



стурных признаков Харалика для решения задачи построения топографической карты с помощью древовидного текстурного классификатора, составленного из нескольких многослойных перцептронов.

Для обучения и тестирования перцептронов, входящих в состав классификатора, был использован набор снимков, полученных при помощи БПЛА вертолетного типа и покрывающих нечерноземные области средней России. Структура перцептронов была подобрана с помощью алгоритмов глобальной оптимизации.

На рисунке 1 приведены результаты сравнения качества обучения простых классификаторов (отдельных многослойных перцептронов), использующих для описания текстур, покрывающих исходные изображения, энергетические характеристики Лавса или текстурные признаки Харалика.

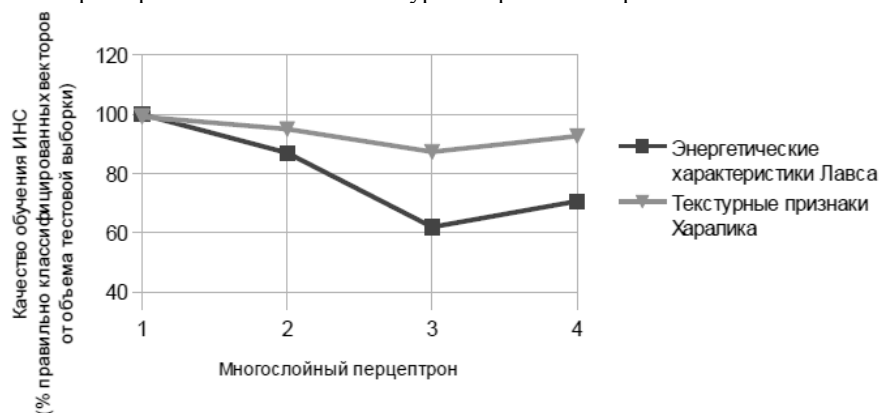


Рисунок 5 - Сравнение качества обучения классификаторов

Таким образом, можно сделать вывод, что текстурные признаки Харалика лучше подходят для решения задач обработки данных, полученных с помощью БПЛА, чем энергетические характеристики Лавса.

К недостаткам текстурных признаков Харалика можно отнести существенный объем вычислений, требуемых для их расчета из-за необходимости анализа нескольких матриц смежности. Данный недостаток, однако, не оказывает существенного влияния на выбор между признаками Харалика и характеристиками Лавса, так как информационная недостаточность последних ведет к существенному увеличению времени обучения классификатора. [3]

Литература

1. Акинин М.В., Акинина Н.В., Никифоров М.Б., Соколова А.В., Таганов А.И. Нейросетевой метод оперативного картографирования с использованием беспилотного летательного аппарата. // Динамика сложных систем - XXI век № 2 (2015). - Москва: Радиотехника. - 2015. - с. 9 - 14.



2. M. Haralick R., K. Shanmugam, I. Dinstein. Textural features for image classification // IEEE Transactions on systems, man and cybernetics. Vol. SMC 3, № 6. 1973. С. 610 – 621.

3. М.В. Акинин, А.А. Логинов, М.Б. Никифоров. Способы описания текстур в задачах построения топографических карт // Материалы XI Международной научно-технической конференции .АВИА — 2013. (том 4). 2013.

А.К. Алимуратов, П.П. Чураков

АДАПТИВНЫЙ МЕТОД ПОВЫШЕНИЯ ЭФФЕКТИВНОСТИ ГОЛОСОВОГО УПРАВЛЕНИЯ

(Пензенский государственный университет)

Увеличение производительности интеллектуальных информационных систем требует от пользователя устойчивых навыков управления. Наиболее адаптивным интерфейсом взаимодействия пользователя с системой голосовое управление, основанное на технологии распознавания речи. Работа в направлении повышения эффективности голосового управления ведется достаточно активно. На сегодняшний день представлено большое количество алгоритмов, повышающих точность распознавания и эффективность голосового управления: «RWTH ASR» [1], «Julius» [2] и «CMU Sphinx» [3].

Как известно, точность распознавания зависит от корректной обработки речевых команд, которая определяется точностью измерения его амплитудных, временных, частотных и энергетических характеристик. Основная причина низкой точности и больших погрешностей в измерениях связана с использованием неадаптивных методов обработки сложных нестационарных речевых сигналов. Исследования существующих методов обработки речевых сигналов [4], выявили перспективность использования адаптивной технологии анализа нестационарных данных – комплементарная множественная декомпозиция на эмпирические моды (КМДЭМ) [5].

Целью данной работы является разработка алгоритма, повышающего точность распознавания речевых команд и эффективность голосового управления за счет применения КМДЭМ. Статья является развитием ранее опубликованных трудов авторов [7].

Комплементарная множественная декомпозиция на эмпирические моды

КМДЭМ представляет собой адаптивную технологию разложения сигнала на эмпирические моды (ЭМ). Адаптивность метода заключается в том, что базисные функции, используемые для разложения, извлекаются непосредственно из исходного сигнала. Аналитическое выражение декомпозиции имеет следующий вид [5]:



$$x(t) = \sum_{i=1}^I IMF_i(t) + r_i(t),$$

где $x(t)$ – исходный сигнал, i – номер ЭМ, I – количество ЭМ, $IMF_i(t)$ – конечное число извлекаемых ЭМ, $r_i(t)$ – результирующий остаток.

Особенность метода КМДЭМ заключается в многократном добавлении к исходному речевому сигналу белого шума с прямыми и инверсными значениями амплитуды и вычислении среднего значения полученных мод, как конечно-го истинного результата независимо от того, сколько сигналов белого шума использовалось:

$$\begin{bmatrix} x_j(t) \\ x_j(t)^* \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \times \begin{bmatrix} x(t) \\ \omega_j(t) \end{bmatrix},$$

$$IMF_i(t) = \frac{\sum_{j=1}^J IMF_{ji}(t)}{J}, \quad r_i(t) = \frac{\sum_{j=1}^J r_{ji}(t)}{J},$$

где $x_j(t)$ – зашумленный белым шумом речевой сигнал, $x_j(t)^*$ – зашумленный инверсным по знаку белым шумом речевой сигнал, $\omega_j(t)$ – добавляемый белый шум, $IMF_{ji}(t)$, $r_{ji}(t)$ – ЭМ и результирующий остаток, полученные при различных декомпозициях, $j=1, 2, \dots, J$ – количество циклов декомпозиций (добавлений к сигналу белого шума).

КМДЭМ в полной мере использует преимущество статистических характеристик белого шума, для обнаружения слабых периодических участков речевых сигналов с минимальным значением остаточного шума.

Алгоритм, повышающий точность распознавания и эффективность голосового управления на основе КМДЭМ

На рисунке 1 представлена блок-схема алгоритма, повышающего распознавание и эффективность голосового управления на основе КМДЭМ.

Суть предложенного алгоритма заключается в уменьшении разницы между поступающими речевыми командами и шаблонами (полученными в ходе обучения), посредством выделения из исходного сигнала только полезной информации – уникальные свойства голоса. Ключевым понятием при выделении полезной информации является формирование набора сигналов на основе информативности ЭМ. При условии, что речевой сигнал имеет конечную энергию, число ЭМ при разложении всегда является конечным. Для абсолютно произвольного сигнала все ЭМ можно разбить на две категории информативные ЭМ с шумовыми и сигнальными составляющими и неинформативные ЭМ с трендовыми и компенсирующими составляющими.

Информативные ЭМ в разложении всегда отражают внутреннюю структуру и особенности речевого сигнала. К их числу относятся шумовые и сигнальные ЭМ. Появление в разложении первых объясняется наличием в исходном сигнале остаточного шума, а вторые связаны непосредственно с полезным сигналом и входящими в него компонентами. Неинформативные ЭМ являются медленно меняющимися функциями. Среди них выделяют трендовые ЭМ, опи-



сывающие истинную динамику среднего значения сигнала и компенсирующие ЭМ возникающие при разложении.

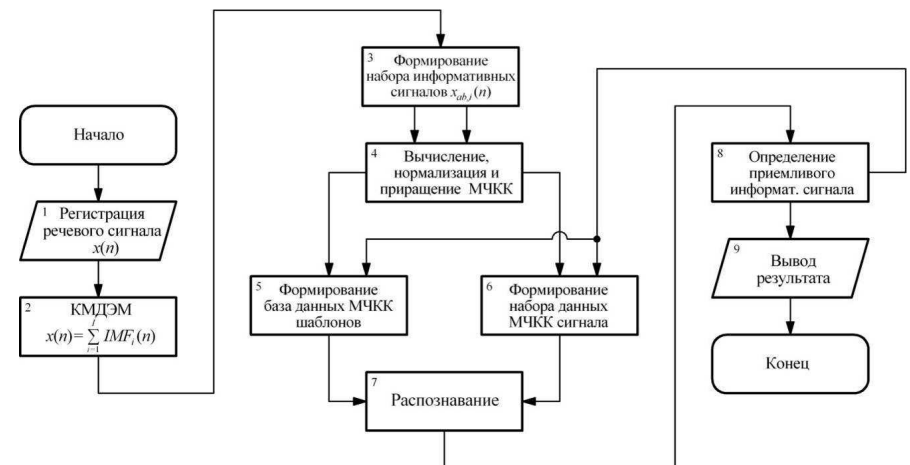


Рисунок 1 – Блок-схема алгоритма, повышающего распознавание и эффективность голосового управления на основе КМДЭМ

Формирование набора информативных сигналов заключается в вычитании из исходного сигнала речевой команды информативных шумовых и неинформативных ЭМ. Информативными шумовыми обычно являются первые две или три ЭМ, в зависимости от интенсивности присутствующего в сигнале шума. Неинформативными являются последние три или четыре ЭМ, в зависимости от общего количества мод (число ЭМ примерно равно двоичному логарифму от числа отсчетов в сигнале). Формирование набора информативных сигналов осуществляется по формуле:

$$x_{ab,i}(n) = x(n) - \left(a \times \sum_{i=0}^2 IMF_{i+1}(n) + b \times \sum_{i=0}^2 IMF_{I-1}(n) \right),$$

где $x_{ab,i}(n)$ – информативный сигнал, $x(n)$ – исходный сигнал речевой команды, i – номер ЭМ, I – количество ЭМ, a , b – коэффициенты, определяющие участие ЭМ в формировании набора информативных сигналов.

Целью формирования набора информативных сигналов является возможность выбора одного сигнала, содержащего максимально большее количество информации об уникальных свойствах голоса.

В качестве информативных параметров в предложенном алгоритме используются мел-частотные кепстральные коэффициенты МЧКК [7], которые наиболее приемлемы для командных систем голосового управления.

В экспериментально-исследовательском комплексе, реализованном в пакете прикладных программ *MATLAB*, проведено исследование предложенного алгоритма распознавания. Цель исследования являлось: определение наиболее



приемлемого информативного сигнала, обеспечивающего наименьшую разницу между поступающей в систему речевой командой и шаблоном. В качестве критерия оценки эффективности распознавания использовалась точность распознавания. В таблице 1 представлены результаты определения наиболее приемлемого информативного сигнала.

Таблица 1 – Зависимость точности распознавания от номера информативного сигнала

Параметр	Исходный речевой сигнал	Номер информативного сигнала				
		1	2	3	4	5
A, %	85,1	92,3	95,2	93,9	97,1	96,2
Δ	2,47	3,88	5,22	4,55	5,84	5,74

В соответствии с полученными результатами наилучшая точность распознавания достигается при использовании информативного сигнала № 4 (рисунок 3).

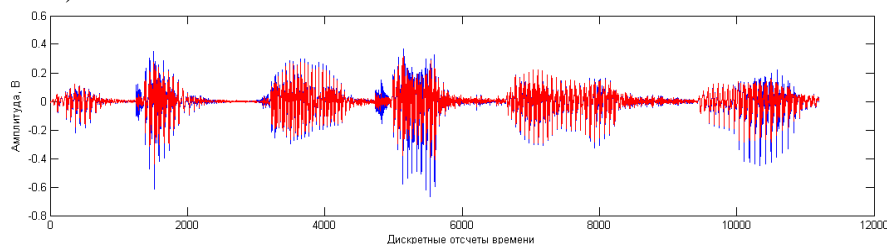


Рисунок 2 – Информативный сигнал № 4

Вывод

Использование разработанного алгоритма распознавания речевых команд на основе КМДЭМ позволит значительно повысить эффективность голосового управления.

Литература

1. R. David, G. Christian, H. Georg, and N. Hermann, “The RWTH Aachen University Open Source Speech Recognition System,” Human Language Technology and Pattern Recognition Computer Science Department, RWTH Aachen University, Germany.
2. A. Lee, T. Kawahara, and K. Shikano, “Julius – an open source real-time large vocabulary recognition engine,” Proc. Europ. Conf. Speech Com. Technol., Aalborg, Denmark, Sep. 2001, pp. 1691–1694.
3. W. Walker, P. Lamere, P. Kwok, Raj R.S. Bhiksha, E. Gouvea, P. Wolf, J. Woelfel, “Sphinx-4: A flexible open source framework for speech recognition,” Tech. Rep. SMLITR-2004-139, Sun Microsystems., Inc., Nov. 2004.
4. Алимуратов А.К. Обзор и классификация методов обработки речевых сигналов в системах распознавания речи / А.К. Алимуратов, П.П. Чураков // Измерение. Мониторинг. Управление. Контроль. - 2015. - № 2 (12). - С. 27 - 35.



5. J.-R. Yeh, J.-S. Shieh, and N.E. Huang, “Complementary ensemble empirical mode decomposition: A novel noise enhanced data analysis method,” Adv. Adapt. Data Anal., vol. 2(2), 2010, pp. 135–156.

6. Алимуратов А.К., Муртазов Ф.Ш. Методы повышения эффективности распознавания речевых сигналов в системах голосового управления. Измерительная техника. - 2015. - № 10. - С. 20 - 24.

7. Алимуратов А.К. Выбор оптимального набора информативных параметров речевых сигналов для систем голосового управления / А.К. Алимуратов, П.П. Чураков, А.Ю. Тычков // Измерение. Мониторинг. Управление. Контроль. - 2013. - № 1 (3). - С. 16 - 20.

Ф.Р. Аляутдинова, В.А. Суздальцев

СТАТИСТИЧЕСКИЕ МЕТОДЫ ПРОГНОЗНОЙ ДИАГНОСТИКИ

(Казанский национальный исследовательский технический университет
им. А.Н. Туполева)

Прогнозная диагностика используется для выявления проблемных ситуаций связанных с отказами, которые могут возникнуть при протекании процессов в системе, и требует немедленного принятия решения для поддержки ее работоспособности в будущем времени.

Возникновению проблемных ситуаций способствует множество взаимно зависимых случайных факторов, поэтому уместно рассматривать задачи прогнозной диагностики в вероятностно статистической постановке, а при разработке методов и алгоритмов их решения использовать результаты, полученные в теории статистических решений [1].

Пусть X – признаковое пространство, а $x=(x_1, x_2, \dots, x_n)$ - вектора признаков текущей ситуации, $x \in X$. Известными считаются априорные вероятности возникновения и отсутствия проблемной ситуации $P(A)$ и $P(A^*)$, $P(A^*)=1-P(A)$. Обозначим через $P(x|A)$ и $P(x|A^*)$ условные вероятности возникновения текущей ситуации с вектором признаков x , при условии, что проблемная ситуация имела место в будущем либо отсутствовала. Решение о том, что развитие текущей ситуации приведет к проблемной ситуации, принимается при выполнении следующего условия:

$$P(A|x) \geq P_{кр}$$

где $P_{кр}$ – пороговая вероятность принятия решения.

Для определения $P(A|x)$ используется формула Байеса:

$$P(A|x) = \frac{P(A) \times P(x|A)}{P(A) \times P(x|A) + P(A^*) \times P(x|A^*)}$$

Значение пороговой вероятности $P_{кр}$ является решением задачи минимизации целевой функции:

$$C(P_{кр}) = c_1 \times p_1(P_{кр}) + c_2 \times p_2(P_{кр}) \rightarrow \min$$