

ОПИСАНИЕ ПОЛЕЗНОЙ МОДЕЛИ К ПАТЕНТУ

(12)

РЕСПУБЛИКА БЕЛАРУСЬ



НАЦИОНАЛЬНЫЙ ЦЕНТР
ИНТЕЛЛЕКТУАЛЬНОЙ
СОБСТВЕННОСТИ

(19) ВУ (11) 4668

(13) U

(46) 2008.10.30

(51) МПК (2006)

G 10L 15/00

(54)

УСТРОЙСТВО СЕГМЕНТАЦИИ РЕЧИ

(21) Номер заявки: u 20080053

(22) 2008.01.29

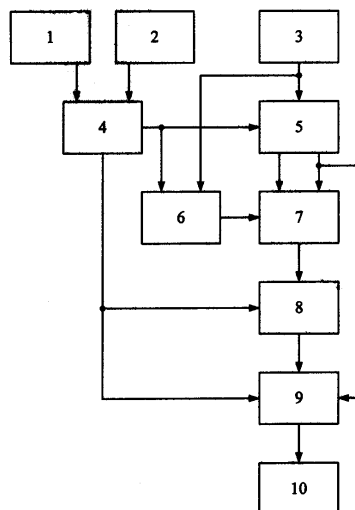
(71) Заявители: Учреждение образования "Белорусский государственный университет информатики и радиоэлектроники"; Государственное научное учреждение "Объединенный институт проблем информатики Национальной академии наук Беларуси" (ВУ)

(72) Авторы: Воробьев Василий Иванович; Давыдов Андрей Геннадьевич; Давыдов Геннадий Владимирович; Лобанов Борис Мефодьевич; Лыньков Леонид Михайлович; Шамгин Юрий Васильевич (ВУ)

(73) Патентообладатели: Учреждение образования "Белорусский государственный университет информатики и радиоэлектроники"; Государственное научное учреждение "Объединенный институт проблем информатики Национальной академии наук Беларуси" (ВУ)

(57)

Устройство сегментации речи, содержащее модуль речевой базы, модуль ввода речи и модуль ввода текста, выход которого подключен к первому входу модуля синтезатора речи по тексту, а выход модуля речевой базы подключен ко второму входу модуля синтезатора речи по тексту, первый выход которого подключен к первому входу блока выделения информативных параметров речи, а второй вход которого подключен к выходу модуля ввода речи, первый выход блока выделения информативных параметров речи подключен к первому входу блока сопоставления характеристических векторов, а второй выход блока выделения информативных параметров речи подключен ко второму входу блока сопоставления характеристических векторов и первому входу блока уточнения меток, второй



ВУ 4668 U 2008.10.30

BY 4668 U 2008.10.30

выход модуля синтезатора речи по тексту подключен к первому входу блока переноса меток и второму входу блока уточнения меток, выход блока сопоставления характеристических векторов подключен ко второму входу блока переноса меток, выход которого подключен к третьему входу блока уточнения меток, а выход блока уточнения меток подключен ко входу модуля вывода разметки входной речи, **отличающееся** тем, что дополнительно содержит блок задания управляющих коэффициентов, выход которого подключен к третьему входу блока сопоставления характеристических векторов, а первый и второй входы соединены с первым выходом модуля синтеза речи по тексту и выходом модуля ввода речи соответственно.

(56)

1. Патент США 5907826, МПК G10L 3/02, опубл. 1999.
2. Патент США 6424946, МПК G10L 15/22, опубл. 2002.
3. Патент США 7010481, МПК G10L 19/04, G10L 19/14, G10L 15/14, опубл. 2006.

Полезная модель относится к устройствам сегментации речи и может быть использована при верификации голоса диктора и обучении системы распознавания речи.

Известно устройство сегментации речи [1], содержащее блок выделения информативных параметров и блок распознавания, соединенных 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.

Известны метод и устройство параллельного распознавания и сегментации речи [2], содержащее модуль ввода речи, модуль речевой базы, блок выделения информативных параметров речи, блок сегментации и блок идентификации диктора и кластеризации. Блок сегментации включает модуль сопоставления характеристических векторов и модуль расстановки меток, включенных последовательно. Недостатком такого устройства является невысокая скорость процесса сегментации для заданной точности.

Наиболее близкими к предлагаемому устройству (прототип) являются метод и устройство сегментации речи [3]. Известное устройство содержит модуль речевой базы, модуль ввода текста, подключенные к модулю синтезатора речи по тексту, один из выходов которого подключен на первый вход блока выделения информативных параметров речи, а на второй вход блока выделения информативных параметров речи подключен модуль ввода речи, при этом первый и второй выходы блока выделения информативных параметров речи подключены на первый и второй входы блока сопоставления характеристических векторов, и один из выходов одновременно подключен на блок сегментации, имеющий в своем составе блок переноса меток и блок уточнения меток, выход блока сопоставления характеристических векторов подключен на второй вход блока сегментации, второй выход модуля синтезатора речи по тексту подключен на третий вход блока сопоставления характеристических векторов и третий вход блока сегментации, выход которого подключен на модуль разметки входной речи.

Недостатком такого устройства является то, что оно не обеспечивает высокой надежности правильной сегментации речи.

Задачей данной полезной модели является создание устройства сегментации речи с высокой точностью сегментации фонограмм с известным началом и концом, а также фонограмм с неизвестными началом и концом.

Задача решается следующим образом. В устройство сегментации речи, содержащее модуль речевой базы, модуль ввода речи и модуль ввода текста, выход которого подклю-

чен к первому входу модуля синтезатора речи по тексту, а выход модуля речевой базы подключен ко второму входу модуля синтезатора речи по тексту, первый выход которого подключен к первому входу блока выделения информативных параметров речи, а второй вход которого подключен к выходу модуля ввода речи, первый выход блока выделения информативных параметров речи подключен к первому входу блока сопоставления характеристических векторов, а второй выход блока выделения информативных параметров речи подключен ко второму входу блока сопоставления характеристических векторов и первому входу блока уточнения меток, второй выход модуля синтезатора речи по тексту подключен к первому входу блока переноса меток и второму входу блока уточнения меток, выход блока сопоставления характеристических векторов подключен ко второму входу блока переноса меток, выход которого подключен к третьему входу блока уточнения меток, а выход блока уточнения меток подключен ко входу модуля вывода разметки входной речи, дополнительно введен блок задания управляющих коэффициентов, выход которого подключен к третьему входу блока сопоставления характеристических векторов, а первый и второй входы соединены с первым выходом модуля синтеза речи по тексту и выходом модуля ввода речи соответственно.

На фигуре представлено схематехническое изображение устройства сегментации речи.

Устройство сегментации речи содержит: модуль речевой базы 1, модуль ввода текста 2, модуль ввода речи 3, модуль синтезатора речи по тексту 4, блок выделения информативных параметров речи 5, блок задания управляющих коэффициентов 6, блок сопоставления характеристических векторов 7, блок переноса меток 8, блок уточнения меток 9, модуль вывода разметки входной речи 10.

Модуль речевой базы 1 и модуль ввода текста 2 подключены соответственно ко второму и первому входам модуля синтезатора речи по тексту 4, первый выход которого подключен к первому входу блока выделения информативных параметров речи 5 и первый вход блока задания управляющих коэффициентов 6, а второй выход модуля синтезатора речи по тексту 4 подключен к первому входу блока переноса меток 8 и второму входу блока уточнения меток 9. Выход модуля ввода речи 3 подключен одновременно ко второму входу блока выделения информативных параметров речи 5 и второму входу блока задания управляющих коэффициентов 6, выход которого подключен к третьему входу блока сопоставления характеристических векторов 7. Первый выход блока выделения информативных параметров речи 5 подключен к первому входу блока сопоставления характеристических векторов 7, а второй выход блока выделения информативных параметров речи 5 подключен параллельно ко второму входу блока сопоставления характеристических векторов 7 и первому входу блока уточнения меток 9. Выход блока сопоставления характеристических векторов 7 подключен ко второму входу блока переноса меток 8, выход которого подключен к третьему входу блока уточнения меток 9, а выход блока уточнения меток 9 подключен ко входу модуля вывода разметки входной речи 10, выход которого является выходом устройства в целом.

Работа устройства сегментации речи происходит следующим образом. С модуля ввода речи 3 речевой сигнал, преобразованный им в цифровую форму, поступает одновременно на второй вход блока выделения информативных параметров 5 и второй вход блока задания управляющих коэффициентов 6. С модуля ввода текста 2 закодированный сигнал текста, соответствующий произнесенному речевому оцифрованному сигналу в модуле ввода речи 3, поступает на первый вход модуля синтезатора речи по тексту 4. На второй вход модуля синтезатора речи по тексту 4 поступает речевой сигнал в цифровой форме соответствующих структурных элементов речи с модуля речевой базы 1. Модуль речевой базы 1 хранит эталонные структурные элементы речи, на которые будет сегментироваться в дальнейшем входной речевой сигнал. Структурными эталонными элементами речи, хранящимися в модуле речевой базы 1, могут быть аллофоны, дифоны, трифоны, полифоны

BY 4668 U 2008.10.30

или другие структурные единицы речи, на которые будет проводиться сегментация входной речи.

Модуль синтезатора речи по тексту 4 синтезирует речевую последовательность из структурных элементов речи, хранящихся в модуле речевой базы 1, в соответствии с закодированным сигналом текста, поступающим с модуля ввода текста 2, и передает ее с первого выхода на первый вход блока выделения информативных параметров речи 5 и первый вход блока задания управляющих коэффициентов 6. Со второго выхода модуля синтезатора речи по тексту 4 сигнал о положении меток, определяющих границы структурных элементов синтезируемой речи, поступает на первый вход блока переноса меток 8 и второй вход блока уточнения (положения) меток 9.

Блок выделения информативных параметров речи 5 выделяет по одному и тому же алгоритму информативные параметры речи для речевого сигнала, поступившего с модуля ввода речи 3, и для синтезированного речевого сигнала, поступившего с модуля синтезатора речи по тексту 4. Речевой сигнал и его информативные параметры для речи, поступившей с модуля ввода речи 3, передаются блоком выделения информативных параметров речи 5 со второго выхода на второй вход блока сопоставления характеристических векторов 7 и на первый вход блока уточнения меток 9. А синтезированный речевой сигнал и его информативные параметры передаются с первого выхода блока выделения информативных параметров речи 5 на первый вход блока сопоставления характеристических векторов 7. Блок выделения информативных параметров речи 5 выделяет информативные параметры, такие как спектр речевого сигнала, скорость изменения спектра, мощность речевого сигнала на его участках (фреймах), кепстр речевого сигнала, меру периодичности речевого сигнала и ее конечную разность.

Блок задания управляющих коэффициентов 6 на основании анализа сегментируемых и синтезированных речевых сигналов вычисляет управляющие коэффициенты и передает их на третий вход блока сопоставления характеристических векторов 7. Вычисление управляющих коэффициентов выполняется на основе анализа длительностей сопоставляемых сигналов.

Блок сопоставления характеристических векторов 7 сравнивает информационные параметры сегментируемой и синтезированной речи, используя метод динамического программирования. Данные сопоставления характеристических векторов передаются на второй вход блока переноса меток 8.

Блок переноса меток 8 на основании положения меток в синтезированной речи и результата сопоставления характеристических векторов синтезированного и сегментируемого речевых сигналов определяет положения меток элементов речи в сегментируемом речевом сигнале. Данные разметки речевого сигнала передаются на третий вход блока уточнения меток 9.

Блок уточнения меток 9 на основе анализа усредненной конечной разности кратковременного спектра сегментируемого речевого сигнала выполняет проверку и уточнение положений меток сегментируемого речевого сигнала.

С выхода блока уточнения меток 9 сегментированный сигнал поступает на вход модуля вывода разметки входной речи 10, который может быть сохранен в памяти или передан исполнительному устройству, примененному в соответствии с решаемой в целом задачей. Это устройство сегментации речи может использоваться как в системе распознавания, так и верификации речи.

Введение блока задания управляющих коэффициентов в устройство сегментации речи позволило повысить точность сегментации речи в 1,5 раза.