

ОПИСАНИЕ ПОЛЕЗНОЙ МОДЕЛИ К ПАТЕНТУ

(12)

РЕСПУБЛИКА БЕЛАРУСЬ



НАЦИОНАЛЬНЫЙ ЦЕНТР
ИНТЕЛЛЕКТУАЛЬНОЙ
СОБСТВЕННОСТИ

(19) ВУ (11) 8194

(13) U

(46) 2012.04.30

(51) МПК

G 10L 15/00 (2006.01)

(54)

УСТРОЙСТВО АНАЛИЗА РЕЧИ

(21) Номер заявки: u 20110835

(22) 2011.10.27

(71) Заявитель: Учреждение образования "Белорусский государственный университет информатики и радиоэлектроники" (ВУ)

(72) Авторы: Зельманский Олег Борисович; Давыдов Андрей Геннадьевич (ВУ)

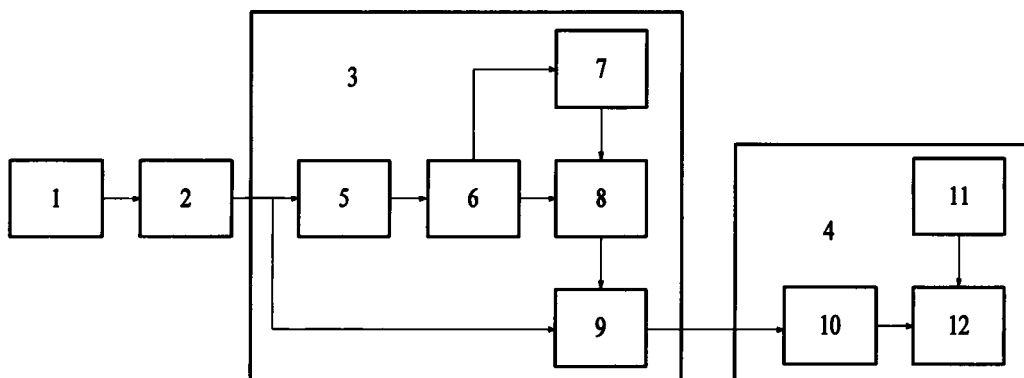
(73) Патентообладатель: Учреждение образования "Белорусский государственный университет информатики и радиоэлектроники" (ВУ)

(57)

1. Устройство анализа речи, содержащее блок поступления голосового сигнала, блок частотного преобразования, **отличающееся** тем, что в него введены блоки сегментации и классификации речи, соединенные последовательно.

2. Устройство по п. 1, **отличающееся** тем, что блок сегментации речи состоит из блока вычисления линейных спектральных частот, блока расчета расстояния между фреймами, блока формирования порогового значения, блока расстановки границ, блока формирования фонетических единиц, соединенных последовательно, при этом вход блока вычисления линейных спектральных частот является входом блока сегментации речи и соединен с входом блока формирования фонетических единиц, выход которого является выходом блока сегментации речи, второй выход блока расчета расстояния между фреймами соединен со вторым входом блока расстановки границ.

3. Устройство по п. 1, **отличающееся** тем, что блок классификации речи состоит из блока вычисления мел-частотных кепстральных коэффициентов, вход которого является входом блока классификации речи, блока вычисления степени соответствия между фонетическими единицами, соединенных последовательно, базы эталонов мел-частотных кепстральных коэффициентов фонетических единиц, соединенной со вторым входом блока вычисления степени соответствия между фонетическими единицами, выход которого является выходом блока классификации речи.



ВУ 8194 U 2012.04.30

(56)

1. Патент США 5907826, МПК G 10L 3/02, 1999.
2. Патент РФ 2403626 С2, МПК G 10L 11/04, 2010.

Полезная модель относится к устройствам анализа речи и может быть использована для сегментации речевого сигнала на фонетические единицы и последующей их классификации.

Известно устройство сегментации речи [1], содержащее блок выделения информативных параметров и блок распознавания, соединенные последовательно. Блок выделения информативных параметров состоит из секции выделения спектральных компонентов, секции выделения частоты основного тона и секции вычисления производной по времени интенсивности основного тона, включенных параллельно.

Недостатком такого устройства является низкая точность сегментации гласных звуков, следующих один за другим.

Наиболее близкими к предлагаемой полезной модели являются анализатор речи, обнаруживающий частоту основного тона, способ анализа речи и программа анализа речи [2]. Известное устройство содержит блок поступления голосового сигнала, предназначенный для получения голосового сигнала от проверяемого, блок частотного преобразования, предназначенный для преобразования упомянутого голосового сигнала в частотный спектр, автокорреляционный блок, предназначенный для вычисления автокорреляционного колебания при сдвиге упомянутого частотного спектра на частотной оси, и блок обнаружения основного тона, предназначенный для вычисления частоты на основе локального интервала между одними гребнями и впадинами упомянутого автокорреляционного колебания.

Недостатком данного устройства является отсутствие возможности сегментации речи на фонетические единицы и классификации этих фонетических единиц.

Задачей предлагаемого технического решения является создание устройства, выполняющего сегментацию речевого сигнала на квазистационарные участки, соответствующие голосовым и шумовым фонемам, и их классификацию.

Вышеуказанная задача решается тем, что в устройство анализа речи, содержащее блок поступления голосового сигнала, блок частотного преобразования, введены блоки сегментации и классификации речи, соединенные последовательно. При этом блок сегментации речи состоит из блока вычисления линейных спектральных частот, блока расчета расстояния между фреймами, блока формирования порогового значения, блока расстановки границ, блока формирования фонетических единиц, соединенных последовательно, при этом вход блока вычисления линейных спектральных частот является входом блока сегментации речи и соединен с входом блока формирования фонетических единиц, выход которого является выходом блока сегментации речи, второй выход блока расчета расстояния между фреймами соединен со вторым входом блока расстановки границ. В свою очередь, блок классификации речи состоит из блока вычисления мел-частотных кепстральных коэффициентов, вход которого является входом блока классификации речи, блока вычисления степени соответствия между фонетическими единицами, соединенных последовательно, базы эталонов мел-частотных кепстральных коэффициентов фонетических единиц, соединенной со вторым входом блока вычисления степени соответствия между фонетическими единицами, выход которого является выходом блока классификации речи.

На фигуре изображена структурная схема устройства анализа речи.

В соответствии с фиг. 1 устройство анализа речи содержит: блок поступления голосового сигнала 1, блок частотного преобразования 2, блок сегментации речи 3 и блок классификации речи 4, соединенные последовательно. Блок сегментации речи 3 содержит:

BY 8194 U 2012.04.30

блок вычисления линейных спектральных частот 5, блок расчета расстояния между фреймами 6, блок формирования порогового значения 7, блок расстановки границ 8, блок формирования фонетических единиц 9, соединенные последовательно. Выход блока частотного преобразования 2 подключен к входу блока вычисления линейных спектральных частот 5, соединенному с входом блока формирования фонетических единиц 9. Второй выход блока расчета расстояния между фреймами 6 соединен со вторым входом блока расстановки границ 8. Блок классификации речи 4 содержит: блок вычисления мел-частотных кепстральных коэффициентов 10, выход которого подключен к первому входу блока вычисления степени соответствия между фонетическими единицами 12, базу эталонов мел-частотных кепстральных коэффициентов фонетических единиц 11, выход которой подключен ко второму входу блока вычисления степени соответствия между фонетическими единицами 12. Вход блока вычисления мел-частотных кепстральных коэффициентов 10 подключен к выходу блока формирования фонетических единиц 9.

Работа устройства происходит следующим образом. Блок частотного преобразования 2 преобразует фреймы (окна) голосового сигнала, полученные в блоке поступления голосового сигнала 1, в частотный спектр, на основе которого в блоке вычисления линейных спектральных частот 5 рассчитывается вектор линейных спектральных корней каждого фрейма. Блок расчета расстояния между фреймами 6 вычисляет расстояние между векторами линейных спектральных корней соседних фреймов. В качестве расстояния используется метрика Минковского четвертого порядка. Полученное значение расстояния сравнивается с пороговым значением в блоке расстановки границ 8, и в случае, если оно превышает пороговое, принимается решение о постановке границы фонетической единицы. В блоке формирования порогового значения 7 на основе получаемых от блока расчета расстояния между фреймами 6 значений расстояний между последовательными фреймами, а также коэффициента отношения величины порогового значения расстояния к среднему значению расстояния между фреймами формируется пороговое значение расстояния, которое поступает в блок расстановки границ 8. На первый вход блока формирования фонетических единиц 9 от блока расстановки границ 8 поступает информация о местоположении границ фонетических единиц, на его второй вход поступает частотный спектр сигнала, на выходе формируются частотные спектры сегментированных фонетических единиц, которые последовательно передаются на обработку в блок вычисления мел-частотных кепстральных коэффициентов 10. Вычисленный вектор мел-частотных кепстральных коэффициентов сегментированной фонетической единицы подается на первый вход блока вычисления степени соответствия между фонетическими единицами 12, на второй вход которого из базы эталонов мел-частотных кепстральных коэффициентов фонетических единиц 11 поступают эталонные вектора мел-частотных кепстральных коэффициентов классифицированных фонетических единиц. С помощью метода динамического программирования блок вычисления степени соответствия между фонетическими единицами 12 определяет, к какому из классов эталонных фонетических единиц наиболее близка анализируемая фонетическая единица, и принимает решение о ее принадлежности данному классу.

Применение предлагаемого технического решения позволяет сегментировать речевые сигналы на участки квазистационарных колебаний, соответствующие голосовым и шумовым фонемам, и классифицировать их.