

УДК 629.7.014-519.05:004.934

doi: 10.32620/reks.2018.3.03

А. Ю. ЛАВРИНЕНКО, Ю. А. КОЧЕРГИН, Г. Ф. КОНАХОВИЧ

Национальный авиационный университет, Киев, Украина

СИСТЕМА РАСПОЗНАВАНИЯ СТЕГАНОГРАФИЧЕСКИ-ПРЕОБРАЗОВАННЫХ ГОЛОСОВЫХ КОМАНД УПРАВЛЕНИЯ БПЛА

Разработана система распознавания стеганографически-преобразованных голосовых команд управления беспилотным летательным аппаратом на основе кепстрального анализа, которая позволяет обеспечить эффективное распознавание и скрытную передачу команд на борт беспилотного летательного аппарата, посредством преобразования голосовых команд управления в своего рода стеганографический вектор признаков, что подразумевает скрытие информации голосового управления беспилотным летательным аппаратом. Синтезирована математическая модель алгоритма вычисления мел-частотных кепстральных коэффициентов и классификатора распознавания голосовых команд управления для решения задачи семантической идентификации и обеспечения скрытности информации управления беспилотного летательного аппарата в канале связи. Разработан программный комплекс, включающий средства для составления базы эталонных голосовых образов субъектов управления для обучения и тестирования системы распознавания стеганографически-преобразованных голосовых команд управления беспилотным летательным аппаратом на основе кепстрального анализа и компьютерные модели предложенных методов и алгоритмов распознавания голосовых команд управления в среде MATLAB. Обосновано и экспериментально доказано целесообразность использования предложенной системы распознавания стеганографически-преобразованных голосовых команд управления беспилотным летательным аппаратом на основе кепстрального анализа. Представлен алгоритм вычисления мел-частотных кепстральных коэффициентов, которые выступают в роли основных признаков распознавания и результатом стеганографического преобразования речи, где для оценивания результатов автоматического распознавания голосовых команд управления используется классификатор, построенный по критерию минимального расстояния в роли которого выступает дисперсия разности математического ожидания мел-частотных кепстральных коэффициентов. Полученные результаты экспериментального исследования позволяют сделать вывод о целесообразности дальнейшего практического применения разработанной системы распознавания стеганографически-преобразованных голосовых команд управления беспилотным летательным аппаратом на основе кепстрального анализа.

Ключевые слова: стеганографическое преобразование речи, голосовые команды управления, распознавание голосовых команд, кепстральный анализ, кепстральные коэффициенты, преобразование Фурье, косинусное преобразование, система голосового управления, беспилотный летательный аппарат.

Введение

Несмотря на бурное развитие вычислительной техники, задача распознавания голосовых команд по-прежнему не может считаться полностью решенной. При этом ее актуальность со временем только увеличивается. Стеганографическое скрытие передаваемой информации также является актуальной задачей в связи с тем, что с каждым годом существенно увеличивается количество передаваемых данных по существующим каналам связи. Наиболее удобными и эффективными средствами взаимодействия человека с техническими объектами такими как беспилотные летательные аппараты (БПЛА) были бы те, которые являются естественными для него. В связи с этим все более актуальной становится потребность в альтернативных, более естественных методах управления БПЛА. Одним из наиболее

естественных методов управления для человека, является управление посредством голосовых команд с обеспечением скрытной стеганографической передачей информации управления на борт БПЛА.

1. Анализ существующих исследований и публикаций

В научном сообществе существуют известные исследования, проведенные за последние годы в области распознавания речи и голосового управления, результаты которых рассмотрены и проанализированные в представленных публикациях [1-4].

Предлагаемые в указанных источниках решения эффективного распознавания речевых сигналов, довольно хорошо справляются с положенной на них задачей, а именно, качественное распознавание речи, в арсенале которых применяются различные

классификаторы распознавания: скрытые Марковские модели [1], нейронные сети [2], модели гауссовых смесей [3], векторное квантование [4], соответственно. Перечисленные методы показывают превосходные результаты по распознаванию речи с точки зрения наилучшего качества распознавания (97 % точности распознавания), но с другой стороны у них есть недостатки: все они являются довольно громоздкими в программной реализации, имеют высокие требования к производительной мощности процессоров, большую задержку в вычислении алгоритмов, требуют достаточных финансовых затрат на производство, что в свою очередь приведет к подорожанию себестоимости таких систем, а также они не позволяют обеспечить стеганографическую скрытность передаваемой речевой информации управления на борт БПЛА, что является по сути самым главным минусом этих систем.

В данной работе авторы предприняли попытку по возможности устранить существующие недостатки и использовать видимые преимущества в выделении информативных признаков распознавания в задачах обеспечения стеганографической скрытности передачи информации управления на борт БПЛА.

2. Постановка проблемы

В процессе анализа существующих известных исследований [1-4], проведенных за последние годы в области распознавания речи (голоса), голосового управления, возникает главная проблема, связанная с разработкой системы распознавания голосовых команд управления БПЛА которая бы соответствовала таким требованиям:

- 1) простота в программной реализации;
- 2) низкие требования к производительности процессоров;
- 3) низкая временная задержка при вычислении алгоритмов;
- 4) небольшие финансовые затраты на производство, что в свою очередь приведет к удешевлению себестоимости таких систем;
- 5) точность распознавания речи должна быть не меньше 90 %.
- 6) скрытность самого факта передачи информации управления т.е. чтобы злоумышленник не понял, что передается речевой сигнал.

Проанализированные существующие методы [1-4] соответствуют только 5-му требованию, во всех остальных случаях они являются полной противоположностью заданным требованиям.

Разработанная авторами статья система распознавания стеганографически-преобразованных голосовых команд управления БПЛА на основе кепстрального анализа решает данную проблему в сущ-

ности своей с помощью алгоритма вычисления информативных признаков распознавания, в качестве которых использовались мел-частотные кепстральные коэффициенты (МЧКК) для семантической идентификации голосовых команд управления, где для оценивания результатов автоматического распознавания голосовых команд управления используется классификатор, построенный по критерию минимального расстояния в роле которого выступает дисперсия разности математического ожидания МЧКК.

Еще один очень важное преимущество разработанной системы является тот факт, что скрытность передаваемой информации на борт БПЛА обеспечивается уже по факту преобразования речевого сигнала в вектор признаков (МЧКК), т.е. осуществляется стеганографическое преобразование речи с уменьшением избыточности речевого сигнала в тысячи раз, что позволит его без труда скрыть любыми стеганографическими методами.

Итак, цель статьи состоит в следующем:

- разработать систему распознавания голосовых команд управления БПЛА на основе кепстрального анализа для создания стеганографически-скрытного диалога между субъектом управления и бортом БПЛА, которая бы удовлетворяла всем перечисленным выше требованиям.

Для достижения поставленной цели необходимо решить следующие задачи:

- синтезировать математическую модель алгоритма вычисления МЧКК и классификатора распознавания голосовых команд управления для решения задачи семантической идентификации и обеспечения скрытности информации управления БПЛА в канале связи;
- синтезировать компьютерную модель разработанной системы распознавания стеганографически-преобразованных голосовых команд управления БПЛА на основе кепстрального анализа и промоделировать в среде MATLAB на примере идентификации произнесенных разными субъектами управления команд: «вверх», «вниз», «вправо», «влево»;
- осуществить сравнительную оценку результатов моделирования.

3. Система распознавания стеганографически-преобразованных голосовых команд управления БПЛА

Разработанная система распознавания стеганографически-преобразованных голосовых команд управления БПЛА на основе кепстрального анализа имеет два режима работы: режим обучения и режим распознавания (тестирования). Эти режимы входят в функциональную схему системы распознавания го-

ловых команд управления БПЛА на основе кепстрального анализа (рис. 1), задача, которой заключается в первичной обработке речевого сигнала и выделения признаков распознавания, в качестве которых используются МЧКК. МЧКК также являются результатом стеганографического преобразования и по своей сути обеспечивают представление речевого сигнала в форму непригодную для его обнаружения в канале связи, т.е. злоумышленник не обнаружит сообщение, скрытое одним из стеганографических методов защиты информации. Дальнейшее поведение системы зависит от режима работы. Если система находится в режиме обучения, полученные на этапе выделения признаков распознавания значения МЧКК, сохраняются в базу эталонных голосовых образов. При нахождении системы в состоянии распознавания, набор значений МЧКК голосовой команды субъекта управления последовательно сравнивается со всеми наборами значений МЧКК из базы эталонных голосовых образов. Задача функции принятия решения определить наилучший результат сравнения по одному из заданных критериев и выдать результат распознавания [5].

В разработанной системе распознавания голосовых команд в качестве информативных признаков распознавания использовались МЧКК для семантической идентификации голосовых команд. Выделенные признаки будут использованы для формиро-

вания базы эталонных голосовых образов или для сравнения с зарегистрированными голосовыми образами (МЧКК) в базе в зависимости от выбранного режима работы системы.

Алгоритм вычисления МЧКК (рис. 2) включает в себя следующие шаги.

На вход алгоритма подаётся речевой сигнал частотный диапазон, которого весьма ограничен и располагается в интервале от 300 до 4000 Гц. Из данного факта следует, что путем моделирования полосового фильтра можно отбросить частотные составляющие, которые находятся за пределами этого диапазона и соответственно, не несут смысловой нагрузки. Сложившийся подход к процедуре обработки речевых сигналов состоит в использовании кратковременного анализа. То есть сигнал разбивается на временные окна фиксированного размера, на которых параметры сигнала не меняются. Для речевого сигнала размер окна обычно выбирается в пределах 10-30 мс. Для более точного представления сигнала между окнами делают перекрытие, равное половине длины окна. Затем к каждому окну применяются алгоритм выделения признаков распознавания (МЧКК) [7]. Исходя из вышесказанного, прошедший предварительную обработку речевой сигнал разбивается на K кадров по N отсчетов, пересекающихся на $1/2$ длины кадра [6].



Рис. 1. Функциональная схема системы распознавания голосовых команд управления БПЛА на основе кепстрального анализа



Рис. 2. Алгоритм вычисления МЧКК

На вход блока дискретного преобразования Фурье (ДПФ) подаётся последовательность отсчётов участка речевого сигнала (K -го кадра), исследуемого на данной итерации, x_0, \dots, x_{N-1} . К данной последовательности применяется весовая функция и затем ДПФ. В качестве весовой функции используется окно Хэмминга, которое имеет следующий вид:

$$w[n] = 0.53836 - 0.46164 \cdot \cos\left(2\pi \frac{n}{N-1}\right), \quad (1)$$

$$n = 0, \dots, N-1,$$

где, N – длина окна, выраженная в отсчетах.

Тогда ДПФ взвешенного речевого сигнала можно записать в виде формулы:

$$X[k] = \sum_{n=0}^{N-1} x[n]w[n]e^{-\frac{2\pi j}{N}kn}, \quad k = 0, \dots, N-1. \quad (2)$$

Значения индексов k соответствуют частотам:

$$f[k] = \frac{F_s}{N}k, \quad k = 0, \dots, N/2, \quad (3)$$

где, F_s – частота дискретизации речевого сигнала.

Полученное представление речевого сигнала в частотной области разбивают на диапазоны с помощью банка (гребёнки) треугольных фильтров. Умножая функцию на фильтр, мы усредняем ее на некотором участке. Каждый треугольный фильтр находит взвешенное среднее тех амплитудных спектральных значений, соответствующих частотам в пределах между нижней и верхней частотой для данного фильтра. Если амплитуда соответствует точно средней частоте полосы, то она умножается на коэффициент равный единице. При передвижении соответствующей амплитудному значению частоты от середины к нижней или верхней границам, коэффициент уменьшается от единицы до нуля. Полученные произведения амплитуд на коэффициенты добавляются и делятся на число амплитудных значений. В результате находим взвешенное среднее для данной полосы частот [7].

Границы фильтров рассчитывают в мел-шкале. Данная шкала является результатом исследований по способности человеческого уха к восприятию звуков на различных частотах.

Перевод в мел-частотную область осуществляют по формуле:

$$M = 1127.01048 \ln(1 + F/700). \quad (4)$$

Обратное преобразование в герцы выражается по формуле:

$$F = 700(e^{M/1127.01048} - 1). \quad (5)$$

Формула для разделения оси частот на треугольные фильтры будет иметь следующий вид:

$$f[m] = \left(\frac{N_f}{F_s}\right)M^{-1} \left(M(F_{\min}) + m \frac{M(F_{\max} - F_{\min})}{N_f + 1} \right), \quad (6)$$

где, N_f – количество мел-фильтров (обычно используют порядка 24 фильтров), F_s – частота дискретизации, $M(F)$ – функция перевода частоты в герцах в частоту в мелах, рассмотренная ранее, $M(F_{\max} - F_{\min})$ – исследуемый диапазон частот в мелах, который разбивается на N_f равномерно распределённых перекрывающихся диапазонов и вычисляются соответствующие границы в области линейных частот.

Составим гребенку треугольных фильтров по следующей формуле:

$$H_m[k] = \begin{cases} 0, & k < f[m-1] \\ \frac{(k - f[m-1])}{(f[m] - f[m-1])}, & f[m-1] \leq k < f[m] \\ \frac{(f[m+1] - k)}{(f[m+1] - f[m])}, & f[m] \leq k \leq f[m+1] \\ 0, & k > f[m+1] \end{cases}, \quad (7)$$

где, $H_m[k]$ – весовые коэффициенты полученных фильтров.

Применение фильтра заключается в попарном перемножении его значений со значениями спектра. Поскольку фильтров у нас N_f , коэффициентов будет столько же. Фильтры применяются к квадратам модулей коэффициентов ДПФ, то есть, нам нужно применить мел-фильтры не к значениям спектра, а к его энергии. Для этого нужно вычислить энергию для каждого окна, после чего прологарифмировать полученные результаты. Считается, что таким образом понижается чувствительность коэффициентов к шумам.

Полученные значения энергии спектра логарифмируются и представляются следующим образом:

$$E[m] = \ln \left(\sum_{k=0}^{N-1} |X[k]|^2 H_m[k] \right), \quad m = 0, \dots, N_f - 1. \quad (8)$$

Заключительным этапом в вычислении МЧКК для уменьшения количества выходных параметров и декорреляции компонентов является ДКП, которое задается следующей формулой:

$$c[n] = \sum_{m=0}^{N_f-1} E[m] \cos \left(\frac{\pi n \left(m + \frac{1}{2} \right)}{N_f} \right), \quad n = 0, \dots, N_f - 1. \quad (9)$$

Это преобразование обладает свойством компактности энергии: большей энергии соответствует меньшее количество информации. Полученный набор значений называется МЧКК. Коэффициент $c[0]$ не используется, так как представляет энергию речевого сигнала.

Таким образом, мы имеем очень небольшой набор значений, который при распознавании успешно заменяет тысячи отсчетов речевого сигнала. Обычно сохраняется только первые несколько элементов (от 8 до 16), по которым в дальнейшем производится идентификация голосовых команд управления [8].

На рис. 3 изображены вычисленные по разработанному алгоритму МЧКК голосовых команд субъекта управления №1: «вверх», «вниз», «вправо», «влево».

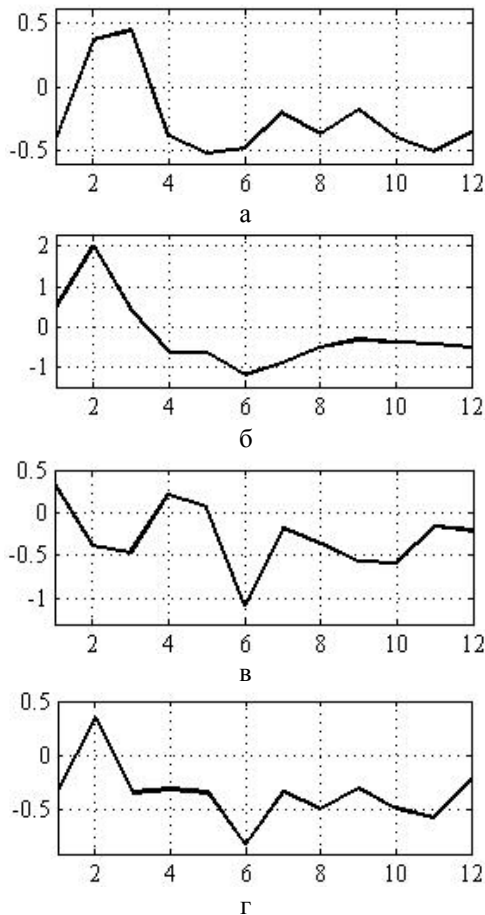


Рис. 3. МЧКК голосовых команд субъекта управления №1: а – «вверх», б – «вниз», в – «вправо», г – «влево»

В этом и заключается главный аспект рассмотренного стеганографического преобразования в решении поставленной задачи. Дело в том, что любое преобразование по своей сути является методом шифрования данных, но в задачах стеганографической защиты информации важную информацию не нужно зашифровывать, ее нужно просто преобразовать таким способом, чтобы не был виден сам факт передачи информации управления на борт БПЛА. В данном случае наглядно видно, что из нескольких тысяч отсчетов речевого сигнала после стеганогра-

фического преобразования получается единицы отсчетов МЧКК, что говорит о возможности эффективного скрытия данной информации любым из возможных методов стеганографии, так как информации стало в тысячи раз меньше чем в первоначальном варианте, а значит и скрыть ее не составит большого труда. Как можно понять из вышеизложенного МЧКК служат как признаки для распознавания голосовых команд управления, так и выступают в роли коэффициентов стеганографического преобразования речевой информации управления БПЛА.

На следующем этапе происходит процедура классификации выделенных МЧКК произнесенных голосовых команд управления по результатам которой, функция принятия решений выдает результат о семантической идентификации голосовых команд. Функция принятия решений задается в заданном пороге допустимого распознавания Θ . В данной системе для оценивания результатов автоматического распознавания голосовых команд управления используется классификатор, построенный по критерию минимального расстояния. В качестве такого показателя выступает дисперсия разности математического ожидания МЧКК сохраненных в базе эталонных голосовых образов с математическим ожиданием МЧКК на уровне тестирования системы.

Дисперсия разности математических ожиданий двух образцов голосовых команд управления (МЧКК), записывается следующим образом:

$$D = \frac{\sum_{i=1}^n \left(\frac{\sum_{i=1}^n x_i}{n} - \frac{\sum_{i=1}^n \bar{x}_i}{n} \right)^2}{n}, \quad (10)$$

где, x_i – МЧКК сохраненные в базе эталонных голосовых образов, \bar{x}_i – МЧКК на уровне тестирования системы, n – количество МЧКК. Решение о семантической идентификации голосовых команд принимается по критерию минимальной дисперсии, то есть наименьшее отклонение сравниваемых МЧКК в определенном пороге распознавания, который задается следующим выражением:

$$\begin{aligned} & \text{if } D_{\min} < \Theta \\ & \text{"identified!"} \\ & \text{else} \\ & \text{"not identified!"} \\ & \text{end} \end{aligned} \quad (11)$$

где, D_{\min} – минимальная дисперсия, $\Theta = 1 - \Delta$ – заданный порог допустимого распознавания (на практике обычно используют $\Delta = 0.80 \dots 0.90$).

Минимальная дисперсия, вошедшая в заданный порог допустимого распознавания, является наилучшим результатом сравнения, а значит, команда идентифицируется (распознается) – "identified!". В противном случае голосовая команда не проходит семантическую идентификацию (не распознается) – "not identified!" [9].

4. Результаты исследований

В статье подробно изложены полученные результаты предварительных экспериментальных исследований, на основе которых сделаны выводы о целесообразности дальнейшего научного и практического применения системы распознавания стеганографически-преобразованных голосовых команд управления БПЛА на основе кепстрального анализа и представленного алгоритма вычисления МЧКК, а также, обоснование научно-технической значимости проведенного экспериментального исследования. Все научно-экспериментальные исследования системы распознавания голосовых команд управления изложенные ниже (табл. 1, 2, 3), были проведены с учетом критерия минимального расстояния, в качестве которого выступает дисперсия разности математических ожиданий сравниваемых МЧКК, в зависимости от которого изменялись значения минимальной дисперсии D_{min} , тем самым давая объективную оценку качества (достоверности) распознавания голосовых команд управления в режиме тестирования системы.

Решение о семантической идентификации голосовых команд принимается по критерию минимальной дисперсии D_{min} в заданном порог допустимого распознавания $\Theta = 1 - \Delta = 0.15$, где $\Delta = 0.85$.

Таблица 1

Результаты тестирования системы распознавания голосовых команд субъекта управления №1

Обучение	Тестирование			
Субъект управления №1	Субъект управления №1			
Голосовые команды управления	Вверх	Вниз	Вправо	Влево
Вверх	0.0311	0.1921	0.2879	0.1479
Вниз	0.5323	0.0648	0.7345	0.4284
Вправо	0.3255	0.5048	0.0123	0.1699
Влево	0.1737	0.1935	0.1648	0.0112

В первом эксперименте (см. табл. 1) были проведены сравнения МЧКК голосовых команд субъек-

та управления №1: «вверх», «вниз», «вправо», «влево», которые сохранены на уровне обучения в базу эталонных голосовых образов с МЧКК голосовых команд того же субъекта управления №1, но уже в режиме тестирования системы (МЧКК произнесенных голосовых команд в режиме тестирования сравниваются с МЧКК голосовых команд, произнесенных ранее в режиме обучения системы). Из полученных результатов (см. табл. 1) видно, что МЧКК голосовых команд субъекта управления №1 соответствуют критерию минимальной дисперсии D_{min} в заданном порог допустимого распознавания $\Theta = 0.15$: «вверх» – $D_{min} = 0.0311$, «вниз» – $D_{min} = 0.0648$, «вправо» – $D_{min} = 0.0123$, «влево» – $D_{min} = 0.0112$, исходя из этого, принимается решение о положительной семантической идентификации произнесенных голосовых команд (голосовые команды распознаются). В других случаях (см. табл. 1) наглядно видно, что значения D_{min} не соответствуют выбранному критерию, а значит, МЧКК произнесенных голосовых команд не совпадают с МЧКК, которые сохранены в базе эталонных голосовых образов, то есть, голосовые команды не распознаются.

Таблица 2

Результаты тестирования системы распознавания голосовых команд субъекта управления №2

Обучение	Тестирование			
Субъект управления №1	Субъект управления №2			
Голосовые команды управления	Вверх	Вниз	Вправо	Влево
Вверх	0.0451	0.3259	0.4090	0.1604
Вниз	0.3770	0.0482	1.1258	0.5460
Вправо	0.2055	0.4988	0.0967	0.1822
Влево	0.1268	0.3056	0.2764	0.0703

Во втором эксперименте (см. табл. 2) были проведены сравнения МЧКК произнесенных голосовых команд субъекта управления №2 в режиме тестирования с МЧКК голосовых команд субъекта управления №1 произнесенных ранее в режиме обучения системы. Из полученных результатов (см. табл. 2) можно сделать выводы, что МЧКК голосовых команд субъекта управления №2 соответствуют критерию минимальной дисперсии D_{min} в заданном порог допустимого распознавания $\Theta = 0.15$: «вверх» – $D_{min} = 0.0451$, «вниз» – $D_{min} = 0.0482$, «вправо» – $D_{min} = 0.0967$, «влево» – $D_{min} = 0.0703$, а значит, принимается решение о

положительном результате распознавания произнесенных голосовых команд. Во всех других случаях голосовые команды не распознаются, так как, результирующие значения не соответствуют заданному критерию распознавания.

Таблица 3

Результаты тестирования системы распознавания голосовых команд субъекта управления №3

Обучение	Тестирование			
Субъект управления №1	Субъект управления №3			
Голосовые команды управления	Вверх	Вниз	Вправо	Влево
Вверх	0.0602	0.1657	0.4547	0.1943
Вниз	0.4099	0.0912	1.1869	0.3772
Вправо	0.2149	0.4521	0.0846	0.1784
Влево	0.1722	0.1946	0.2922	0.0785

В третьем эксперименте (см. табл. 3) сравнивались МЧКК произнесенных голосовых команд субъекта управления №3 в режиме тестирования с МЧКК голосовых команд субъекта управления №1 произнесенных ранее в режиме обучения системы, которые сохранены в базу эталонных голосовых образов команд управления. Полученные значения результатов сравнения: «вверх» – $D_{\min} = 0.0602$, «вниз» – $D_{\min} = 0.0912$, «вправо» – $D_{\min} = 0.0846$, «влево» – $D_{\min} = 0.0785$, полностью соответствуют критерию $D_{\min} < \Theta$, где $\Theta = 0.15$, а значит, принимается решение о положительном результате распознавания произнесенных голосовых команд. Что же касается других полученных результирующих значений, то они не соответствуют заданному критерию распознавания, а значит, голосовые команды не распознаются.

Выводы

В данной статье представлена разработанная система распознавания стеганографически-преобразованных голосовых команд управления БПЛА на основе кепстрального анализа, которая позволяет обеспечить соответствие всем заданным требованиям (см. 2. Постановка проблемы) для такого рода систем.

Осуществлена сравнительная оценка результатов моделирования. В первом эксперименте (см. табл. 1) были проведены сравнения МЧКК голосовых команд субъекта управления №1: «вверх», «вниз», «вправо», «влево», которые сохранены на уровне обучения в базу эталонных голосовых обра-

зов с МЧКК голосовых команд того же субъекта управления №1, но уже в режиме тестирования системы. Из полученных результатов (см. табл. 1) видно, что МЧКК голосовых команд субъекта управления №1 соответствуют критерию минимальной дисперсии D_{\min} в заданном пороге допустимого распознавания $\Theta = 0.15$: «вверх» – $D_{\min} = 0.0311$, «вниз» – $D_{\min} = 0.0648$, «вправо» – $D_{\min} = 0.0123$, «влево» – $D_{\min} = 0.0112$, исходя из этого, принимается решение о положительной семантической идентификации произнесенных голосовых команд.

Во втором эксперименте (см. табл. 2) были проведены сравнения МЧКК произнесенных голосовых команд субъекта управления №2 в режиме тестирования с МЧКК голосовых команд субъекта управления №1 произнесенных ранее в режиме обучения системы. Из полученных результатов (см. табл. 2) можно сделать выводы, что МЧКК голосовых команд субъекта управления №2 соответствуют критерию минимальной дисперсии D_{\min} в заданном пороге допустимого распознавания $\Theta = 0.15$: «вверх» – $D_{\min} = 0.0451$, «вниз» – $D_{\min} = 0.0482$, «вправо» – $D_{\min} = 0.0967$, «влево» – $D_{\min} = 0.0703$, а значит, принимается решение о положительном результате распознавания произнесенных голосовых команд.

В третьем эксперименте (см. табл. 3) сравнивались МЧКК произнесенных голосовых команд субъекта управления №3 в режиме тестирования с МЧКК голосовых команд субъекта управления №1 произнесенных ранее в режиме обучения системы. Полученные значения результатов сравнения: «вверх» – $D_{\min} = 0.0602$, «вниз» – $D_{\min} = 0.0912$, «вправо» – $D_{\min} = 0.0846$, «влево» – $D_{\min} = 0.0785$, полностью соответствуют критерию $D_{\min} < \Theta$, где $\Theta = 0.15$, а значит, принимается решение о положительном результате распознавания произнесенных голосовых команд.

Дальнейшее исследование рассмотренной системы распознавания стеганографически-преобразованных голосовых команд управления БПЛА позволит обеспечить решения ряда важных научно-технических задач в военном и гражданском применении, а именно:

– интегрирования в современные БПЛА модуля голосового управления, что позволит повысить удобство использования БПЛА в решении различного рода военных и гражданских задач, в частности использование БПЛА в роли разведчика в группах быстрого реагирования;

– забезпечення захищеного голосового управління функціями БПЛА, що дозволить уникнути можливості перехвату управління БПЛА при розв'язанні поставлених задач.

Література

1. Cuiling, L. *English Speech Recognition Method Based on Hidden Markov Model [Text]* / L. Cuiling // *Smart Grid and Electrical Automation (ICSGEA), IEEE International Conference*. – 2016. – P. 94–97.
2. Lu, L. *Small-Footprint Highway Deep Neural Networks for Speech Recognition [Text]* / L. Lu, S. Renals // *Transactions on Audio, Speech, and Language Processing*. – 2017. – № 7. – P. 1502–1511.
3. Hazem, Z. *A study of speech recognition system based on the Hidden Markov Model with Gaussian-Mixture [Text]* / Z. Hazem, Y. Zouhir, K. Ouni // *Electrical Sciences and Technologies in Maghreb, IEEE International Conference*. – 2014. – P. 1–5.
4. Zhao, Y. *Speaker-Dependent Isolated-Word Speech Recognition System Based on Vector Quantization [Text]* / Y. Zhao, L. Zhu // *Computer Network, Electronic and Automation (ICCNEA), IEEE International Conference*. – 2017. – P. 133–137.
5. Rashmi, N. *An improved method for reversible data hiding steganography combined with cryptography [Text]* / N. Rashmi, K. Jyothi // *Inventive Systems and Control, IEEE 2nd International Conference*. – 2018. – P. 81–84.
6. Sharma, D. *Steganography of the keys into an encrypted speech signal using Matlab [Text]* / D. Sharma, De. Sharma // *Computing for Sustainable Global Development, IEEE 3rd International Conference*. – 2016. – P. 721–724.
7. Liu, J. *Least-significant-digit steganography in low bitrate speech [Text]* / J. Liu, K. Zhou, H. Tian // *Communications, IEEE International Conference*. – 2012. – P. 1133–1137.
8. Lavrynenko, O. *Method of Voice Control Functions of the UAV [Text]* / O. Lavrynenko, G. Konakhovych, D. Bakhtiiarov // *Methods and Systems of Navigation and Motion Control (MSNMC), IEEE 4th International Conference*. – 2016. – P. 47–50.
9. Lavrynenko, O. Yu. *Compression algorithm of voice control commands of UAV based on wavelet*

transform [Text] / O. Yu. Lavrynenko, G. F. Konakhovych, D. I. Bakhtiiarov // *Electronics and control systems*. – 2018. – № 1. – P. 17–22.

References

1. Cuiling, L. *English Speech Recognition Method Based on Hidden Markov Model. Smart Grid and Electrical Automation (ICSGEA), IEEE International Conference*, 2016, pp. 94–97.
2. Lu, L., Renals, S. *Small-Footprint Highway Deep Neural Networks for Speech Recognition. Transactions on Audio, Speech, and Language Processing*, 2017, vol. 25, no. 7, pp. 1502–1511.
3. Hazem, Z. *A study of speech recognition system based on the Hidden Markov Model with Gaussian-Mixture. Electrical Sciences and Technologies in Maghreb, IEEE International Conference*, 2014, pp. 1–5.
4. Zhao, Y., Zhu, L. *Speaker-Dependent Isolated-Word Speech Recognition System Based on Vector Quantization. Computer Network, Electronic and Automation (ICCNEA), IEEE International Conference*, 2017, pp. 133–137.
5. Rashmi, N., Jyothi, K. *An improved method for reversible data hiding steganography combined with cryptography. Inventive Systems and Control, IEEE 2nd International Conference*, 2018, pp. 81–84.
6. Sharma, D., Sharma, De. *Steganography of the keys into an encrypted speech signal using Matlab. Computing for Sustainable Global Development, IEEE 3rd International Conference*, 2016, pp. 721–724.
7. Liu, J., Zhou, K., Tian, H. *Least-significant-digit steganography in low bitrate speech. Communications, IEEE International Conference*, 2012, pp. 1133–1137.
8. Lavrynenko, O., Konakhovych, G., Bakhtiiarov, D. *Method of Voice Control Functions of the UAV. Methods and Systems of Navigation and Motion Control (MSNMC), IEEE 4th International Conference*, 2016, pp. 47–50.
9. Lavrynenko, O.Yu., Konakhovych, G.F., Bakhtiiarov, D.I. *Compression algorithm of voice control commands of UAV based on wavelet transform. Electronics and control systems*, 2018, vol. 55, no. 1, pp. 17–22.

Поступила в редакцію 15.07.2018, рассмотрена на редколлегии 12.09.2018

СИСТЕМА РОЗПІЗНАВАННЯ СТЕГANOГРАФІЧНО-ПЕРЕТВОРЕНИХ ГОЛОСОВИХ КОМАНД УПРАВЛІННЯ БПЛА

О. Ю. Лавриненко, Ю. А. Кочергін, Г. Ф. Коначович

Розроблено систему розпізнавання стеганографічно-перетворених голосових команд управління безпілотним літальним апаратом на основі кепстрального аналізу, яка дозволяє забезпечити ефективне розпізнавання та потайливу передачу команд на борт безпілотного літального апарату, за допомогою перетворення голосових команд управління в свого роду стеганографічний вектор ознак, що має на увазі приховування інформації голосового управління безпілотним літальним апаратом. Синтезована математична модель алгоритму обчислення мел-частотних кепстральних коефіцієнтів і класифікатора розпізнавання голосових команд управління для вирішення завдання семантичної ідентифікації та забезпечення скритності інформації

управління безпілотного літального апарату в каналі зв'язку. Розроблено програмний комплекс, що включає засоби для складання бази еталонних голосових образів суб'єктів управління для навчання і тестування системи розпізнавання стеганографічно-перетворених голосових команд управління безпілотним літальним апаратом на основі кепстрального аналізу і комп'ютерні моделі запропонованих методів і алгоритмів розпізнавання голосових команд управління в середовищі MATLAB. Обгрунтовано та експериментально доведено доцільність використання запропонованої системи розпізнавання стеганографічно-перетворених голосових команд управління безпілотним літальним апаратом на основі кепстрального аналізу. Представлений алгоритм обчислення мел-частотних кепстральних коефіцієнтів, які виступають в ролі основних ознак розпізнавання і результату стеганографічного перетворення мови, де для оцінювання результатів автоматичного розпізнавання голосових команд управління використовується класифікатор, побудований за критерієм мінімальної відстані в ролі якого виступає дисперсія різниці математичного очікування мел-частотних кепстральних коефіцієнтів. Отримані результати експериментального дослідження дозволяють зробити висновки про доцільність подальшого практичного застосування розробленої системи розпізнавання стеганографічно-перетворених голосових команд управління безпілотним літальним апаратом на основі кепстрального аналізу.

Ключові слова: стеганографічне перетворення мови, голосові команди управління, розпізнавання голосових команд, кепстральний аналіз, кепстральні коефіцієнти, перетворення Фур'є, косинусне перетворення, система голосового управління, безпілотний літальний апарат.

SYSTEM OF RECOGNITION THE STEGANOGRAPHIC-TRANSFORMED VOICE COMMANDS OF THE UAV CONTROL

O. Lavrynenko, Y. Kocherhin, G. Konakhovych

It is created the system of recognition the steganographic-transformed voice commands of unmanned aerial vehicle control based on a cepstral analysis. It provides effective recognition and hidden commands transmission of the board of an unmanned aerial vehicle, by converting voice control commands into a kind of steganographic characteristics vector, which implies the concealment of voice control information of an unmanned aerial vehicle. The mathematical model of the algorithm for calculating the mel-frequency cepstral coefficients and the recognition classifier of voice control commands for the solution of the problem of semantic identification and securing the control information of the unmanned aerial vehicle in the communication channel is synthesized. A software package has been developed that includes tools for compiling the base of reference voice images of subjects of management for training and testing the system for recognizing steganographic-transformed voice commands of the unmanned aerial vehicle control based on the cepstral analysis and computer models of the proposed methods and algorithms for recognition voice control commands in the MATLAB environment. The expediency of applying the proposed system for recognizing steganographic-transformed voice commands of the unmanned aerial vehicle control based on a cepstral analysis is substantiated and experimentally proved. An algorithm is presented for calculating the mel-frequency cepstral coefficients that appear in the role of the main features of recognition and the result of the steganographic transformation of speech, where for the evaluation of automatic recognition of voice commands using the results of classifier constructed by the criterion of minimum distance in the role which acts as the variance of the difference of the expectation of a mel-frequency cepstral coefficients. The obtained results of the experimental research allow to draw a conclusion about the expediency of further practical application of the developed system of recognition the steganographic-transformed voice commands of the unmanned aerial vehicle control based on the cepstral analysis.

Keywords: steganographic speech transformation, voice command control, voice command recognition, cepstral analysis, cepstral coefficients, Fourier transform, cosine transform, voice control system, unmanned aerial vehicle.

Лавриненко Александр Юрьевич – аспирант, ассистент кафедры телекоммуникационных систем, Национальный авиационный университет, Киев, Украина, e-mail: oleksandrlavrynenko@gmail.com.

Кочергин Юрий Анатольевич – канд. техн. наук, доцент, главный инженер, Национальный авиационный университет, Киев, Украина, e-mail: tko@nau.edu.ua.

Конахович Георгий Филимонович – д-р техн. наук, проф., зав. каф. телекоммуникационных систем, Национальный авиационный университет, Киев, Украина, e-mail: tko@nau.edu.ua.

Oleksandr Lavrynenko – PhD student, Assistant to Department of Telecommunication Systems, National Aviation University, Kyiv, Ukraine, e-mail: oleksandrlavrynenko@gmail.com.

Yurii Kocherhin – PhD, Associate Professor, National Aviation University, Kyiv, Ukraine, e-mail: tko@nau.edu.ua.

Georgy Konakhovych – Doctor of Engineering, Professor, Head of Department of Telecommunication Systems, National Aviation University, Kyiv, Ukraine, e-mail: tko@nau.edu.ua.