

УДК 681.324

С.А. Бутенков, Al Dhouyani Saud

Таганрогский государственный радиотехнический университет, г. Таганрог, Россия
saab@tsure.ru

Применение нейронных сетей для семантической интеллектуальной сегментации

Работа связана с развитием интеллектуального подхода к одной из важнейших задач систем технического зрения – сегментации исходного изображения на части, составляющие элементы объекта и фона. Предлагаемый подход к решению этой задачи основывается на систематическом введении методов мягких вычислений и имеет целью повышение коэффициента интеллектуальности систем машинного зрения. Работа состоит из двух частей, в первой из которых предлагается метод разделения пикселей изображений объекта и фона, основанный на использовании семантических пространств яркостей, а во второй – кластеризации полученных пикселей на основе нового типа нейронной сети.

Введение

Сегментация является важнейшим этапом интеллектуального анализа изображений и является ключевым моментом, определяющим уровень машинного интеллекта (MIQ по определению Л. Задэ [1]) для интеллектуальных систем технического зрения (СТЗ). Процедура сегментации заключается в отделении визуально важных частей изображения (*объекта интереса*) от незначачих (*фона изображения*) [2]. Полный процесс сегментации формально включает две основных задачи – отделение пикселей объекта от пикселей фона и кластеризацию на полученных множествах пикселей. Часто обе задачи решаются с помощью единого итерационного алгоритма.

Можно отметить, что выполнение сегментации связано с наличием высокой степени неопределенности, поскольку сами понятия объекта и фона являются сложными и не всегда могут быть выражены четкими понятиями [3], [4]. Свойства объекта и фона также часто плохо формулируются в терминах евклидовой геометрии [4].

С появлением основополагающих работ Л. Задэ [5], [6], в которых были провозглашены идеи *мягких вычислений*, были созданы теоретические основы разработки новых подходов к сегментации изображений, развитые в работах [1], [7], [8]. В этих работах была введена теоретическая база для построения алгоритмов нечеткой сегментации, а также предложены практические программные решения этой задачи.

Задачи обработки изображений относятся к классу сложных и вычислительно затратных вычислительных задач в силу сравнительно большого числа пикселей изображения и весьма сложных алгоритмов обработки [2]. Следовательно, имеется объективная необходимость, во-первых, в развитии методов гранулирования визуальной информации с целью уменьшения объема данных, необходимых для представления изображений, во-вторых – в развитии методов параллельных и нейровычислений для решения задач интеллектуального анализа изображений.

Целью работы является решение задач, связанных с разработкой новых методов грануляции и интеллектуальной обработки визуальной информации, ориентированных на использование нейросетевых вычислений. Это позволит в полной мере реализовать

все потенциальные преимущества мягких вычислений перед числовыми алгоритмами [3], [5], [6]. Отметим, что в развитии основных идей предыдущих работ по этой тематике [8] используется геометрический подход к гранулированию и обработке многомерной информации, реализующий основные идеи Л. Задэ о представлении информации нечеткими графами [3].

В качестве основной цели следует указать создание методологической базы для нейросетевых методов [7], перехода от алгоритмических методов гранулирования изображений [4] к наиболее полно реализующим преимущества мягкого подхода к обработке информации [5]. Важной задачей является также исследование вопроса о создании новых видов нейронных сетей, ориентированных на решение задач указанного класса на основе так называемых моделей «стеклянного ящика» [9].

Постановка задачи. Теоретической основой для развиваемого в наших работах подхода является геометрическое истолкование задач и методов гранулирования многомерной информации [4], [8], поэтому в качестве базиса для активационных функций подобных сетей естественно выбрать R-функции, введенные в работах В.Л. Рвачева [10] и использованные в наших работах для геометрического моделирования [11].

В работе разрабатываются методики кластеризации информационных элементов (в том числе и изображений), основанные на использовании перцептуального подхода к интеллектуальной обработке информации [12], [13].

Необходимо разработать также архитектуру и алгоритмы обучения нечетких нейросетей, ориентированных на решение геометрических задач [14], основанные на аналитических методах вычислительной геометрии [11].

Одномерное и пространственное гранулирование

С точки зрения теории информационной грануляции (ТИГ) Л. Задэ можно отметить, что оба этапа сегментации (отделение пикселей и их кластеризация) сводятся к грануляции, но в первом случае – к классической одномерной грануляции на одномерных семантических пространствах (СП) [1], а во втором – к пространственной грануляции на пикселях.

Поскольку в ТИГ идеология пространственной грануляции только начинает разрабатываться [4], [8], то здесь необходима разработка новых подходов.

Информационной гранулой называется подмножество универсума, на котором определено отношение сходства, неразличимости и т.п. [3]. Множество гранул, которое содержит все объекты универсума, называется *гранулированием универсума*. Подмножество $A \subseteq U$ называется *составной* (не элементарной) *гранулой*, если оно представляет собой объединение *атомарных гранул* [5].

Определив на плоскости проекции произвольной гранулы A , обозначаемые как $pr_x A$ и $pr_y A$, зададим *инкапсулирующую декартову гранулу* для произвольного A как $G^+ = pr_x A \times pr_y A$. Гранула G^+ является точной верхней гранью конечного множества всех гранул, содержащих A .

С понятием инкапсулирующей гранулы тесно связано фундаментальное понятие *аппроксимирующего графа отношения*. Согласно [5], граф на плоском множестве задается как $f^* = A_1^x \times A_1^y + \dots + A_n^x \times A_n^y = \sum_i A_i^x \times A_i^y$, $i = 1, \dots, n$, где операция «+» означает дизъюнкцию в широком смысле слова. Отметим, что в настоящей работе речь идет о декартовых координатах в отличие от лингвистических переменных по [1], [12].

Для применения методов топологии к анализу *непересекающихся* гранул изображения представим инкапсулирующую гранулу на плоскости в виде грасмановского примитивного элемента [13]:

$$G^+(g^1, g^2) = \begin{pmatrix} \min(x_1^1, x_1^2) & \min(y_1^1, y_1^2) & 1 \\ \min(x_2^1, x_2^2) & \max(y_1^1 + h^1, y_1^2 + h^2) & 1 \\ \max(x_1^1 + w^1, x_1^2 + w^2) & \min(y_1^1, y_1^2) & 1 \end{pmatrix}, \quad (1)$$

где параметры инкапсулирующей гранулы G^+ определяются как

$$\begin{aligned} x^+ &= \min(x_1^1, x_1^2), \quad y^+ = \min(y_1^1, y_1^2), \\ w^+ &= \max(x_1^1 + w^1, x_1^2 + w^2) - \min(x_1^1, x_1^2), \\ h^+ &= \max(y_1^1 + h^1, y_1^2 + h^2) - \min(y_1^1, y_1^2). \end{aligned}$$

Методы гранулирования изображений с сохранением энтропии исходного изображения были разработаны в ряде наших предыдущих работ [4], [8] и др., поэтому здесь мы не будем их описывать.

Переходя к бинарным отношениям на множестве гранул изображения, отметим, что основой кластеризации являются отношения эквивалентности или сходства.

Для точечных объектов эти отношения представляются количественно с помощью метрик (например, евклидовой) и с помощью скалярного произведения радиусов-векторов. Для двумерных объектов имеют место проблемы при построении транзитивных отношений [12], поэтому нами предлагается перцептуальное отношение сходства для двумерных объектов [13].

Отношения сходства и иерархическая кластеризация

В предыдущих работах нами были разработаны новые типы отношений сходства на декартовых гранулах [13], позволяющие строить процедуры генерализации для гранулированной информации, в том числе и изображений. Кластеризация представлена в [12] как преобразование данного взвешенного нечеткого отношения сходства [6] в нечеткое отношение эквивалентности, которое определяет иерархию нечетких отношений эквивалентности и, как результат, иерархию разбиений множества объектов на классы эквивалентности. Основной целью введения иерархии является сохранение классов сходства.

Пусть задано множество пространственных гранул X универсума U и $S: X \times X \rightarrow [0,1]$ – нечеткое отношение (сходства), т.е. S удовлетворяет на $X \times X$ условиям:

- симметричности $S(g_1, g_2) = S(g_2, g_1) \quad \forall g_1, g_2 \in X \times X$;
- рефлексивности $S(g_1, g_1) \geq S(g_2, g_3) \quad \forall g_1, g_2, g_3 \in X \times X$.

Отношения сходства, удовлетворяющие условию (\vee, \wedge) -транзитивности $S(g_1, g_2) \geq \min\{S(g_1, g_3), S(g_2, g_3)\}$, называются взвешенными нечеткими отношениями эквивалентности.

Отношению сходства двойственно отношение различия $D: X \times X \rightarrow [0,1]$, удовлетворяющее на $X \times X$ условиям:

- симметричности $D(g_1, g_2) = D(g_2, g_1) \quad \forall g_1, g_2 \in X \times X$;
- антирефлексивности $D(g_1, g_1) = 0 \quad \forall g_1 \in X \times X$.

Отношение различия, удовлетворяющее свойству (\vee, \wedge) -транзитивности $D(g_1, g_2) \leq \max\{D(g_1, g_3), D(g_2, g_3)\}$, в кластерном анализе называется ультраметрикой [12].

С учетом введенных обозначений иерархическая процедура кластеризации на пространственных гранулах может рассматриваться как преобразование отношения сходства S во взвешенное отношение эквивалентности E , которое определяет иерархию отношений эквивалентности и, следовательно, иерархию разбиений множества рассматриваемых объектов на классы эквивалентности [12].

Теперь по аналогии с [12] мы можем ввести схему двухэтапной иерархической кластеризации на пространственных гранулах в виде

$$E = Q(S) = TC(F(S)), \quad (2)$$

где F – оператор коррекции нечеткого отношения сходства S [6], удовлетворяющий условию $F(S) \subseteq S$. Оператор TC определяет процедуру (\vee, \wedge) -транзитивного замыкания взвешенных отношений сходства [6]. Выражение (2) описывает алгоритм иерархической пространственной кластеризации, использованный в данной работе и развивающий идеологию работы [7].

Разработка нейросетевого базиса кластеризации

В ранее предложенных алгоритмических решениях для задачи гранулирования использовался базис дискретной однопроцессорной системы. Однако для задач обработки изображений предпочтительными являются базисы с распараллеливанием вычислений, например, нейросетевые [9]. На основе ранее введенных основных положений аналитического подхода к решению основных задач компьютерной графики [11] нами предложен новый тип активационных функций однослойной гранулированной сети [14], основанный на теории R-функций [10], и разработана параметризованная модель нейрона с нелинейной активационной функцией вида

$$R(w, x) = \left| \frac{w_2 - w_1}{2} \right| - \left| (x) - \left\| \frac{w_2 - w_1}{2} + w_1 \right\| \right|, \quad (3)$$

где w – вектор весов, x – значение координаты. Подобная модель основана на использовании так называемого семейства операций R_0 [10], [11], позволяющего наиболее просто реализовать функцию активации как программно, так и аппаратно.

Структура сети с нейронами (3) соответствует предложенной в [14] типовой решетчатой структуре сети для решения двумерных геометрических задач. Выходная функция сети Y реализуется с помощью выходного слоя, описываемого соотношением

$$Y = \begin{cases} 1 & R(a_1, a_2, x)_1 \& R(a_1, a_2, x)_2 \& R(a_1, a_2, x)_h > 0 \\ 0 & \end{cases}. \quad (4)$$

Для предложенной архитектуры сети введен алгоритм обучения без учителя, основанный на использовании экстремальных свойств функций (3).

Шаг 1

Устанавливаются начальные значения

$$w_1 = \gamma \quad (\gamma - \text{максимальный размер гранулы})$$

$$w_2 = (X_1)_h;$$

$$x = (X_h);$$

$$M(w_1, w_2) = \frac{w_2 - w_1}{2}.$$

Шаг 2

Вычисляется $R(\mathbf{w}, x)$;

Если $R(\mathbf{w}, x) < 0$;

$$w_1 = w_1 + R(\mathbf{w}, x);$$

Вычисляется $R(\mathbf{w}, x)$;

Если $R(\mathbf{w}, x) < 0$

$$w_2 = w_2 - R(\mathbf{w}, x).$$

Шаг 3

Повторяются шаги 1 – 2 для каждой гранулы.

Данный алгоритм выполняется параллельно для каждой гранулы исходного изображения, множество которых первоначально задается разбиением изображения на классы эквивалентности путем введения отношений строгого порядка на координатных осях. Получаемые классы эквивалентности описываются декартовыми гранулами на плоскости [5], что позволяет реализовывать на них операции, сходные с предложенными Л. Задэ операциями *вычислений со словами* (ВС) [4]. По аналогии операции над гранулированными представлениями изображений можно назвать *вычислениями с формами* (или фигурами).

В результате завершения предложенного алгоритма, который является разновидностью алгоритма с выбыванием, для некоторых нейронов размеры гранулы, представленные весовыми коэффициентами, могут свестись к 0 (произошло выбывание нейрона).

Работа предложенного алгоритма иллюстрируется следующими рисунками, показывающими кластеризацию одной гранулы. Рис. 1 показывает исходное состояние нейрона, когда размеры гранулы равны максимальным (γ).

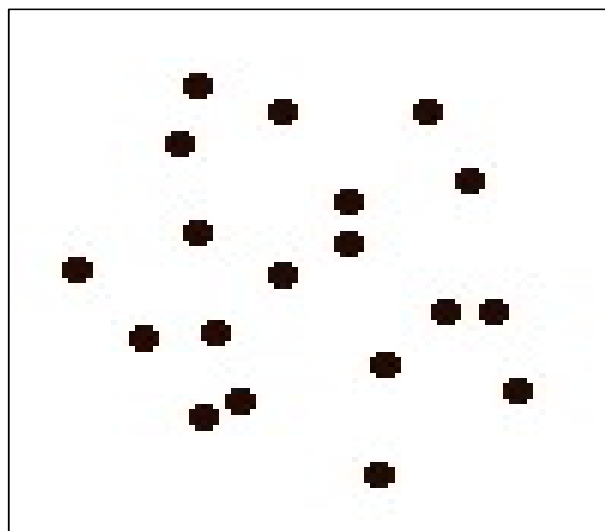


Рисунок 1 – Исходные данные для кластеризации, размещенные внутри квадратной гранулы

На первом шаге обучения происходит пересчет весовых коэффициентов в соответствии с рис. 2.

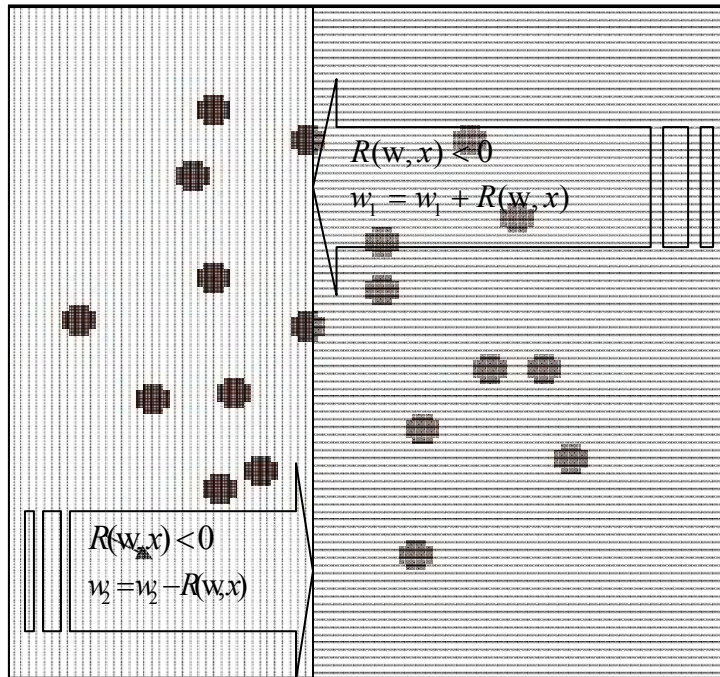


Рисунок 2 – Первый шаг пересчета весовых коэффициентов нейрона

После завершения работы алгоритма по оси ординат веса нейрона будут соответствовать ординатам инкапсулирующей гранулы [4] для представленного на рис. 1 множества пикселей.

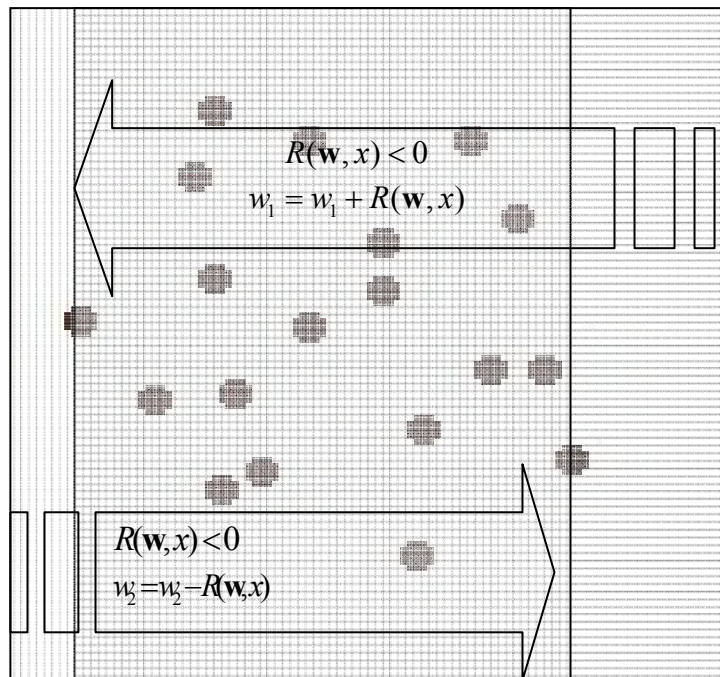


Рисунок 3 – Проекция инкапсулирующей гранулы, полученная после завершения алгоритма

После завершения работы алгоритма по оси ординат веса нейрона будут соответствовать ординатам инкапсулирующей гранулы [4] для представленного на рис. 1 множества исходных пикселей. На следующем этапе происходит обучение нейрона по координатам оси абсцисс.

Предложенная методика кластеризации с помощью нейронной сети позволяет представить исходное множество пикселей изображения, полученное в результате отделения изображения объекта от изображения фона [1] в виде множества декартовых гранул (каноническая форма представления изображения [4]), для которого можно использовать перцептуальные операции, предложенные в [13].

Результаты экспериментального исследования

Для решения задачи моделирования был использован пакет MatLab, с помощью которого был создан программный экспериментальный стенд для статистических исследований настраиваемой модели.

Стенд позволяет изучать действие различных алгоритмов обучения на различных видах экспериментально полученных данных (в основном, изображений).

Следующие рисунки изображают результаты отделения изображения объекта от фона без всякой задаваемой пользователем предварительной информации о свойствах объекта. Рис. 4а представляет исходное малоразмерное (108x153) изображение, а рис. 4б – результат пороговой сегментации [2].

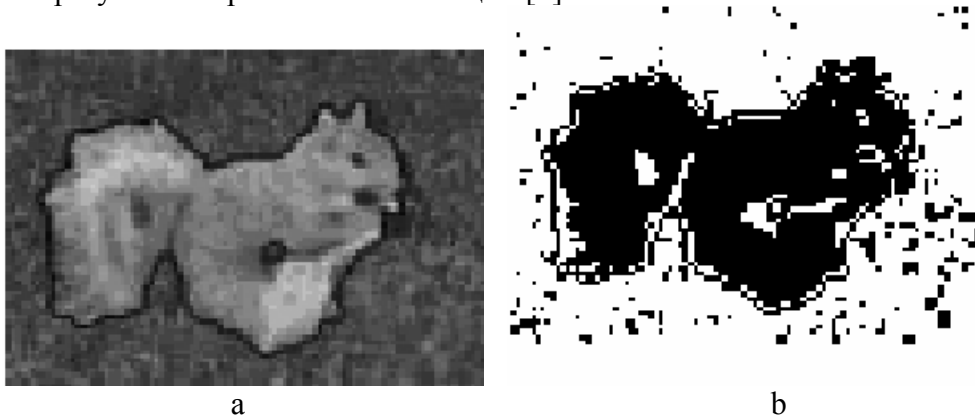


Рисунок 4 – Исходное изображение неоднородно окрашенного объекта на неоднородном фоне и результат пороговой сегментации



Рисунок 5 – Этапы отделения объекта от фона с помощью разработанных методов нечеткой сегментации

Рис. 5 представляет результаты отделения объекта от фона с помощью предложенного авторами двухэтапного метода [1]. На рис. 5а представлен результат нечеткого отделения пикселей, рис. 5б показывает результат «уборки мусора» на бинарном изображении, выполненный по разработанному авторами методу. При этом рост суммарной энтропии изображения составляет 18 %, но после фильтрации она уменьшается до 7 % от исходной, что говорит о высокой корректности разработанных методов.

Выводы

В результате применения нового подхода к построению систем сегментации изображений удалось создать новую методику мягкой сегментации и вычислительный нейросетевой базис для ее реализации.

Экспериментальное исследование подтвердило высокую эффективность полученных методик в условиях работы с изображениями, имеющими высокую степень неопределенности – малые размеры, низкий контраст, высокое отношение сигнал/шум и т.д.

В качестве ведущего направления дальнейших исследований перспективно исследование нечетких метрических функций на канонических представлениях изображений [8], позволяющих улучшить качество кластеризации по (2) на нижнем уровне.

Необходимо также развивать теорию применения новых функций активации для совершенствования структуры и алгоритмов обучения нейронных сетей обработки многомерной информации. Другим направлением развития метода является переход к адаптивной нечеткой сегментации с использованием разработанных авторами методов выделения так называемых гипербоксов для пространственных данных [9].

Литература

1. Бутенков С.А., Зюзерова Н.С., Кривша В.В. Семантический подход в задаче разделения изображений неоднородно окрашенных объектов и неоднородного фона. Настоящий номер.
2. Rosenfeld A. From Image Analysis to Computer Vision // Computer Vision and Image Understanding. – 2001. – Vol. 84. – P. 298-324.
3. Zadeh L. Towards a Theory of Fuzzy Information Granulation and its Centrality in Human Reasoning and Fuzzy Logic // Fuzzy Sets and Systems. – 1997. – Vol. 90. – P. 111-127.
4. Бутенков С.А. Грануляция и инкапсуляция в системах эффективной обработки многомерной информации // Искусственный интеллект. – 2005. – № 4. – С. 106-115.
5. Zadeh L. From Computing with Numbers to Computing with Words – From Manipulation of Measurements to Manipulation of Perceptions // IEEE Trans. Circuits and Systems – Fundamental Theory and Applications. – 1999. – Vol. 45, № 1. – P. 105-119.
6. Zadeh L.A. Similarity relations and fuzzy orderings // Information Sciences. – 1973. – Vol. 3. – P. 177-200.
7. Бутенков С.А., Аль-Доуяни С.Х.С. Синергетический подход к задаче генерализации при обучении нечетких нейронных сетей для обработки изображений // Искусственный интеллект. – 2005. – № 4. – С. 319-324.
8. Butenkov S., Krivsha V., Al Dhouyani S. Granular Computing in Computer Image Perception: basic issues and Glass Box models // Proc. IASTED Conf. In Artificial Intelligence and applications “AIA 2006”. – Innsbruck (Austria). – February 16-18 2006. – P. 811-816.
9. Бутенков С.А. Модели «стеклянного ящика» в задачах интеллектуальной обработки информации // Сб. трудов Научной сессии МИФИ-2006. – 23-27 января Москва. – 2006. – Т. 3. – С. 71-73.

10. Рвачев В.Л. Теория R-функций и некоторые её приложения. – Киев: Наукова думка, 1982.
11. Бутенков С.А. Обобщенные геометрические модели и новая парадигма обработки геометрической информации // Известия ТРТУ. – 2002. – № 3. – С. 150-157.
12. Baturshin I., Rudas T., Klimova A. On general scheme of invariant clustering procedures based on fuzzy similarity relation // Proc. International Conf. on Fuzzy Sets and Soft Computing in Economics and Finance (FSSCEF 2004). – St. Petersburg (Russia). – June 17-20, 2004. – Vol. I. – P. 122-129.
13. Бутенков С.А., Каркищенко А.Н., Кривша В.В. Перцептуальный подход к анализу гранулированных изображений в интеллектуальных системах // Искусственный интеллект. – 2003. – № 3. – С. 209-218.
14. Бутенков С.А., Аль-Доуяни С.Х.С. О принципах построения оптимальных нечетких нейронных сетей // Сб. трудов Научной сессии МИФИ-2006. – Москва. – 23-27 zydfhz2006. – Т. 3. – С. 78-79.

С.А. Бутенков, Al Dhoyani Saud

Застосування нейронних мереж для семантичної інтелектуальної сегментації

Робота присвячена дослідженню нових підходів до розширення сфери застосування нейронних мереж. Дається аналіз існуючих обмежень, спричинених використанням моделі синергетичної мережі. Запропоновано синергетичний алгоритм, який дозволяє використовувати методи для побудови промислових систем.

S.A. Butenkov, Al Dhoyani Saud

This paper deals with the especial kind of neuron nets for the whole processing of granulated images. The main problem is related with the activation functions and the learning process stability under the noised initial images. We propose the very common approach to neuron nets architecture and learning algorithms, based on R-functions concept. As a result, the provided algorithms are very robust and efficient for the corrupted and small-size images.

Статья поступила в редакцию 23.06.2006.