

Предлагаемое изобретение относится к области приборостроения и может быть использовано при автоматическом распознавании речевого сигнала.

Известен способ спектрального представления вокализованного речевого сигнала, основанный на формировании совокупности отсчетов амплитудных и фазовых спектров сигналов [1].

Недостатком известного способа является большая сложность и низкая достоверность анализа невокализованных звуков.

Известен также способ обработки речевого сигнала, основанный на интегральном преобразовании речи [2].

Недостатком этого способа является низкая достоверность анализа речевого сигнала из-за применения спектроанализатора с низкой степенью разрешения по частоте.

Наиболее близким к предлагаемому техническим решением, принятым в качестве прототипа, является способ обработки речевых сигналов, включающий сегментацию речевого сигнала с последующим функциональным преобразованием сегментов: и спектральным анализом полученного результирующего сигнала [3].

К недостаткам этого способа относится низкая достоверность анализа речевого сигнала. Это объясняется тем, что в указанном изобретении применение только одного спектрального базиса для анализа сегментов вызывает большие погрешности в представлении этого речевого сегмента из-за отсутствия возможности определения оптимальной системы базисных функций. Другими словами, в данном способе обработки речевого сигнала предполагается использование только одной экспоненциальной системы базисных функций, с помощью которой представляются абсолютно все сегменты речевого сигнала. Поскольку тонкая структура этих сегментов существенно отличается друг от друга, то можно утверждать, что спектральное представление значительно различающихся по своим структурным свойствам речевых сегментов будет иметь различную погрешность представления в выбранном в прототипе гармоническом спектральном базисе. По этой причине невозможно осуществить точную сегментацию речевого сигнала на основании полученного его спектрального изображения до тех пор, пока не будет достигнута минимальная погрешность спектрального представления этого речевого сигнала. Кроме того, прототип лишен возможности на этапе анализа речевых сегментов определять величину погрешности их представления в выбранном спектральном базисе, что на этапе распознавания и принятия решений отрицательно сказывается на достоверности правильного распознавания.

В основу изобретения поставлена задача создать способ обработки речевого сигнала, который путем определения оптимальных систем базисных функций для представления каждого сегмента речевого сигнала с последующим вычислением погрешности аппроксимации этих сегментов спектральными функциями выбранного базиса, позволил бы обеспечить получение более достоверного спектрального описания распознаваемого речевого сигнала,

Поставленная задача решается тем, что в способе, основанном на сегментации речевых сигналов с последующим функциональным преобразованием сегментов и спектральным анализом полученного результата, дополнительно осуществляют поиск оптимальной системы базисных функций, синтез речевого сигнала с учетом вычисленных оптимальных спектральных коэффициентов, задержку исходного речевого сигнала на время разложения и последующего синтеза этого сигнала по оптимальному для него системе базисных функций, сравнивают задержанный и синтезированный речевые сигналы и, по результатам сравнения, судят о величине погрешности спектрального представления обрабатываемого речевого сигнала.

На чертеже приведена структурная схема устройства для осуществления предлагаемого способа.

Сущность предлагаемого способа заключается в следующем. Согласно [4], тонкая структура различных фонем речевого сигнала имеет существенные отличия друг от друга. Существующее многообразие систем спектральных функций (например, система функций Фурье, Уолша, Хаара, Шаудера, Виленкина-Крестенсона и т. д., система коэффициентов Лаггера, Чебышева, Эрмита и т. д.) позволяет подобрать для каждого сегмента речевого сигнала такой функциональный базис, разложение по которому оптимальным образом учитывало бы структурные особенности обрабатываемой фонемы. Следовательно, признаки речевого сегмента, полученные в этом функциональном базисе, будут более точно отождествлять описываемый речевой сегмент, чем аналогичные признаки любого другого функционального базиса, в котором данный сегмент описывается с большей погрешностью. Использование данных о величине погрешности спектрального описания речевых сегментов дает возможность на этапе их распознавания применять принцип "разумной достаточности" при выборе процедур обработки и сравнения распознаваемого речевого сигнала с эталонными. В результате чего уровень сложности процедуры обработки и оценки искомого речевого сигнала будет соответствовать степени точности спектрального описания этого сигнала. С этой целью обрабатываемый речевой сигнал задерживают на время, необходимое для поиска оптимальной системы базисных функций для его разложения и последующего синтеза обрабатываемого речевого сигнала, представленного в оптимальном функциональном базисе, сравнивают обрабатываемый и синтезированный речевые сигналы между собой, в результате чего вычисляют значение погрешности, с которой осуществлено спектральное представление обрабатываемого речевого сигнала.

Известно устройство для осуществления способа-прототипа, содержащее вход, два выхода, анализатор спектра и блок сегментации, выход которого соединен с первым выходом устройства, а второй выход устройства подключен к выходу анализатора спектра [3].

Работа известного устройства заключается в следующем. Обрабатываемый речевой сигнал сегментируют, после чего осуществляют спектральный анализ полученных сегментов, который является основой для последующей идентификации распознаваемого речевого сигнала.

Недостатком известного устройства-прототипа является низкая достоверность спектрального анализа распознаваемого речевого сигнала.

В основу изобретения поставлена задача создать устройство, которое путем поиска оптимальных систем базисных функций, предназначенных для спектрального разложения распознаваемых речевых сигналов и

последующего синтеза исходного сигнала с учетом оптимальных спектральных коэффициентов, с последующим вычислением величины расхождения исходного сигнала от обработанного позволило бы обеспечить получение на выходе устройства более достоверное спектральное описание исходного распознаваемого речевого сигнала.

Поставленная задача решается тем, что в устройство, содержащее вход, два выхода, анализатор спектра и блок сегментации, выход которого соединен с первым выходом устройства, а второй выход устройства подключен к выходу анализатора спектра, дополнительно введены блок сравнения, блок задержки и дополнительный выход, кроме того, анализатор спектра сделан адаптивным и выполнен с одним дополнительным выходом, причем вход устройства подключен ко входам адаптивного анализатора спектра и блока задержки, выход блока задержки и дополнительный выход анализатора спектра подключены ко входам блока сравнения, выход которого подключен к дополнительному выходу устройства, вход блока сегментации подключен к выходу анализатора спектра.

Структурная схема предлагаемого устройства приведена на чертеже.

Устройство содержит адаптивный анализатор спектра 1, блок сегментации 2, блок задержки 3, блок сравнения 4.

Элементы устройства соединены следующим образом. Вход устройства соединен со входами адаптивного анализатора спектра 1 и блока задержки 3, дополнительный выход адаптивного анализатора спектра 1 и выход блока задержки 3 подключены ко входам блока сравнения 4, выход которого подключен к дополнительному выходу устройства, первый выход устройства соединен с выходом блока сегментации 2, вход которого подключен к выходу адаптивного анализатора спектра 1 и ко второму выходу устройства.

При этом в качестве адаптивного анализатора спектра 1 можно использовать анализатор спектра, приведенный в [5], в котором вход необходимо подключить ко входу устройства, выход 1 необходимо подключить ко второму выходу предлагаемого устройства, а выход 2 соединить с входом блока сравнения 4.

В качестве блока сравнения 4 может использоваться любой сумматор двух аналоговых сигналов, один из которых проинвертирован с помощью усилителя, либо любое другое известное устройство, предназначенное для вычитания аналоговых сигналов, описанное, например, в [6].

В качестве блока сегментации 2 можно использовать устройство, принцип построения и алгоритм работы которого описан в [7].

Схема задержки 4 является типовым элементом радиоэлектронных систем и описана, например, в [5].

Устройство работает следующим образом.

Анализируемый речевой сигнал поступает на входы адаптивного анализатора спектра 1 и блока задержки 3. Адаптивный анализатор спектра 1 позволяет не только вычислять спектральные коэффициенты в различных базисах функций, но и определять систему базисных функций, оптимальную для исследуемого сигнала. Определенные таким образом оптимальные спектральные коэффициенты для каждого сегмента речи поступают с выхода адаптивного анализатора спектра 1 на второй выход устройства и, одновременно, на блок сегментации 2, на выходе которого формируются сегментирующие речевой сигнал импульсы в моменты изменения значений спектральных коэффициентов. Помимо вычисления оптимального спектрального базиса для разложения речевого сигнала, адаптивный анализатор спектра 1 после определения этого оптимального базиса синтезирует, на основе вычисленных коэффициентов, исходный речевой входной сигнал, который с дополнительного выхода адаптивного анализатора спектра 1 поступает на вход блока сравнения 4, на второй вход которого поступает естественный (несинтезированный) входной сигнал, задержанный в блоке задержки 3 на время преобразования и синтеза адаптивного анализатора спектра 1. Применение блока задержки 3 позволяет выровнять по времени синтезированный и естественный входные сигналы на входах блока 4. На выходе блока сравнения 4 будет наблюдаться величина погрешности представления речевых участков в оптимальном спектральном базисе.

Таким образом, в результате работы предлагаемого устройства по обработке и анализу речевых сигналов на первом выходе устройства будут формироваться импульсы сегментации, разделяющие входной речевой сигнал, на втором выходе устройства будут наблюдаться спектральные коэффициенты, имеющие минимальную погрешность представления каждого речевого сегмента, а на третьем выходе устройства будет наблюдаться величина этой погрешности представления.

Проведенный такой анализ речевого сигнала является более объективным и достоверным, поскольку он базируется не только на использовании как в процессе выбора параметров речевого сигнала для распознавания, так и в процессе самого распознавания речевого сигнала такого его спектрального изображения, которое имеет минимальную погрешность, но также вычислении величины этой погрешности.

