

ФОРМИРОВАНИЕ СИСТЕМЫ ИНФОРМАТИВНЫХ ПРИЗНАКОВ В ЗАДАЧАХ ИДЕНТИФИКАЦИИ ДИКТОРОВ

Федоров А.В., Омельченко А.В., Пресняков И.Н.

Харьковский национальный университет радиоэлектроники
61166, Харьков, пр. Ленина, каф. сетей связи, тел. (057) 702-14-29

E-mail: stativka@kture.kharkov.ua

In the present work the problems of shaping of a system of informative indications for identification of the speakers on singularities of their voice are considered. The primary map of a voice call represents a standard digital stream with a transfer rate 64 Kbps. The methods of a linear prediction of speech estimate parameters of the model of a speech tract as an audio pipe. The obtained estimations of parameters of model use as informative indications in the task of identification. The problems of correlation and statistical independence of selected spaces of informative indications researched. The program security on research of performances of voice calls and shaping of systems of informative indications is created.

Выбор параметров речевого сигнала, способных наилучшим образом описать индивидуальность голоса, является, пожалуй, самым важным этапом при построении автоматических систем опознавания голосов. Такие параметры сигнала, называемые признаками индивидуальности, помимо эффективности представления информации об особенностях голоса и речи диктора, должны обладать рядом других свойств. Во-первых, они должны быть легко измеряемы и малоависимы от мешающих факторов окружающей среды (шумов и помех), во-вторых, они должны обладать стабильностью во времени и, в-третьих, не должны поддаваться имитации. Создание системы признаков, удовлетворяющей всем этим требованиям, является трудно решаемой задачей. Поэтому, в первую очередь, производят отбор признаков по их способности адекватнее и эффективнее произвести описание индивидуального голоса, уделяя меньше внимания другим факторам.

Исследование речевого сигнала показывает, что признаки индивидуальности, как правило, составляют большой массив данных, эффективность которых не может быть заранее оценена, до тех пор, пока не будет выяснена избыточность этого массива. Кроме того, пространство исходного описания, имеющее высокую размерность, существенно осложняет формирование решающего правила и создает большие трудности вычислительного характера. Отсюда возникает задача сокращения исходного числа параметров изучаемого объекта до такого числа, которое обеспечит получение необходимого результата. Отобранные таким образом параметры будем называть информативными признаками. Создание и отбор признаков из исходного пространства может происходить эвристически или на основе соответствующих определенным закономерностям алгоритмов. Эвристически найденные признаки могут оказаться сильно коррелированными между собой и повторять информацию. Для создания некоррелированной системы признаков обычно прибегают к специальным алгоритмам, оптимизирующим некоторый функционал, характеризующий разделяющие свойства системы признаков. При этом может максимизироваться надежность идентификации при заданном числе признаков, или минимизироваться число признаков при сохранении точности идентификации дикторов. Одним из таких преобразований может быть декоррелирующее преобразование Карунена-Лоэва [1].

Эффективность отдельного признака или системы признаков может быть оценена также через энтропию [2], которая характеризует количество информации, привносимой в результате использования данного признака или системы признаков.

Рассмотрим теперь непосредственно процесс нахождения требуемой подсистемы. Очевидно, оптимальное решение этой задачи может быть получено путем полного перебора всевозможных вариантов, что сопряжено с большими вычислительными трудностями. Их можно преодолеть, применяя приближенные методы решения задачи.

Например, используя результаты теории планирования эксперимента, рассматривая процесс идентификации как пассивный эксперимент [3].

Согласно [4] основные физические проявления индивидуальности следует искать в спектральных или формантных, а также временных и амплитудных характеристиках сигнала. Такие системы признаков не используют данных о строении речеобразующего тракта говорящего, которые содержатся в предварительно обработанном речевом сигнале. Методами линейного предсказания речи [5] удастся оценить параметры модели речеобразующего тракта в виде акустической трубы, которая простирается от голосовой щели до губ и изменяется по форме с течением времени. Полученные оценки параметров модели речевого тракта могут быть использованы в качестве признаков индивидуальности в задачах идентификации дикторов. Кроме того, важная информация об индивидуальности абонентов сосредоточена в динамических характеристиках частоты основного тона и сигналах возбуждения голосового тракта, их корреляционных характеристиках.

Настоящая работа посвящена вопросам формирования систем информативных признаков включающих в себя: 1) коэффициенты авторегрессии, 2) коэффициенты отражения звуковой волны от стыка секций акустической трубы, 3) оценки параметров голосового тракта в виде акустической трубы.

Для выяснения статистической зависимости исследуемых информативных признаков, для коэффициентов авторегрессии и отражения, а также для оценок площадей поперечного сечения акустических труб дикторов были построены нормированные ковариационные матрицы до и после декоррелирующего преобразования Карунена-Лозва. Сравнительный анализ полученных данных показывает, что информативные признаки сильно коррелированы, что необходимо учитывать при построении решающих правил идентификации дикторов.

На рис.1 и рис. 2 показаны координаты собственных векторов ковариационных матриц коэффициентов отражения и площадей поперечных сечений акустической трубы одного диктора.

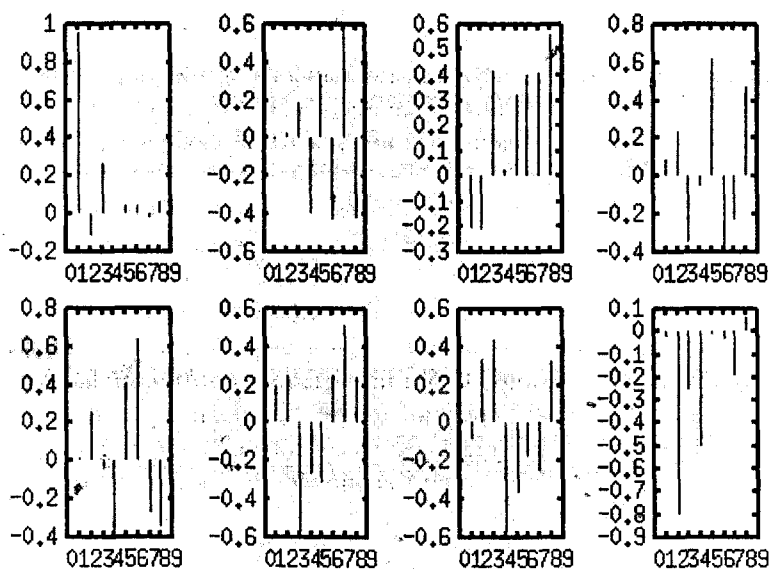


Рис. 1. Координаты собственных векторов ковариационных матриц коэффициентов отражения звуковой волны от стыка секций акустической трубы

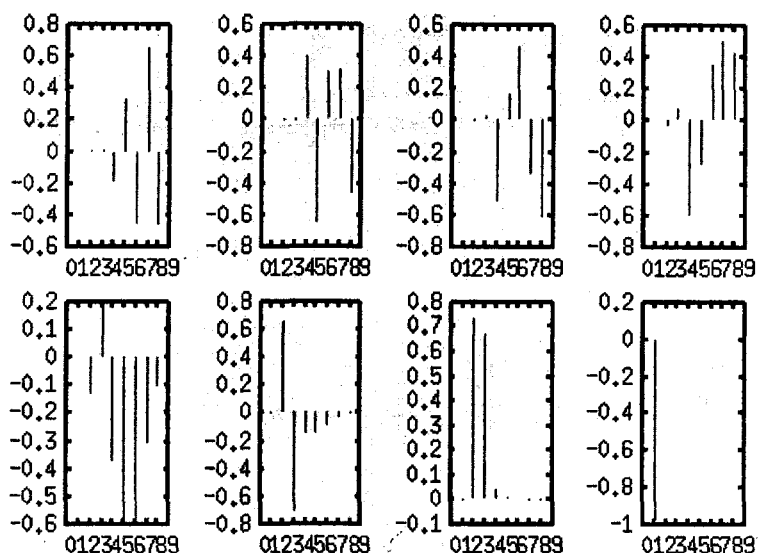


Рис. 2. Координаты собственных векторов ковариационных матриц площадей поперечных сечений акустической трубы

Здесь вертикальными линиями показаны 8-мь координат 8-ми собственных векторов расположенных в порядке возрастания соответствующих им собственных чисел. Номера векторов отложены по оси абсцисс, а значения – по оси ординат. Из анализа приведенных графических зависимостей видно, что у площадей сечений коррелированы только соседние значения. Это дает возможность упростить алгоритмы идентификации, т.к. позволяет учитывать меньшее число членов ковариационной матрицы, что особенно важно при обучении по выборкам малого объема.

Для сравнения рассматриваемых систем информативных признаков по критерию информационной эффективности для четырех различных дикторов по выборкам длительностью порядка 40с вычислялись оценки дифференциальной энтропии. На рис.3 показаны гистограммы декоррелированных значений первых четырех площадей поперечных сечений акустических труб диктора. По оси ординат отложено число значений, попавших в соответствующий интервал гистограммы.

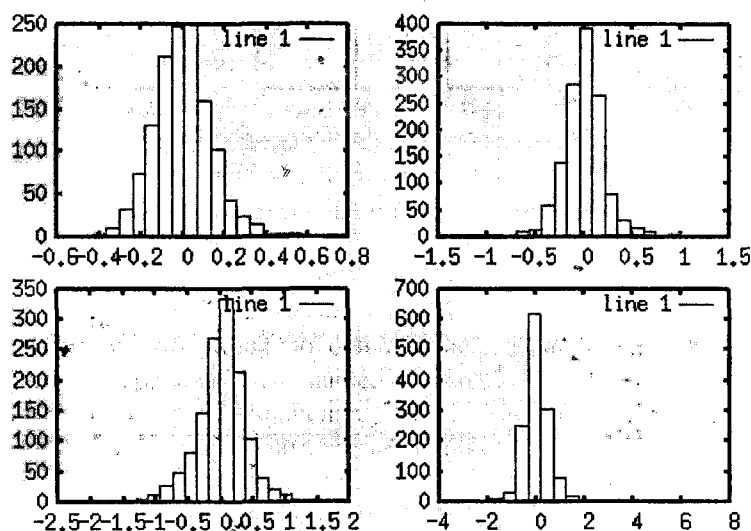


Рис. 3. Гистограммы первых четырех площадей поперечных сечений акустической трубы диктора

Используя критерий Колмогорова-Смирнова, была подтверждена гипотеза о нормальности распределения декоррелированных информативных признаков. Значения дифференциальной энтропии определялись по декоррелированным признакам в предположении нормального распределения по формуле [2]:

$$H = \sum_{i=1}^n H_i, \quad H_i = \log_2 \sqrt{2 \cdot \pi \cdot e \cdot \sigma_i^2},$$

где n - порядок модели линейного предсказания, σ_i^2 - дисперсия i -го подпризнака пространства признаков. Здесь $n = 8$.

Результаты вычислений сведены в табл. 1.

Таблица 1 – Оценки энтропии для четырех различных дикторов, дв.ед.

№ диктора/признак	Коэффициенты авторегрессии	Коэффициенты отражения	Площади поперечных сечений
1	-45.950	-44.665	-24.733
2	-48.247	-46.959	-14.275
3	-53.473	-51.562	-21.856
4	-53.174	-50.664	-30.559

Таким образом, с точки зрения информационной эффективности наилучшими следует считать системы информативных признаков состоящие из площадей поперечных сечений акустических труб дикторов.

В работе выполнен анализ различных систем информативных признаков включающих в себя: 1) коэффициенты авторегрессии, 2) коэффициенты отражения звуковой волны от стыка секций акустической трубы, 3) оценки параметров голосового тракта в виде акустической трубы с точки зрения их статистической независимости и информационной эффективности. Было показано что рассматриваемые системы информативных признаков в значительной степени коррелированы и для построения эффективных систем идентификации дикторов исходные признаки необходимо декоррелировать. Использование в качестве информативных признаков площадей поперечных сечений акустических труб дикторов дает возможность упростить алгоритмы идентификации, т.к. позволяет учитывать меньшее число членов ковариационной матрицы. Кроме того, выполненные экспериментальные исследования по идентификации дикторов показали, что наилучшие результаты достигаются при использовании коэффициентов отражения и площадей поперечных сечений акустических труб дикторов.

Литература

1. Пугачев В.С. Теория вероятностей и математическая статистика. – М.: Наука, 1979.
2. Шеннон К. Работы по теории информации и кибернетике. – М.: Издательство иностранной литературы, 1963.
3. Асатурян В.И. Теория планирования эксперимента. – М.: Радио и связь, 1983.
4. Рамишвили Г.С. Автоматическое распознавание говорящего по голосу. – М.: Радио и связь, 1981.
5. Маркел Дж. Д., Грей А.Х. Линейное предсказание речи: – М.: Связь 1980.