммы синусоид с различной амплитудой и частотой. Функция *dct2* в приложении инструментария обработки изображений реализует 2D-DCT изображений. Одна из особенностей дискретного преобразования Фурье состоит в том, что некоторые локальные участки изображения можно охарактеризовать небольшим количеством коэффициентов дискретного преобразования Фурье. Это свойство очень часто используется при разработке методов сжатия изображений. Например, DCT является основой международного стандарта, который используется в алгоритме сжатия изображений с потерями JPEG. Название формата “JPEG” состоит из первых букв названия рабочей группы, которая принимала участие в разработке этого стандарта (Joint Photographic Experts Group).

DCT обладает тем свойством, что для типичного изображения большинство визуально значимая информация об изображении сосредоточена в нескольких коэффициентах. Извлеченные коэффициенты DCT могут быть использованы как тип подписи, которая является полезной для задач распознавания, таких как распознавание образов и лиц [4, 5].

2D-DCT матрицы A с размерами $M \times N$ определяется согласно следующему выражению

$$B_{pq} = \alpha_p \alpha_q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A_{mn} \cos\left(\frac{\pi(2m+1)p}{2M}\right) \cos\left(\frac{\pi(2n+1)q}{2N}\right), \quad (1)$$

где $0 \leq p \leq M-1$ и $0 \leq q \leq N-1$;

Значения B_{pq} называют коэффициентами DCT матрицы A . DCT - обратимое преобразование, и 2D-IDCT (двумерное обратное дискретное косинусное преобразование) определяется следующим образом:

$$A_{mn} = \sum_{p=0}^{M-1} \sum_{q=0}^{N-1} \alpha_p \alpha_q B_{pq} \cos\left(\frac{\pi(2m+1)p}{2M}\right) \cos\left(\frac{\pi(2n+1)q}{2N}\right), \quad (2)$$

где $0 \leq m \leq M-1$ и $0 \leq n \leq N-1$;

Значения α_p и α_q в (1) и (2) имеют вид:

$$\alpha_p = \begin{cases} \frac{1}{\sqrt{M}}, & p = 0 \\ \sqrt{\frac{2}{M}}, & 1 \leq p \leq M-1 \end{cases} \quad \alpha_q = \begin{cases} \frac{1}{\sqrt{N}}, & q = 0 \\ \sqrt{\frac{2}{N}}, & 1 \leq q \leq N-1 \end{cases} \quad (3)$$

Предлагаемый метод использует матрицу преобразования DCT в MATLAB окне инструментария обработки изображений. Этот метод является эффективным для небольших квадратных входов, таких как блоки изображения с размером 8×8 пикселей. $M \times M$ матрица преобразования T определяется по формуле:

$$T_{pq} = \begin{cases} \frac{1}{\sqrt{M}} & p = 0, 0 \leq q \leq M-1 \\ \sqrt{\frac{2}{M}} \cos \frac{\pi(2q+1)p}{2M} & 1 \leq p \leq M-1, 0 \leq q \leq M-1 \end{cases} \quad (4)$$

Предварительная обработка изображений

Входные изображения предварительно обрабатываются в MATLAB. Процесс изготовления изображения показан на рис.2. Предварительная обработка изображений включает в себя следующие этапы:

- уровни автоматической регулировки оттенка и насыщения;
- регулировка яркости и отличие от фиксированного масштаба;
- разбавление 24 бит RGB цвета в 8-битные оттенки серого;
- сокращение размера изображений до 512×512 пикселей;
- сохранение изображений в формате JPEG.

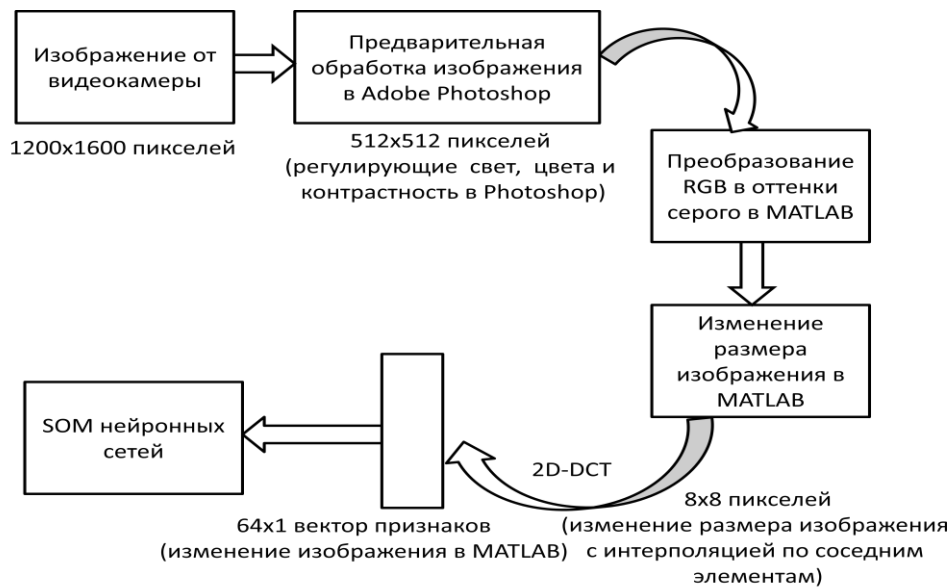


Рис. 2. Процесс изготовления изображения

2D-DCT сжатие изображений

Интерполяция по соседним элементам выполняется с помощью MATLAB Image Processing Toolbox для изменения размера предварительных изображений из размера 512×512 пикселей до блоков изображений с размером 8×8 пикселей, как показано на рис. 2. Предлагаемый метод проектирования рассчитывает 2D-DCT блоков изображений размера 8×8 пикселей с использованием '8' из 64 коэффициентов DCT для маскировки. Остальные 56 коэффициентов отбрасываются (устанавливать на нуль). Затем изображение реконструируется путем вычисления 2D-IDCT каждого блока с использованием DCT метода расчета матрицы преобразования (DCT transform matrix computation method). Наконец, на выходе получается набор массивов. Каждый массив имеет размер 8×8 пикселей и представляет собой одно изображение. Эти этапы представлены на рис.3 для типового изображения. Эмпирически, в левом верхнем углу каждая 2D-DCT матрица содержит наиболее важные данные, потому что они соответствуют низкочастотным компонентам в обработанном блоке изображения [6,7].

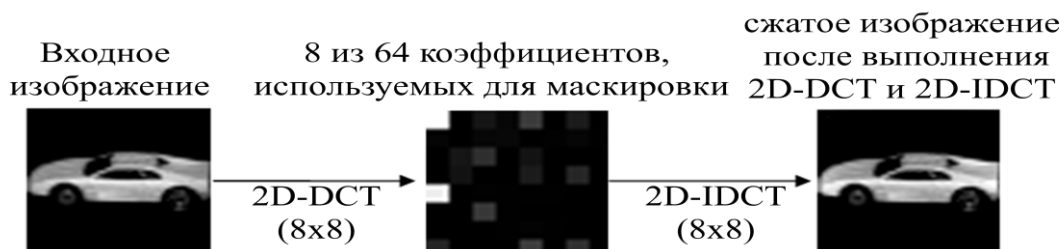


Рис. 3. Вычисление 2D-DCT изображения

Распознавание образов

Распознавание образов является одним из важных шагов в обработке изображений и из немногих биометрических методов, которые обладают достоинствами, как высокой точности, так и низкого вмешательства. Если есть образец, первый шаг в распознавании образов заключается в выборе набора признаков или атрибутов из всех доступных признаков, которые будут использоваться для классификации образца. Далее, исходный образец должен быть преобразован в представление, которым легко манипулировать программно. После того, как данные обработаны с целью удаления шумов, находят признаки в данных как относящиеся к соответствующему образцу. На стадии классификации данные классифицируются на основе измерений сходства с другими образцами. Процесс распознавания образов заканчивается, когда данным на основе их принадлежности к классу назначается метка. В системе распознавания образов, у нас есть база данных изображений, хранящихся в системе. Когда получили новое изображение, оно сравнивается с базой данных изображений, уже хранящихся или не хранящихся в системе. Сначала в эксперименте строилась база данных различных образов, как показано на рис.4 и блок-схема распознавания образов на рис.5, SOM работает в двух режимах: обучение и отображение.

- Обучение строит карту с помощью входных примеров. Это конкурентный процесс, называемый также квантизацией векторов.

- Отображение автоматически классифицирует новый входной вектор.

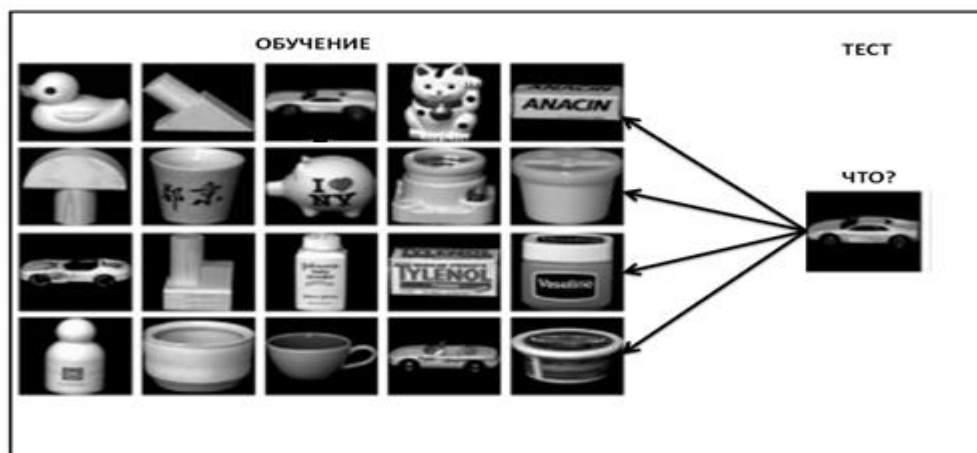
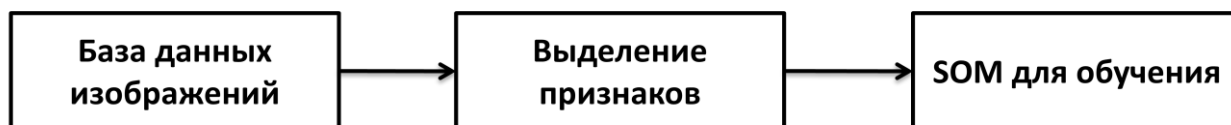


Рис. 4. Сравнение нового изображения с базой данных

А. Обучение



Б. Отображение

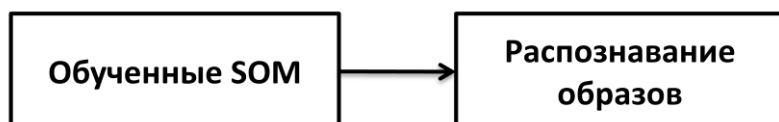


Рис. 5. Блок-схема распознавания образов

Самоорганизующиеся карты

Самоорганизующиеся карты (Self Organizing Maps – SOM) также известны как карта Кохонена. это одна из разновидностей нейросетевых алгоритмов. Основным отличием данной технологии от рассмотренных нами ранее нейросетей, обучаемых по алгоритму обратного распространения, является то, что при обучении используется метод обучения без учителя, то есть результат обучения зависит только от структуры входных данных. В нейронных алгоритмах обучения без учителя и SOM Кохонена широко использованы в области распознавания образов. Важным отличием алгоритма SOM является то, что в нем все нейроны (узлы, центры классов...) упорядочены в некоторую структуру (обычно двумерную сетку). При этом в ходе обучения модифицируется не только нейрон-победитель, но и его соседи, но в меньшей степени. Все нейроны в некоторой окрестности нейрон-победителя обновляются с использованием правила Кохонена. Правило Кохонена позволяет веса нейронов для обучения входного вектора, и из-за этого оно полезно в приложениях распознавания. Поэтому в этой системе SOM используется для классификации векторов DCT на группы, чтобы распознавать объект во входном изображении присутствует или нет в базе данных изображений [7]. SOM может быть одномерной, двумерной или многомерной. Количество входных соединений в сети SOM зависит от количества атрибутов, которые будут использоваться в классификации [8].

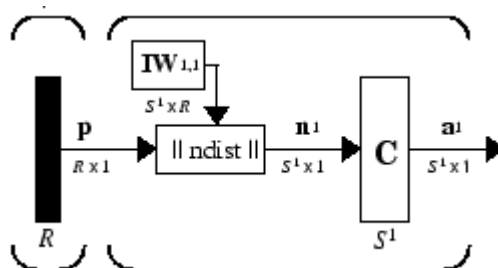


Рис. 6. Архитектура SOM

На рис. 6 входным вектором \mathbf{P} является строка пикселей DCT сжатия изображения. Блок $\|\text{dist}\|$ принимает входной вектор \mathbf{P} и матрицу входных весов $\mathbf{IW}^{1,1}$, которая производит вектор, имеющий S^1 элементов. Элементы являются отрицательными расстояниями между входным вектором и векторы $\mathbf{IW}^{1,1}$ формируются из строки матрицы входного веса. Блок

$\|dist\|$ вычисляет входную сеть n^1 конкурентного слоя путем нахождения Евклидова расстояния между входным вектором P и весовыми векторами. Функция конкурентной передачи C принимает сеть входного вектора для слоя и возвращается выходные нейроны 0 для всех нейронов за исключением победителя, нейрон связан с самым положительным элементом входной сети n^1 . Выход победителя – 1 . Весовой вектор нейрона ближе к входному вектору и имеет наименее отрицательную сеть входа, поэтому для выхода победитель – 1 . Таким образом, функция конкурентной передачи C производит 1 на выходе элемента a_i^1 , соответствующего i^* «победитель». Все другие элементы выхода в $a^1 - 0$ [3].

$$n^1 = -\|IW_{1,1} - p\|$$

$$a^1 = \text{compet}(n^1)$$

SOM подразумевает использование упорядоченной структуры нейронов. Обычно используются одно и двумерные сетки. При этом каждый нейрон представляет собой n -мерный вектор-столбец $w = [w_1, w_2, \dots, w_n]^T$, где n определяется размерностью исходного пространства (размерностью входных векторов).

Обучение без учителя

Во время обучения, нейрон с весами, наиболее близкими к вектору входных данных, объявляется победителем. Тогда веса всех нейронов в окрестности нейрона-победителя корректируются на величину, обратно пропорционально Евклидову расстоянию. Кластеры и классификаторы набора данных используются на основе набора атрибутов.

Алгоритм обучения можно резюмировать следующим образом :

Инициализация (Ш0): Выбор случайных значений для начальных весовых векторов $w_j(0)$, весовые векторы, различные для $j = 1, 2, \dots, l$ где l – это общее число нейронов.

$$w_i = [w_{i1}, w_{i2}, \dots, w_{il}]^T \in \mathcal{R}^n \quad (1)$$

Выборка (Ш1): выборка x из входного пространства с определенной вероятностью.

$$x = [x_1, x_2, \dots, x_l]^T \in \mathcal{R}^n \quad (2)$$

Сходство в соответствие (Ш2): Найти нейрон-победитель $i(x)$ в момент времени t , $0 < t \leq n$ с помощью минимального критерия Евклидова расстояния:

$$i(x) = \arg \min_j \|x(n) - w_j\|, j = 1, 2, \dots, l \quad (3)$$

Обновление (Ш3): настройка весового вектора всех нейронов с помощью формулы обновления:

$$w_j(n+1) = w_j(n) + \eta(n)h_{j,i}x(n)(x(n) - w_j(n)), \quad (4)$$

где $\eta(n)$ является параметр коэффициента обучения, и $h_{j,i}x(n)$ – функция окрестности (neighborhood function) сосредоточена вокруг нейрона-победителя $i(x)$. Как $\eta(n)$, так и $h_{j,i}x(n)$ изменяются динамически в процессе обучения для достижения наилучших результатов.

Ш4. Повторяется шаг Ш1 до тех пор, пока выходные значения не наблюдается в карте признаков.

Обучение изображений отображается в меньшей размерности с использованием сети SOM и матрицы весов каждого изображения, храненного в базе данных обучения. При распознавании обученные изображения восстановлены с использованием весовых матриц.

Распознавание осуществляется через необученные тестовые изображения с использованием евклидова расстояния как мера сходства. Обучение и тестирование для нашей системы проводились с использованием MATLAB Neural Network Toolbox.

Этап обучения

На этапе обучения, обозначенные DCT-векторы, которые представляются SOM по одному. Для каждого узла, число "победитель" записано вместе с меткой входной выборки. Для каждого узла, число "победы" записано вместе с меткой входной выборки. Для узлов весовые векторы обновляются, как описано в этапе обучения без учителя. В конце этого этапа, каждый узел SOM имеет два записанных значений: суммарное число побеждающих раз для объекта, присутствующего в базе данных изображений и суммарное число побеждающих раз для объекта, не присутствующего в базе данных изображений [6].

Этап тестирования

На этапе тестирования, каждый входной вектор сравнивается со всеми узлами SOM и нейрон-победитель найден на основе минимального евклидова расстояния, как указано в этапе Ш2 [6]. Окончательный выход системы на основе её распознавания отображает, если тестовое изображение присутствует или не присутствует в базе данных изображений.

База данных изображения разделяется на два подмножества: обучение и тестирование. Во время обучения SOM использовались 20 изображений. Рис. 4 показывает базу данных обучения и тестирования изображение. Предварительно обработанные полутоновые изображения с размером 8×8 пикселей изменены в MATLAB, чтобы сформировать 64×1 массив с 64 строками и 1 столбцом для каждого изображения. Этот метод выполняется на всех 5 тестовых изображениях, чтобы сформировать входные данные для тестирования системы распознавания. В работе авторами получены экспериментальные результаты на основе MATLAB (рис.7).

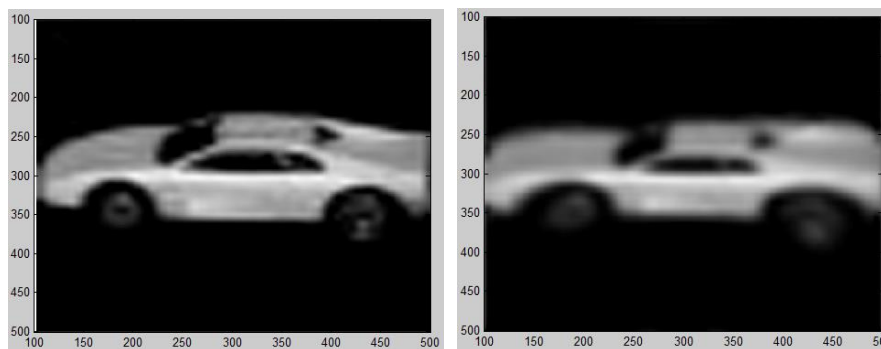


Рис. 7. Экспериментальные результаты на основе MATLAB

Заключение

Таким образом, данная работа посвящена применению нейросетевого алгоритма Кохонена для классификации образов без учителя, т.е. кластеризации. При такой постановке задачи дан набор объектов, каждому из которых сопоставлена строка таблицы (вектор значений признаков). Требуется разбить исходное множество на классы, т.е. для каждого объекта найти класс, которому он принадлежит. В результате получения новой информации о классах возможна коррекция существующих правил классификации объектов.

ЛИТЕРАТУРА

1. S. Samarasinghe. Neural Networks for Applied Sciences and Engineering: From Fundamentals to Complex Pattern Recognition. Auerbach Publications, 2006.
2. L. Ma, Y. Xiao, and K. Khorasani, "A new facial expression recognition technique using 2D DCT and k-means algorithm", in Proc. International Conference on Image Processing, Oct. 2004, pp. 1269-1272.
3. A. Abdallah, M. Abou El-Nasr, and A. Lynn Abbott, "A New Face Detection Technique using 2D DCT and Self Organizing Feature Map" in Proc. of World Academy of Science, Engineering and Technology, Vol. 21, May 2007, pp. 15-19.
4. J. Nagi, "Design of an Efficient High-speed Face Recognition System", Department of Electrical and Electronics Engineering, College of Engineering, Universiti Tenaga Nasional, March 2007.
5. D. Kumar, C.S. Rai, and S. Kumar, "Face Recognition using Self-Organizing Map and Principal Component Analysis" in Proc. on Neural Networks and Brain, ICNNB 2005, Vol. 3, Oct 2005, pp. 1469-1473.
6. Тант Зин Пью, Кочетков М.П. Обработка изображений в системах технического зрения робототехнических комплексов// Информационно-управляющие вычислительные системы: алгоритмы, аппаратные и программные средства: межвузовский сборник/ Под ред. В.А. Бархоткина. – М.: МИЭТ, 2011. – 148с. С. 130-135.
7. Тант Зин Пью. Адаптивные алгоритмы распознавания// Микроэлектроника и информатика – 2013. 20-я Всероссийская межвузовская научно-техническая конференция студентов и аспирантов: Тезисы докладов. – М.: МИЭТ, 2013г. – 323с. С. 221.

Рецензент: НИУ МИЭТ, д.т.н., профессор О.И. Лисов.