

ПРИМЕНЕНИЕ ИМПУЛЬСНЫХ И РЕКУРРЕНТНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ВЫДЕЛЕНИЯ ДВИЖУЩИХСЯ ОБЪЕКТОВ НА ВИДЕОИЗОБРАЖЕНИИ

В.В. Чередов, И.А. Емельянов, П.С. Зайцев, И.А. Евсеенко

Предложены модели импульсной и рекуррентной нейронных сетей для выделения движущихся объектов на видеоизображении, позволяющие создать детектор движения.

Ключевые слова: пиксель, видеоизображение, нейрон, нейронная сеть.

1. Введение

Исследования по искусственным нейронным сетям связаны с тем, что способ обработки информации человеческим мозгом в корне отличается от методов, применяемых обычными цифровыми компьютерами. Мозг представляет собой чрезвычайно сложный, нелинейный, параллельный компьютер (систему обработки информации). Он обладает способностью организовывать свои структурные компоненты, называемые нейронами, так, чтобы они могли выполнять конкретные задачи (такие как распознавание образов, обработку сигналов органов чувств, моторные функции) во много раз быстрее, чем могут позволить самые быстродействующие современные компьютеры. Примером такой задачи обработки информации может служить обычное зрение. В функции зрительной системы входит создание представления окружающего мира в таком виде, который обеспечивает возможность взаимодействия с этим миром. Более точно, мозг последовательно выполняет ряд задач распознавания (например, распознавание знакомого лица в незнакомом окружении). На это у него уходит около 100-200 миллисекунд, в то время как выполнение аналогичных задач даже меньшей сложности на компьютере может занять несколько дней.

2. Модель импульсной нейронной сети

Структура импульсной нейронной сети, применяемая для выделения движущихся объектов, представлена на рис. 1 [3].

Входной слой нейронов является аналогом слоя фоторецепторов сетчатки глаза, поэтому далее будем называть нейроны первого слоя рецепторами. Каждому пикселю (x, y) входного кадра видеоизображения соответствует свой рецептор $N_r(x, y)$. Скрытый слой представляет собой аналог слоя интернейронов сетчатки глаза. Он состоит из двух независимых друг от друга массивов нейронов N_1 и N_2 . Они имеют такие же размеры, как и первый слой, и соединены синаптическими связями как с входным слоем нейронов (рецепторов) N_r , так и с выходным слоем нейронов N_{out} .

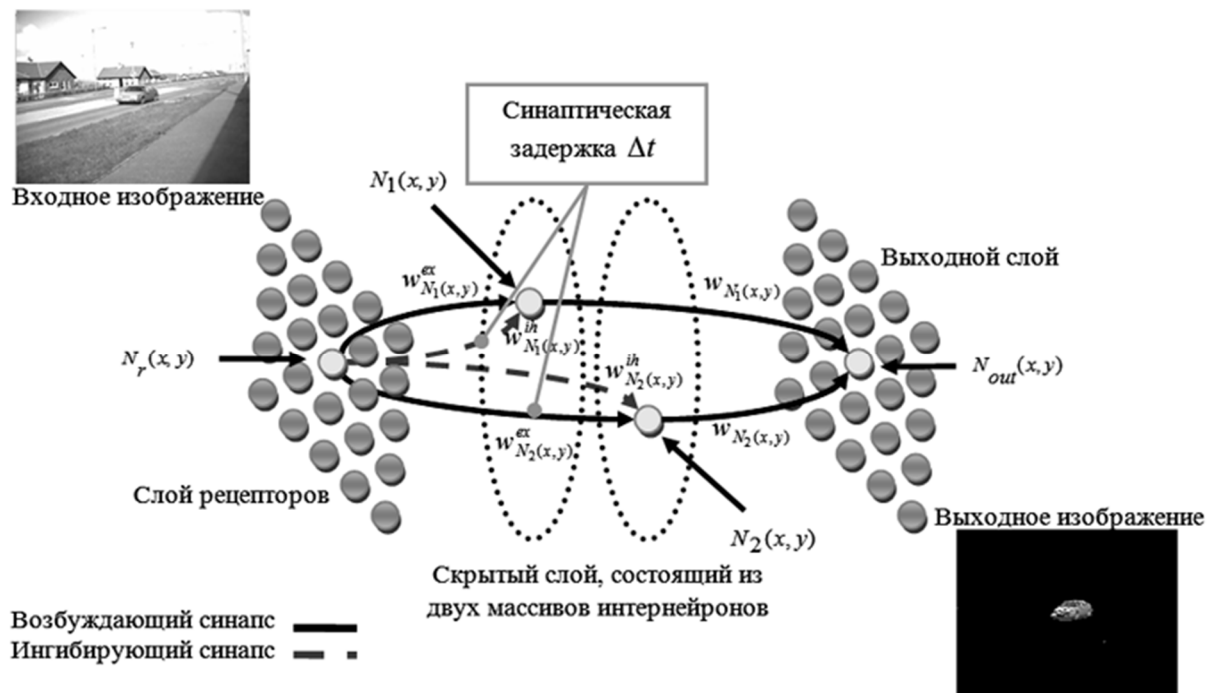


Рис. 1. Модель импульсной нейронной сети для детектора движения

Каждый рецептор $N_r(x, y)$ соединен с интернейроном $N_1(x, y)$ посредством возбуждающего синапса, передающего сигналы без задержек, и ингибирующего синапса, передающего сигналы с синаптической задержкой Δt . Также каждый рецептор $N_r(x, y)$ соединен с интернейроном $N_2(x, y)$ посредством возбуждающего синапса, передающего сигналы с синаптической задержкой Δt , и ингибирующего синапса, передающего сигналы без задержек.

Допустим, что $I_{nr}(x, y, t)$ – сила тока, текущего от рецептора $N_r(x, y)$ в момент времени t . Если ток, идущий от рецептора $N_r(x, y)$ стабилен, т.е. $I_{nr}(x, y, t) = I_{nr}(x, y, t - \Delta t)$, а возбуждающий и ингибирующий входы интернейрона $N_1(x, y)$ сбалансированы путем настройки весов $w_{N_1(x,y)}^{ex}$ и $w_{N_1(x,y)}^{ih}$ соответственно, то интернейрон $N_1(x, y)$ будет находиться в состоянии покоя. Аналогичная ситуация складывается и с интернейроном $N_2(x, y)$. Если же сила тока рецептора $N_r(x, y)$ растет, т.е. $I_{nr}(x, y, t) > I_{nr}(x, y, t - \Delta t)$, то баланс нарушается, поскольку сигнал приходящий с возбуждающего синапса сильнее, чем задержанный на период Δt сигнал, приходящий с ингибирующего синапса. Интернейрон $N_1(x, y)$ начинает генерировать импульсы (спайки). Если же сила тока рецептора снижается, т.е. $I_{nr}(x, y, t) < I_{nr}(x, y, t - \Delta t)$, то интернейрон $N_1(x, y)$ не реагирует, но зато начинает генерировать импульсы интернейрон $N_2(x, y)$, поскольку задержанный на период Δt сигнал, приходящий с возбуждающего синапса, сильнее, чем сигнал, приходящий с ингибирующего синапса. Другими словами, нейронная сеть, начинает реагировать на изменение яркости пикселей, которая может быть вызвана прохождением движущегося объекта по статическому фону.

Выходной слой нейронной сети N_{out} имеет такие же размеры, как входной и скрытый слои. Каждый нейрон данного слоя $N_{out}(x, y)$ соответствует каждому пикселю (x, y) выходного кадра видеоизображения. Интернейроны $N_1(x, y)$ и $N_2(x, y)$ соединены возбуждающими синапсами без задержек с выходным нейроном $N_{out}(x, y)$. Он продуцирует сигналы только в том случае, когда получает импульсы от интернейрона $N_1(x, y)$ или $N_2(x, y)$, иначе он находится в состоянии покоя. Величина яркости серой шкалы каждого пикселя (x, y) выходного кадра видеоизображения пропорциональна частоте генерации импульсов выходным нейроном $N_{out}(x, y)$ и имеет значение 0 (черный), если выходной нейрон $N_{out}(x, y)$ не генерирует никаких сигналов с течением определенного периода времени T . В противном случае яркость пикселя (x, y) будет выше 0 (рис.2).



Рис. 2. Выделение движущегося объекта с помощью импульсной нейронной сети

3 Модель рекуррентной нейронной сети

Рекуррентные нейронные сети – это класс нейронных сетей с обратной связью между различными слоями нейронов [1, 2]. Особенностью таких сетей является передача сигналов с выходного (скрытого) слоя на входной слой. Особый интерес представляют многослойные рекуррентные сети, которые являются развитием однонаправленных сетей персептронного типа за счет добавления обратных связей. На каждом слое такой сети присутствует элемент единичной задержки, который позволяет считать входной поток сигналов однонаправленным. Такая рекуррентная нейронная сеть функционирует как однонаправленная персептронная сеть. Алгоритм обучения такой сети является более сложным из-за зависимости сигналов в момент времени t от их значения в предыдущие моменты.

Среди рекуррентных сетей (*Recurrent MultiLayer Perceptron*, рекуррентная сеть Эльмана, *Real Time Recurrent Network* и т.д.) наиболее интересной является сеть типа *RTRN (Real Time Recurrent Network)*, предложенная Р. Вильямсом и Д. Зипсером и предназначенная для обработки сигналов в реальном времени [3]. Сеть *RTRN* является частным случаем сети Эльмана. Структура сети *RTRN* представлена на рис. 3.

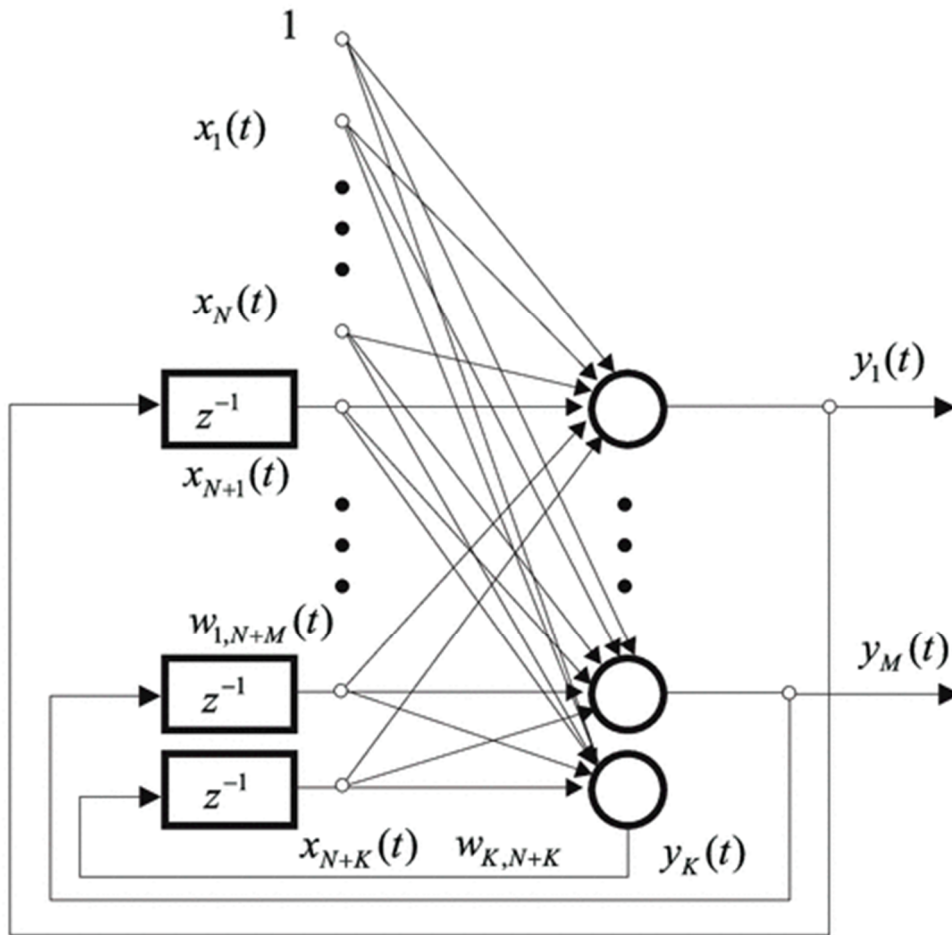


Рис. 3. Структура сети RTRN

Сеть содержит N входных узлов, K скрытых нейронов и K соответствующих им узлов контекстного слоя. Из K скрытых нейронов только M составляют выход сети. Все сигналы выходных нейронов подсоединены как входы нейронной сети через элементы задержки z^{-1} . Обозначим взвешенную сумму i -го нейрона скрытого слоя u_i , а выход этого нейрона – y_i . Вектор $x(t)$ и смещенный (задержанный) на один цикл вектор $y(t-1)$ образуют расширенный вектор активации $x(t)$, возбуждающий нейроны сети:

$$x(t) = [1, x_1(t), x_2(t), \dots, x_N(t), y_1(t-1), \dots, y_K(t-1)]^T \quad (1)$$

После описания входного вектора сети в момент t можно определить состояние всех нейронов согласно зависимостям:

$$u_i = \sum_{k=0}^{N+K} w_{ik} x_k \quad (2)$$

$$y_i(t) = f(u_i(t)) \quad (3)$$

Причем $f()$ обозначает непрерывную функцию активации нейрона (сигмоидальную). Также мы определили $w_i = [w_{i0}, \dots, w_{i, N+K}]$ как вектор весов i -того нейрона и $W = [w_1, \dots, w_K]$ - матрица весов нейронов. Для обучения *RTRN* сети используется алгоритм обучения Вильяма-Зипсера, а именно необходимо минимизировать следующий критерий

$$J(t) = \frac{1}{2} \sum_{i=1}^K = \frac{1}{2} \sum_{i=1}^K [d_i(t) - f(x(t)w_i(t))]^2, \quad (4)$$

где $\varepsilon_i(t)$ определяется по выражению (5)

$$\varepsilon_i(t) = \begin{cases} d_i(t) - y_i(t) & | i = 1 \dots M; \\ 0 & | i > M. \end{cases} \quad (5)$$

После минимизации критерия (4) веса $w_{\alpha\beta}$ (вес из матрицы W) могут быть определены следующим образом

$$w_{\alpha\beta}(t+1) = w_{\alpha\beta}(t) - \eta \sum_{i=1}^K \varepsilon_i(t) \frac{dy_i(t)}{dw_{\alpha\beta}}. \quad (6)$$

Предложенную модель рекуррентной нейронной сети можно использовать для обнаружения движущихся объектов на видеоизображении. Видеоизображение представляет собой последовательность кадров, каждый кадр представляет собой набор пикселей, заданных соответствующими параметрами. Будем считать, что на вход нейронной сети поступает последовательность кадров, а на выходе ожидается получить обработанную последовательность кадров, на котором только движущийся объект выделен из общего изображения.

Набор пикселей входного кадра видеоизображений будет соответствовать входному потоку сигналов для входных узлов нейронной сети, таким образом, примем количество входных узлов сети N равным количеству пикселей одного кадра. Так как на выходе нейронной сети ожидаем получить обработанный кадр, то устанавливаем количество нейронов скрытого слоя $K = N$. Выходы скрытого слоя нейронов являются выходными сигналами нейронной сети, а также являются входными сигналами для элементов задержки z^{-1} (контекстный слой).

На вход каждого нейрона скрытого слоя поступает сигнал от каждого входного узла и от каждого элемента задержки контекстного слоя. Таким образом, каждый нейрон скрытого слоя в момент времени t обладает информацией о пикселях текущего кадра и обработанного (предыдущего) кадра в момент времени $t - 1$. Обработывая информацию о пикселях текущего и предыдущего (обработанного) кадра скрытый нейрон выдает соответствующий сигнал.

В начальный момент времени t_0 , когда элементы задержки не имеют информацию об обработанном кадре, на вход сети поступает первый кадр видеоизображения. В этот

же момент происходит процесс обучения (установки весов нейронной сети), который направлен на установку сети в равновесное положение. Равновесное положение достигается тогда, когда входной и обработанный кадры идентичны, т.е. на видеоизображении нет движения.

При появлении движения на видеоизображении (имеется отличие на двух последовательно идущих кадрах) настроенная нейронная сеть выйдет из состояния равновесия, что отобразится на обработанном кадре. На следующем шаге нейронная сеть будет пытаться войти в состояние равновесия на основе обработанного кадра. Такой процесс будет продолжаться до тех пор, пока на видеоизображении присутствует движение.

Одним из вариантов развития (модернизации) предложенной модели рекуррентной нейронной сети для детектора движения является уменьшение избыточности в связях между скрытым слоем нейронов и входным слоем сигналов, а также контекстным слоем. Такая избыточность в связях затрудняет обучение сети и обработку видеоизображений высокого разрешения. Предлагается уменьшить количество связей и связать каждый нейрон скрытого слоя с соответствующим входным элементом входного слоя (пикселем) и соседними с этим пикселем элементами. В качестве соседних пикселей возможно использование 4 или 8 соседних пикселей, соответствующих 4- или 8-связности (рис. 4).

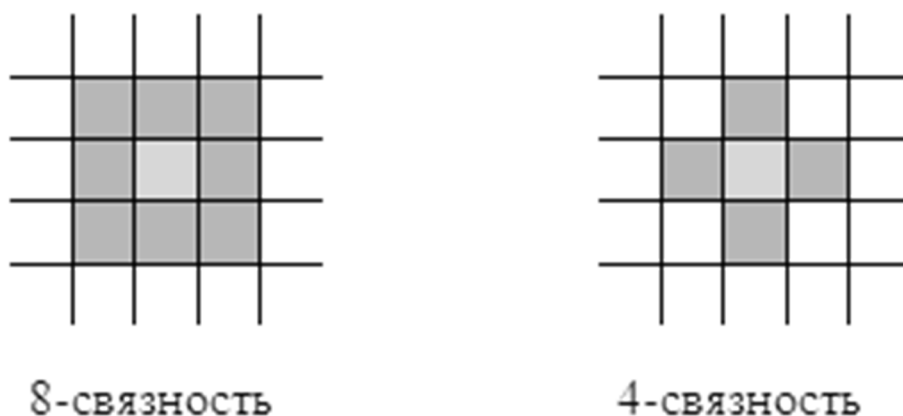


Рис. 4. Соседние пиксели при 8- и 4-связности

Предложенный вариант уменьшения избыточности позволит сократить время обучения сети, при уменьшении количества связей необходимо выполнять меньшее количество расчетных операций, что снизит требования к вычислительным ресурсам детектора движения. Так же сократится время обработки каждого кадра и время реакции датчика на движение, что позволит использовать предложенную методику для видеоизображений с высокой частотой кадров.

4. Заключение

В работе описаны основные перспективные направления исследований в области детектирования движения на основе искусственного интеллекта с применением нейронных сетей.

Предложены модели импульсной и рекуррентной нейронных сетей для выделения движущихся объектов на видеоизображении.

Недостатком предлагаемого подхода для детектирования движения на видеоизображении является то, что при увеличении разрешения изображения число нейронов в

сети резко возрастает, что приводит к заметным задержкам при обработке потока сигналов нейронов.

Предложенный подход для обнаружения и выделения движущихся объектов является попыткой симитировать способности человеческого глаза достаточно быстро выделять движущиеся объекты и превзойти существующие детерминистские методы по скорости выделения движущихся объектов и экономии вычислительных ресурсов. А разрабатываемый на базе данного подхода детектор движения, как программный модуль, сможет найти достойное применение в области цифровой обработки видеоизображения. Предполагается использование данного детектора в автоматизированных системах управления дорожным движением, как альтернатива существующим детекторам, даже с учетом возможного улучшения последних путем использования параллельных вычислений для одновременной обработки сегментов видеоизображения и выделения движущихся объектов в пределах каждого из них.

Элементы импульсной нейронной сети могут быть реализованы аппаратно, или программно с применением современных технологий параллельных вычислений на базе графических процессоров. Это может значительно ускорить процесс выделения движущихся объектов на видеоизображении.

Литература

1. Бендерская, Е.Н. Рекуррентные нейронные сети в задачах распознавания образов / Е.Н. Бендерская, К.В. Никитин // Матриалы IX Международной конференции «Интеллектуальные системы и компьютерные науки» (23-27 октября 2006 г.). – М.: МГУ им. М.В. Ломоносова, 2006. – Т. 1. – Ч. 1. – С. 60-69.
2. Колесницкий, О.К. Аппаратная реализация элементов импульсных нейронных сетей с использованием биспин-приборов / О.К. Колесницкий, И.В. Бокоцей, С.С. Яремчук // Научная сессия МИФИ – 2010. Сборник научных трудов. – М.: МИФИ, 2010. – Ч.1: XII Всероссийская научно-техническая конференция. Нейроинформатика – 2010. – С. 121-132.
3. Мельников, И.И. Детектор движения на основе импульсных нейронных сетей / И.И. Мельников, К.А. Демиденков, И.А. Евсеенко, И.А. Емельянов // Информационные технологии. – № 7. – 2013 г. – С. 57-60.

Чередов Василий Васильевич

Магистрант

Белорусско-Российский университет, г. Могилев

Тел.: +375(29)1905706

E-mail: agitator@uvaga.by

Емельянов Игорь Александрович

Выпускник магистратуры

Белорусско-Российский университет, г. Могилев

Тел.: +375(29)3077342

E-mail: bujhm24@gmail.com

Зайцев Павел Сергеевич

Студент

Белорусско-Российский университет, г. Могилев

Тел.: +375(29)5477410

E-mail: pavel.zaitsau@hotmail.com

Евсеенко Игорь Антонович

Доцент кафедры автоматизированных систем управления, к-т техн. наук

Белорусско-Российский университет, г. Могилев

Тел.: +375(29) 844-32-13

E-mail: 327igor@rambler.ru