

# Нейросетевой алгоритм обнаружения малоразмерных объектов на облачных фонах

Н.Ю. Шубин, В.С. Муравьев, С.И. Муравьев

Факультет автоматики и информационных технологий в управлении

Рязанский государственный радиотехнический университет, Рязань, Россия, [aitu@rsreu.ru](mailto:aitu@rsreu.ru)

## Аннотация

Постоянное совершенствование датчиков изображений и средств вычислительной техники приводит к расширению сферы практического применения методов и алгоритмов анализа изображений. Так, комплексы обработки изображений реального времени начали активно применяться в системах автоматического обнаружения и сопровождения объектов. Области их использования являются: управление дорожным движением, контроль производственных процессов, исследование Земли из космоса, робототехника, медицина и целый ряд других.

Особое внимание в последние два десятилетия уделяется проблеме обнаружения малоразмерных воздушных объектов. В работе предложен нейросетевой алгоритм решения данной задачи, ориентированный на анализ изображений в темпе их поступления от видеодатчика. Описана процедура обучения нейронной сети, приведены результаты сравнительных экспериментальных исследований.

*Ключевые слова:* нейронная сеть, обнаружение, малоразмерный объект, бинарное изображение.

## 1. ВВЕДЕНИЕ

Одной из важных задач, решаемых в оптико-электронных системах наблюдения за воздушным пространством, является обнаружение, оценка параметров и сопровождение объектов. Необходимо отметить, что ранее для этих целей использовались преимущественно радиолокационные станции (РЛС). Однако РЛС присущ ряд существенных недостатков, к которым можно отнести присутствие “слепой зоны”, чувствительность к постановкам помех, высокую стоимость и значительные габариты. Кроме этого, во многих случаях применение активных способов наблюдения нежелательно, ввиду ограничений по энергопотреблению и необходимости обеспечения скрытности средств обзора воздушного пространства. Одним из способов решения указанных проблем является использование современных телевизионных и тепловизионных датчиков для получения изображения объекта с последующим анализом поступающей видеoinформации. Зачастую интересные объекты на изображениях могут быть малоразмерными или точечными, так как они имеют небольшие габариты и находятся на значительном удалении от датчика изображения.

Отметим ряд особенностей, усложняющих решение задачи обнаружения и оценки параметров воздушных объектов по сравнению с классическими задачами технического зрения:

- присутствие геометрических преобразований изображений, возникающих вследствие изменения ориентации датчика в пространстве;
- высокая динамика фона;
- требование обеспечения высокой вычислительной

эффективности алгоритма, связанное с необходимостью обработки в реальном масштабе времени последовательности изображений при ограниченных аппаратных ресурсах.

– возможность работы алгоритма в замкнутом контуре сопровождения.

Вместе с тем при разработке алгоритма обнаружения малоразмерных объектов необходимо принимать во внимание свойства облачного фона, которому свойственно наличие корреляционных связей между соседними элементами изображения. В данной работе предлагается нейросетевой алгоритм обнаружения объектов, основанный на пространственной обработке изображений. Рассматриваемый подход может использоваться в таких областях как контроль и мониторинг воздушного пространства, аэронавигация, метеорологические исследования, в военных и специальных приложениях.

## 2. ОПИСАНИЕ АЛГОРИТМА

Для успешного решения задачи обнаружения объектов, необходимо располагать как можно более полной информацией о функциях яркости объектов и фона, а также знать параметры распределения шума. Оптимальный в смысле выбранного критерия качества алгоритм можно построить только при наличии полной априорной информации о свойствах объекта и фона. В большинстве случаев требуемая информация бывает недоступна. Реальные условия всегда характеризуются той или иной степенью неопределенности в отношении информации о фоновой обстановке, которая обусловлена как невозможностью точного предсказания свойств объекта и фона, так и непредвиденными изменениями этих свойств во времени. Проблемы, связанные с математическим синтезом оптимальных алгоритмов приводят к необходимости поиска эвристических подходов.

Для решения задач классификации широкое распространение получил метод опорных векторов [1], однако высокое качество решения задачи достигается за счет увеличения вычислительной сложности, что затрудняет его использование в системах обработки изображений реального времени. На практике в случае отсутствия формализованной информации о фоновой обстановке могут применяться нейронные сети (НС), которые обобщают в процессе обучения основные отличия необходимые для выработки классификационных решений о принадлежности точек изображения объекту либо фону.

Первоначальный этап обработки в большинстве случаев заключается в формировании многомерного признакового пространства. Общеизвестно, что качество решения задачи классификации в большей степени зависит от выбранной для описания системы признаков. Не вызывает сомнения целесообразность использования телевизионных, тепловизионных и радиолокационных яркостных

изображений. Располагая этими естественными признаками, путем их линейной либо нелинейной обработки получают пространственные, градиентные, спектральные, текстурные и ряд других характеристик [2]. Однако вопрос о требуемой размерности признакового пространства остается открытым. Так, с целью выделения наиболее информативных свойств возможно выполнять редукцию многомерного пространства с помощью преобразования главных компонент [3].

Представим совокупность полученных  $M$  признаков в виде трехмерного массива чисел размерностью  $N_x \times N_y \times M$ , где  $N_x, N_y$  – ширина и высота наблюдаемого изображения. Для определенности положим  $M$  нечетным и введем в рассмотрение многомерное окно размерностью  $L \times L \times M$  с центром в точке  $(i, j, (M-1)/2)$ , перемещающееся вдоль координат  $i$  и  $j$ . На каждом шаге из исходного массива в пределах окна выбираются данные, из которых формируется входной вектор из  $N_{вх} = L^2 M$  элементов, поступающий на вход нейронной сети. Нейронная сеть может осуществлять отнесение по нечеткому правилу центрального элемента к классу “объект” или “фон”, а в соответствующий элемент выходного двумерного изображения записывать значение, соответствующее степени истинности. Таким образом, после завершения обхода получается полутоновое изображение. Можно утверждать, что предложенный подход к классификации имеет много общего с процедурой многомерной фильтрации [4].

Во многих случаях на практике потребителя интересует бинарное изображение, которое может быть получено путем глобальной пороговой обработки результата нейронечеткой классификации. Пороговое значение может быть задано в зависимости от заданной степени истинности.

После получения бинарного изображения можно уменьшить степень фрагментации сегментов (связных областей) путем морфологической обработки, заключающейся в последовательном применении операций морфологического закрытия и открытия с квадратными структурирующими элементами. Переход от итогового бинарного изображения к представлению результатов обработки в виде списка параметров связных областей изображения производится с помощью процедуры разметки и параметризации. Исходя из параметров найденных сегментов и имеющейся априорной информации, можно принять решение о наличии интересующих объектов в обрабатываемой области кадра.

Важным вопросом является выбор структуры НС, который осуществляется в соответствии с особенностями и сложностью задачи. Для решения некоторых отдельных типов задач, на сегодняшний день, уже существуют оптимальные архитектуры [5]. Если же задача не может быть сведена ни к одному из известных типов, разработчику приходится решать проблему синтеза новой архитектуры сети. При этом обычно руководствуются несколькими ключевыми принципами:

- вычислительные возможности сети возрастают с увеличением числа ячеек сети, плотности связей между ними и числа выделенных слоев;
- введение обратных связей наряду с увеличением возможностей сети к решению многих практических задач ставит вопрос о ее динамической устойчивости;
- сложность алгоритмов функционирования сети способствует улучшению способности к обобщению данных.

Однако в большинстве случаев приемлемый вариант получается на основе интуитивного подбора, и дать подробные рекомендации не представляется возможным. В данной работе в качестве базовой структуры сети был выбран перцептрон с одним скрытым слоем и сигмоидальными активационными функциями (рисунок 1), что было обусловлено изученностью его свойств, простотой реализации и скоростью работы. Количество нейронов в скрытом слое  $N_{скр}$  удовлетворяет неравенству  $N_{скр} < N_{вх} - 1$ .

### 3. КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ

#### 3.1 Обучение сети

Планируется, что разрабатываемый алгоритм будет использован в бортовых системах обработки информации, что накладывает существенные ограничения на вычислительную трудоемкость.

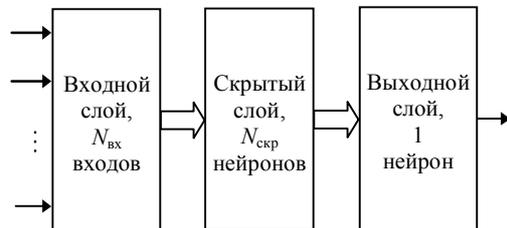


Рисунок 1. Структура трехслойного перцептрона.

В текущем варианте реализации в качестве базового признака было выбрано яркостное изображение. Полученные результаты, характеризующие работоспособность разрабатываемого алгоритма, носят промежуточный характер, и в ближайшей перспективе предполагается выполнять комплексирование мультиспектральных изображений. Для обучения сети использовался метод сопряженных градиентов, который обеспечивает компромисс между качеством обучения и вычислительной эффективностью.

С целью проведения обучения НС был создан обширный банк видеопоследовательностей ИК диапазона. В качестве датчиков применялись охлаждаемые тепловизоры и болометрические приборы, на основе сканирующих линеек и матричного типа, чувствительные к излучению с длинами волн 3-5 мкм и 8-14 мкм. Выбранные сюжеты характеризовались наличием как однородного, так и облачного фонов, отношение сигнал/шум (отношение локального контраста объекта к СКО шума) достигало значения 2,5 и менее. Каждый сюжет содержал один объект, что объясняется необходимостью проектирования алгоритма для работы в замкнутой системе автоматического сопровождения объектов. Учитывая размеры объектов и статистические характеристики фона, значение  $L$  было принято равным 7. В качестве реальных объектов выступали мини-БПЛА, Ми-8, МиГ-31 и другие. Из тестовой базы было отобрано 50 характерных кадров, содержащих различные типы сцен. Примеры исходного и инвертированного бинарного изображений приведены на рисунке 2.

На каждом кадре формировались обучающие примеры, т.е. участки изображений размерами  $7 \times 7$ , представляемые в виде вектора, содержащего 49 элементов.

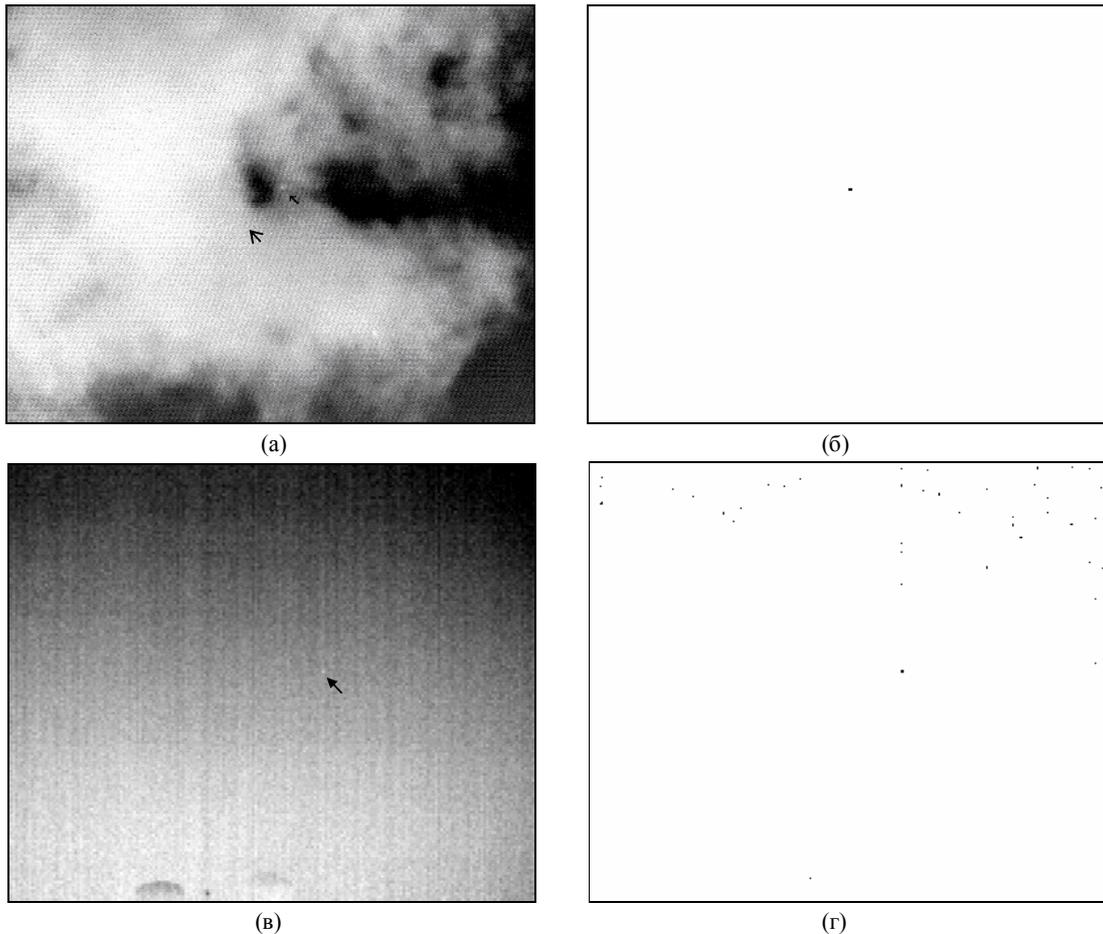


Рисунок 2. Примеры: исходных изображений (а, в) и инвертированных бинарных изображений (б, г). Стрелками обозначено положение объекта.

Выбор фоновых участков производился произвольным образом из разных областей исходного кадра. Участки изображений, содержащих объект, выбирались так, чтобы яркость центра объекта отличалась от точек окрестного фона. Для фона и для объекта формировалось одинаковое число обучающих примеров, по 250 для каждого. Обучающие примеры для разных классов подавались на вход сети поочередно. Максимальное количество циклов обучения (эпох) составило несколько тысяч, минимальная ошибка обучения задавалась на уровне  $10^{-6}$ . Для улучшения процесса обучения входные данные нормализовывались на основе найденных оценок математического ожидания и дисперсии. По результатам проведенного компьютерного моделирования  $N_{скр}$  было взято равным 16.

### 3.2 Экспериментальные исследования

Экспериментальная проверка алгоритма проводилась на натуральных тестовых видеопоследовательностях ИК диапазона продолжительностью от 500 до 1000 кадров. Для оценки работоспособности предложенного алгоритма использовалась следующая методика. На  $n$ -м кадре рассчитывался бинарный параметр правильного обнаружения  $N_{no}^{(n)}$  по следующему правилу

$$N_{no}^{(n)} = \begin{cases} 1, & \text{если } \left| \frac{\mathcal{E}_y^n}{i_{yэ}^n} - i_{yэ}^n \right| \leq 1, \left| \frac{\mathcal{E}_x^n}{j_{xэ}^n} - j_{xэ}^n \right| \leq 1, n = \overline{1, N}; \\ 0, & \text{иначе,} \end{cases}$$

где  $(i_{yэ}^n, j_{xэ}^n)$  – эталонные значения координат центров объектов на  $n$ -м кадре, в качестве которых использовались данные, полученные человеком-экспертом,  $(\mathcal{E}_y^n, \mathcal{E}_x^n)$  – измеренные значения координат центров объектов,  $N$  – число кадров в видеосюжете.

Зная  $N_{no}^{(n)}$  на каждом кадре, можно вычислить частоту правильного обнаружения  $P_{no}$  как

$$P_{no} = \frac{1}{N} \sum_{n=1}^N N_{no}^{(n)}.$$

Дополнительно оценивалась частота ложных тревог  $P_{лм}$  как усредненное отношение количества точек, неверно отнесённых к объекту, к общему количеству точек изображения, не принадлежащих объекту. При варьировании порогового коэффициента можно построить кривую зависимости  $P_{no} = f(P_{лм})$ , называемую рабочей характеристикой обнаружения. Пример зависимостей  $P_{no} = f(P_{лм})$  полученных для двух видеосюжетов

представлен на рисунке 3 (характерные кадры из тестовых видеопоследовательностей представлены на рисунке 2). При проведении сравнительных исследований использовался алгоритм обнаружения воздушных объектов, описанный в [6]. Суть подхода состоит в применении параллельной двумерной фильтрации наблюдаемого изображения двумя фильтрами с масками разного вида и размерности, подстройки размерности маски фильтра для лучшего обнаружения интересующего объекта, сравнения обработанного изображения с глобальным порогом, морфологической фильтрации бинарного изображения. Данный алгоритм выбран для сравнения, так как он успешно используется в системах обработки изображений реального времени семейства “Охотник”, выпускаемых ФГУП ГРПЗ [7], а его эффективность подтверждена опытом практического применения.

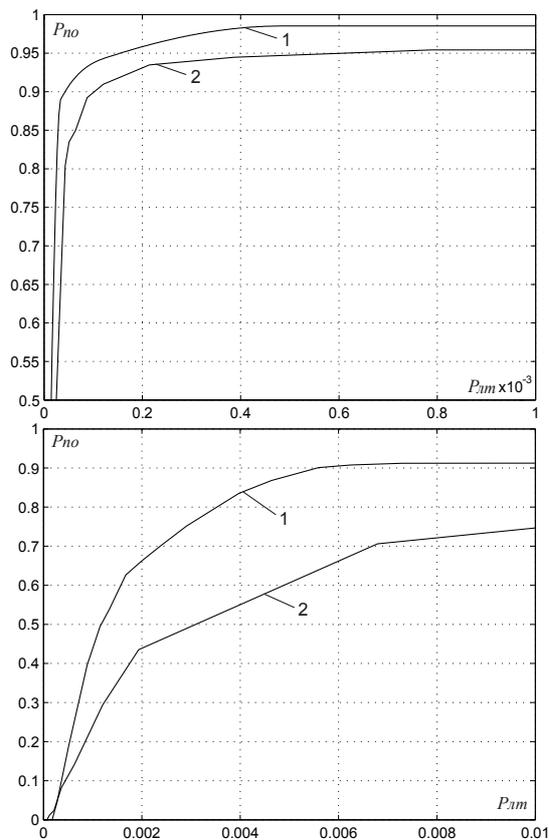


Рисунок 3. Графики зависимости  $P_{no} = f(P_{dm})$ . Кривая 1 получена для нейросетевого алгоритма, кривая 2 – для предложенного ранее подхода.

Анализ полученных результатов показывает, что прирост частоты правильного обнаружения при фиксированной частоте ложных тревог в среднем составляет около 10%. Исследования показывают, что значение СКО ошибки измерения координат центров объектов на изображении не превышает 1,5 пикселей.

#### 4. ЗАКЛЮЧЕНИЕ

В работе предложен нейросетевой алгоритм обнаружения малоразмерных воздушных объектов, наблюдаемых на фоне

облачного неба. Алгоритм ориентирован на реализацию в системах обработки изображений реального времени. Приведены предварительные результаты сравнительных экспериментальных исследований, свидетельствующие о возможности применения рассматриваемого подхода. Улучшения работы алгоритма планируется достичь путем увеличения размерности признакового пространства и использования других архитектур нейронной сети.

#### 5. СПИСОК ЛИТЕРАТУРЫ

- [1] Хайкин С. Нейронные сети: полный курс. Пер. с англ., 2006.– 1104с.
- [2] Heesung Kwon, Sandor Z. Der, Nasser M. Nasrabadi. Adaptive multisensor target detection using feature-based fusion // *Optical Engineering*, vol. 41, no. 1, 2002, pp. 69-80.
- [3] Image Fusion: algorithms and applications, edited by Tania Sathaki, London, Academic Press, 2008, 500 p.
- [4] Mukul V. Shirvaikar, Mohan M. Trivedi. A neural network filter to detect small targets in high clutter background // *IEEE Transactions on Neural Networks*, vol.6, no.1, 1995, pp. 252-257.
- [5] Markos Markou, Sameer Singh. Novelty detection: a review – part 2: neural network based approaches // *Signal Processing*, vol. 83, 2003, pp. 2499 – 2521.
- [6] Алпатов Б.А., Блохин А.Н., Муравьев В.С. Алгоритм обработки изображений для систем автоматического сопровождения воздушных объектов // *Цифровая обработка сигналов*. – 2010. – №4. – С.11-17.
- [7] Алпатов Б.А., Блохин А.Н., Костяшкин Л.Н., Романов Ю.Н., Шапка С.В. Семейство многофункциональных систем обработки видеозаписей “Охотник” // *Цифровая обработка сигналов*. – 2010. – №4. – С.44-51.

#### NEURAL NETWORK ALGORITHM FOR SMALL OBJECT DETECTION ON CLOUDY BACKGROUND

##### Abstract

Continuous improvement of video sensors and computer equipment leads to an expansion of practical applications of the image analysis. The real-time image processing systems have been actively used for automatic object detection and tracking. They are in high demand in traffic management, control of product manufacturing, remote sensing, robotics, medicine and other fields.

Much attention in last two decades is paid to the problem of small target detection on cloudy background. In this work the neural network approach with the orientation on image analysis of the real scale video sequences has been suggested. The network learning procedure has been described and the preliminary results of comparative experimental research have been also presented.

##### Авторы

Шубин Н.Ю., аспирант кафедры Автоматики и информационных технологий в управлении (АИТУ) Рязанского государственного радиотехнического университета (РГРТУ)

Муравьев В.С., к.т.н., научный сотрудник кафедры АИТУ РГРТУ

Муравьев С.И., к.т.н., доцент кафедры АИТУ РГРТУ.