УДК 506.1, 519.7

Ю.Ю. Шамаева[1], А.А. Подорожняк[2]

[1]*Харьковский национальный университет им. В.Н. Каразина, Харьков*
[2]*Харковский университет Воздушных Сил им. И. Кожедуба, Харьков*

НЕЙРОСЕТЕВОЙ МЕТОД РАСПОЗНАВАНИЯ КОНЦЕПТОВ ЭМОЦИЙ

Рассмотрен подход к использованию нейросетевых технологий при решении задач, возникающих в процессе распознавания концептов эмоций в текстовой и вербальной информации. Предлагается для выделения эмоциональной насыщенности речи и текста использовать трехслойную прямонаправленную нейросеть с радиальными функциями возбуждения, позволяющую реализовать распознавание в режиме реального времени. Представлены результаты моделирования работы предложенного метода.

Ключевые слова: *эмоциональный концепт, прямонаправленная нейросеть, распознавание речи, распознавание текста.*

Постановка проблемы и анализ литературы

Совершенствование существующих и разработка принципиально новых высокоточных систем распознавания текста и речи является в настоящее время важной задачей в теории и практике совершенствования систем распознавания. Известно достаточно большое количество публикаций, посвященных данному вопросу, а так же разработано немало различных (и достаточно эффективных) технологий распознавания [1 – 5]. Наиболее популярными методами в решении подобного класса задач являются статистические методы распознавания, использующие априорные знания о частоте применения знаков, звуков и образов в текстовой и вербальной информации, а также спектральные характеристики речи. Кроме того, достаточно часто употребляемыми являются так называемые лингвистические методы распознавания, основанные на заданном словаре и правилах обработки информации.

Однако, наряду с очевидными достоинствами, лингвистические методы распознавания обладает рядом существенных недостатков, таких как:

– повышенная чувствительность данных систем к возможной изменениям в используемом словаре и различным модификациям правил произнесения (написания) слов и предложений для разных авторов текста и речи;

– количество вычислений, используемых при таких методах, достаточно велико и резко возрастает при необходимости увеличения точности распознавания, что обуславливает определенные трудности при обеспечении режима реального времени.

Таким образом, актуальными являются исследования, связанные с разработкой систем распознавания, в той или иной степени свободных от указанных недостатков.

В данной статье предлагается новый подход к построению систем распознавания речи и текста, основанный на использовании нейронной сети. Приведем определение для нейронной сети, данное в [4]: искусственная нейронная сеть (ИНС) – это существенно параллельный распределенный процессор, который обладает способностью к сохранению и репрезентации опытного знания. Она сходна с мозгом в двух аспектах:

– знания приобретаются сетью в процессе обучения;

– для сохранения знаний используются силы межнейронных соединений, называемые так же синаптическими весами.

Общепризнанным является тот факт, что ИНС – параллельная, существенно распределенная и адаптивная технология обработки информации. При этом вычислительная мощь ИНС определяется двумя фундаментальными свойствами, присущими только нейросетям [6]:

– существенно параллельной распределенной структурой (с возможностью модификации собственной топологии);

– способностью к обучению и репрезентации полученных знаний.

Благодаря исключительно высокой скорости обработки информации, способности к обучению и репрезентации знаний, а также способности нейросетей осуществлять сложные нелинейные преобразования "вход-выход", ИНС в настоящее время являются одной из лучших технологий, применяемых для обработки сигналов. В течение последних лет ИНС успешно применялись в различных областях науки и техники [7, 8], таких как обработка эмоционально окрашенных текстовых блоков [9], обработка и анализ речи, классификация образов, спектральный анализ, оценка параметров, оптимизация и др.

Целью настоящей статьи является разработка нейросетевого подхода к построению систем распознавания эмоциональной насыщенности текстовой и вербальной информации – нейросетевых систем

© Ю.Ю. Шамаева, А.А. Подорожняк

распознавания концептов эмоций (НССРКЭ).

Таким образом, именно возможность репрезентации знаний, полученных нейросетью в процессе обучения и способность правильно реагировать не только на сигналы, предъявленные в процессе тренировки, но также генерировать правильные выходы для входных сигналов, которые не были задействованы в процессе обучения, в совокупности с колоссальной скоростью собственно работы (обобщения) и послужили определяющими факторами при выборе ИНС, как наилучшего кандидата для разработки системы распознавания, удовлетворяющей поставленной цели.

1. Формирование исходных данных

Распознавание с помощью ИНС осуществляется по данным "вход-выход", полученным в ходе эксперимента с объектом распознавания (текстовые данные с различной эмоциональной насыщенностью) и определенными эмоциональными концептами. Объединив эти данные, можно создать полный набор тренировочных шаблонов (НТШ), описывающих наличие заданных эмоциональных концептов во входной информации.

Для обучения необходим следующий НТШ:

$$\Xi:\{\mathbf{x_i}, \mathbf{d_i}\}\,|_{\mathbf{i=1,...N}}\,,\ \mathbf{x}\in\Re^{p},\ d\in\Re^{m}, \qquad (1)$$

где $\mathbf{x_i}$ – входной шаблон; $\mathbf{d_i}$ – выходной шаблон; N – количество шаблонов; p – размер входного шаблона; m – размер выходного шаблона.

Формирование НТШ $\{\mathbf{x}, \mathbf{d}\}$ осуществляется, как правило, путем моделирования в лабораторных условиях с использованием реальной информации о текстовом (вербальном) значении различных слов и выражений (в том числе идиом) в контексте выявления эмоциональных концептов. При этом группы тестовых входных воздействий и желаемых откликов системы управления объединяются в НТШ, наиболее полно описывающие свойства искомых эмоциональных концептов, которые затем используются для обучения ИНС. Важно заметить, что исследователь свободен в выборе количества и структуры тестовых шаблонов, которые (исходя из личного опыта и интуиции исследователя) могут иметь место в качестве входных сигналов в условиях собственно функционирования НССРКЭ.

Ниже приведена методика создания НТШ для обучения ИНС, лежащей в основе НССРКЭ.

Для учета возможных реконфигураций объекта распознавания необходимо проанализировать (предсказать) возможные (умышленные и естественные) искажения эталонной информации (ЭИ), в качестве которой выступает эмоционально окрашенные текст или речь, и сформировать набор текстовых блоков Ψ (рис. 1), в котором:

D – количество возможных реконфигураций объекта распознавания;

$\mathbf{I_0}$ – нерекофигурированная (эталонная) информация;

$\mathbf{I_1}$ – информация, подвергнутая реконфигурации первого рода и т.д.

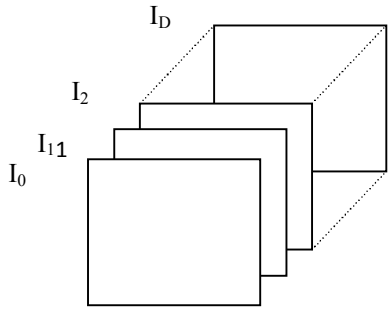

Рис. 1. Набор входных эмоционально окрашенных текстовых блоков для обучения ИНС НССРКЭ

Подвергнув каждый из эталонных эмоционально окрашенных текстовых блоков, принадлежащих пространству Ψ, трансформирующему преобразованию по всему континууму значений параметров (возможные умышленные и неумышленные ошибки в произношении, написании слов, а также особенности различных диалектов) сформируем выборку $\mathbf{x_i}$, i = 1, 2, …, N, называемую *входной выборкой.*

Сопоставив каждому эмоционально окрашенному текстовому блоку $\mathbf{x_i}$ вектор соответствующего эмоционального концепта $\mathbf{d_{il}}$, называемого *вектором желаемого выхода*, сформируем набор тренировочных шаблонов (НТШ) $\{x_i,d_i\}$, описывающих поведение системы распознавания.

2. Описание распознающей нейросети

Задача построения нейросетевой системы распознавания опирается на общую проблему интерполяции [3], которую для рассматриваемого случая можно сформулировать следующим образом:

По данному набору N эмоционально окрашенных текстовых блоков $\{\mathbf{x_i} \in \Re^{P}\,|\,i = 1, 2,\ldots, N\}$ и соответствующему набору векторов эмоциональных концептов (ВЭК) $\{\mathbf{d_i} \in \Re^{3}\}$, которая удовлетворяет следующему условию:

$$\mathbf{F(x_i) = d_i},\ i = 1, 2, \ldots, N, \qquad (2)$$

где i – размерность эмоционально окрашенного текстового блока.

Следует отметить, что при такой формулировке проблемы, интерполяционная поверхность (т.е. функция **F**) должна пройти через ***все*** точки НТШ.

Однако, на практике, условие (2) можно переписать как:

$$\|\mathbf{d_i - F(x_i)}\| \le \varepsilon, \qquad (3)$$

где ε – ошибка интерполяции; $\|.\|$ – евклидово расстояние (норма).

С учетом вышеизложенного, задачу данной работы можно сформулировать так:

по набору тренировочных шаблонов $\{\mathbf{x_i}, \mathbf{d_i}\}$, i = 1, 2,…, N необходимо сформировать ИНС, которая обладает следующими свойствами:

– при поступлении на вход системы текущего эмоционально окрашенного текстового блока **u**, совпадающего с любым эмоционально окрашенный текстовый блоком из входной выборки $\mathbf{x_i}$, i = 1, 2,…, N ИНС должна генерировать на выходе вектор эмоционального концепта $\mathbf{v} = F(\mathbf{u}) = F(\mathbf{x_i}) = \mathbf{d_i}$;

– при поступлении на вход системы текущего эмоционально окрашенного текстового блока **u**, не совпадающего ни с одним из $\mathbf{x_i}$, i = 1, 2,…, N, однако, принадлежащего пространству входных сигналов, ИНС должна генерировать на выходе вектор эмоционального концепта $\mathbf{v} = \mathbf{F(u)}$ такой, чтобы среднеквадратическая ошибка распознавания не превосходила некоторой наперед заданной величины, т.е. $\|\mathbf{v} - \mathbf{d_i}\| \le \varepsilon_{доп}$.

РБФ-сети. В настоящее время можно выделить 4 основных вида ИНС, отличающихся друг от друга архитектурой, правилами обучения и решаемыми задачами:

– однослойные прямонаправленные сети;
– многослойные прямонаправленные сети;
– рекуррентные сети;
– полностью связанные сети.

Проведенный анализ показал, что для задач распознавания целесообразно использовать трехслойную прямонаправленную сеть с радиально-базисными функциями (РБФ-сеть).

В общем случае РБФ-сеть состоит из трех слоев нейронов (рис. 2). Первый – "входной" слой, служит для приема и ретрансляции входного сигнала. Нейроны второго – "скрытого" слоя, осуществляют нелинейное преобразование входных сигналов. Нейроны третьего – "выходного" слоя осуществляют суммирование взвешенных выходных сигналов скрытого слоя и формируют выход сети.

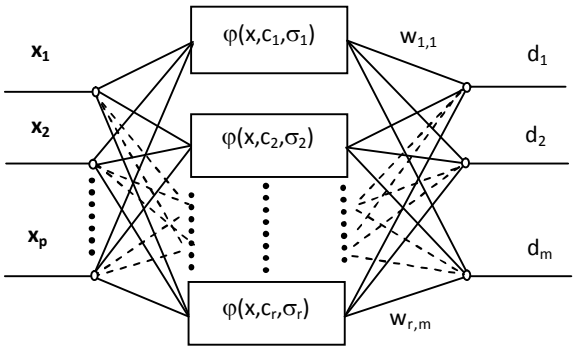

Рис. 2. Архитектура p-r-m РБФ-сети.

Рассмотрим процедуру обучения РБФ-сети по исходным данным, заданным в виде набора шаблонов, как она описана в [4]: пусть P – количество входов сети, R – количество нейронов скрытого слоя, M – количество выходов сети.

Введем следующие обозначения:

$\mathbf{x} = (\mathbf{x_1}, \mathbf{x_2},\ldots, \mathbf{x_p})$ – пространство входных сигналов;

$\mathbf{c} = (\mathbf{c_1}, \mathbf{c_2},\ldots, \mathbf{c_p})$ – вектор координат центров активационных функций нейронов скрытого слоя;

$\boldsymbol{\sigma_j}$, j = 1, 2,…, R – ширина окна активационной функции j-го нейрона скрытого слоя;

$\varphi(x, c, \sigma) = \exp\left(-\|x_i - c_j\|^2 / \sigma_j^2\right)$ – радиально-симметричная активационная функция нейрона скрытого слоя;

w_{ij} – вес связи между i-м нейроном выходного и j-м нейроном скрытого слоя.

Алгоритм синтеза и обучения РБФ-сети может быть представлен в виде:

1. Выберем размер скрытого слоя R, равным количеству тренировочных шаблонов. Синаптические веса нейронов скрытого слоя примем равными единице.

2. Разместим центры активационных функций нейронов скрытого слоя в точках **x** пространства входных сигналов сети, которые входят в набор тренировочных шаблонов Ξ:

$$\mathbf{c_j} = \mathbf{x_j},\ j = 1, 2, \ldots, R.$$

3. Выберем ширины окон активационных функций нейронов скрытого слоя $\boldsymbol{\sigma_j}$, j = 1, 2,…, R достаточно большими, но так, чтобы они не накладывались друг на друга в пространстве выходных сигналов.

4. Определим веса нейронов выходного слоя сети w_{ij}, i = 1, 2,…, R; j = 1,2,…, M, для чего предъявим сети весь набор тренировочных шаблонов Ξ и в результате получим набор линейных уравнений, который можно записать в матричном виде:

$$\Phi \bullet w = D, \qquad (4)$$

где w – матрица (размера R×M) синаптических весов; D – матрица (размера R×M) желаемых выходов (выходных шаблонов); Φ – интерполяционная матрица (размера R×R), элементы которой

$\varphi_{ij} = \exp\left(-\|x_i - c_j\|^2 / \sigma_j^2\right)$, где i = 1,…, R, j=1,…, R.

Решение системы уравнений в виде $w = \Phi^{-1} \bullet D$ обеспечивает прохождение интерполяционной поверхности через все точки тренировочного набора шаблонов.

Нейросетевая система распознавания. По своей сути НССРКЭ представляет собой частный случай классической системы распознавания [2], однако, ее характерной особенностью является отсутствие априорных знаний о математической модели объекта управления (плотности вероятностей случайных внешних воздействий либо дифференциальных уравнений, описывающих поведение системы); доступными для анализа и синтеза являются только ее входы и выходы.

Результаты моделирования. С использованием приведенного алгоритма синтеза и обучения была синтезирована ИНС, которая при испытании ее тестовыми шаблонами, а также эмоционально окрашенными текстовыми блоками не использовавшимися при обучении сети (полученными из тестового эмоционально окрашенного текстового блока применением неучтенной в НТШ реконфигурации), выдавала вектор эмоциональных концептов **v** такой, что экспертная оценка совпадала с полученной в 85% случаев.

Выводы

В данной статье была представлена НССРКЭ на основе РБФ-сети для решения задач распознавания эмоциональных концептов в речи и тексте. Предложенная система позволяет успешно решать поставленную задачу в случае реконфигурации (наличии случайных и преднамеренных ошибок в речи и тексте, различных диалектных особенностях и т.д.) объекта распознавания.

Однако, следует отметить, что структура РБФ-сети требует использования достаточно большого размера скрытого слоя. Данный факт определяет основной недостаток РБФ-сетей – *проклятие размерности* [4, 5, 10], т.к. с увеличением размерности входного пространства (количества тренировочных шаблонов) число необходимых нейронов скрытого слоя возрастает экспоненциально. Кроме того, при классической постановке задачи синтеза РБФ-сети остаются открытыми вопросы выбора ширин активационных функций и формирования НТШ, адекватно описывающих поведение объекта распознавания.

Авторами была также решена задача оптимизации РБФ-сетей (в смысле минимизации размера скрытого слоя без потери репрезентативности НТШ и выбора ширин активационных функций) на основе использования генетических алгоритмов. Результаты исследования эволюционного синтеза РБФ-сетей для задачи построения НССРКЭ оптимальной структуры, надеемся, будут опубликованы в следующей статье.

Список литературы

1. *Шульга Е.Н. Компьютерная герменевтика / Е.Н. Шульга // Вопросы философии. – 2007. – № 2. – С. 97-107.*
2. *Хакен Г. Синергетика, межуровневые нейронные сети и конгитивные карты / Г. Хакен, Дж. Португали. // Синергетика и психология: Выпуск 3: Конгитивные процессы; под ред. В.И. Аршинова, И.Н. Трофимовой, В.М. Шендяпина. – М.: Когито-Центр, 2004. – С. 129-154.*
3. *Нейросетевая система наведения летательных аппаратов / В.В. Сапожников, А.А. Подорожняк, С.И. Клевец и др. // Системи обробки інформації: зб. наук. пр. – Х.: НАНУ, ПАНМ, ХВУ, 2000. – Вип. 1 (7). – С. 129-138.*
4. *Haykin S. Neural networks: a comprehensive foundation / S. Haykin. – New York: Macmillan College Publishing Company, 1994. – 691 p.*
5. *Kussul N. Comparative analysis of neural networks and statistical approaches to remote sensing image classification / N. Kussul, S. Skakun, O. Kussul // Intern. Sci. J. of Computing. – 2006. – 5, № 2. – P. 93-99.*
6. *Рутковская Д. Нейронные сети, генетические алгоритмы и нечеткие системы / Д. Рутковская, М. Пилинський, Л. Рутковский; пер. с польск. И.Д. Рудинского. – М.: Горячая линия-Телеком, 2004. – 452 с.*
7. *Любченко Н.Ю. Анализ построения операционной семантики стратифицированных семантических сетей WEB-приложений / Н.Ю. Любченко // Системи обробки інформації: зб. наук. пр. – Х.: ХУ ПС, 2008. – Вип. 3 (70). – С. 102-105.*
8. *Нейросетевой метод представления и нейросетевое распознавание частотно-временных векторов речевой информации [Электронный ресурс] / Ю. Баяковский, А. Жирков, Д. Корчагин, А. Лукин, А. Крылов // Программирование. – М.: МГУ, 2003. – Вип. 4. – С. 41-52. – Режим доступа к документу: http:// graphics.cs.msu.ru/ en/publications/text/prog2003zh.pdf.*
9. *Капитонова Т.А. Нейросетевое моделирование в распознавании образов: философско-методологические аспекты / Т.А. Капитонова. – Минск: Белорусская наука, 2009. – 376 с.*
10. *Шамаева Ю.Ю. Этнокультурное своеобразие конгитивной модели радости в английском, украинском и русском языках / Ю.Ю. Шамаева // Нова філологія. – 2002. – № 4 (15). – С. 142-148.*

Поступила в редколлегию 28.08.2009

Рецензент: д-р техн. наук, проф. Л.Ф. Купченко, Харьковский университет Воздушных Сил им. И. Кожедуба, Харьков.

НЕЙРОМЕРЕЖЕВИЙ МЕТОД РОЗПІЗНАВАННЯ КОНЦЕПТІВ ЕМОЦІЙ

Ю.Ю. Шамєва, А.О. Подорожняк

Розглянуто підхід до використання нейромережевих технологій при вирішенні задач, які виникають в процесі розпізнавання концептів емоцій в текстовій і вербальній інформації. Пропонується для виділення емоційній насиченості мови і тексту використовувати тришарову прямо спрямовану нейромережу з радіальними функціями збудження, таку, що дозволяє реалізувати розпізнавання в режимі реального часу. Представлено результати моделювання роботи запропонованого методу.

Ключові слова: *емоційний концепт, прямо спрямована нейромережа, розпізнавання мови, розпізнавання тексту.*

NEURONET METHOD OF RECOGNITION OF EMOTIONS KONCEPT

J.Y. Shamaeva, A.A. Podorognyak

Approach is considered to using of neuronet technologies for the decision of tasks of arising up in the process of recognition of koncepts emotions in text and verbal information. It is suggested for a selection the emotional saturation of speech and text to use three-layered straight-directed neuronet with the radial functions of excitation, allowing to realize recognition in the mode of the real time. The results of design of work of the offered method are presented.

Keywords*: emotional koncept, neuronet, recognition of language, recognition of text.*