Изобретение относится к системам анализа речи и может быть использовано для определения эмоционального состояния человека по голосу, применительно к задачам криминалистики, медицины, системам контроля и управления доступом и др. Технический результат заключается в повышении достоверности и воспроизводимости оценок эмоционального состояния диктора. Способ заключается в записи речевого сигнала и его последующей обработке, включающей в себя расчет коэффициентов интегрального преобразования путем свертки речевого сигнала с ядром преобразования, и последующем анализе полученных коэффициентов на основе меры различимости. Коэффициентами интегрального преобразования являются коэффициенты локального вейвлет-спектра непрерывного вейвлет-преобразования речевого сигнала, и мерой различимости является евклидова невязка между локальными спектрами непрерывного вейвлет-преобразования. 1 з.п. ф-лы, 1 ил.

Изобретение относится к системам анализа речи и может быть использовано для определения эмоционального состояния человека по голосу, применительно к задачам криминалистики, медицины, системам контроля и управления доступом и др.

Известен способ обнаружения эмоций (Патент РФ №2287856, G06N 5/00, G10L 15/00, G06K 9/00, 2006), заключающийся в записи и анализе голосового сигнала, при котором выявляется изменение во времени, во-первых, интенсивности голоса, во-вторых, его темпа как характеристики скорости изменения голосового сигнала и, в-третьих, его интонации, характеризующей изменение интенсивности в каждом отрезке голосового сигнала. По полученным трем величинам изменения параметров голосового сигнала автоматически определяются состояния по меньшей мере гнева, печали и удовольствия.

Указанный способ характеризуется низкой достоверностью и плохой воспроизводимостью результатов, из-за неустойчивости к вариациям произнесения речевого материала диктором, так как результат детектирования эмоций существенным образом зависит от набора и характеристик голосовых фильтров, применяемых для установления специфических текстозависимых частотных компонент речевого сигнала;

отсутствия критерия оценки численного значения характерного времени речевых сегментов, соответствующего конкретной эмоции, а также математической неопределенности критерия выявления значимого рассогласования для различных эмоций «рисунка изменения интенсивности» в каждом слове голоса.

В другом способе определения эмоционального состояния человека по голосу (Горшков Ю.Г. Новые решения речевых технологий безопасности / Ю.Г.Горшков // Специальная техника. - 2006. - №4. - С.1-13) из записанного речевого сигнала с применением вейвлет-анализа выделяются и анализируются одновременно его основные параметры и кратковременные высокочастотные составляющие.

Недостатком данного способа является ограниченность его применения и неоднозначность получаемых результатов из-за отсутствия критериев различия и мер различимости, необходимых для детектирования эмоций, а также неопределенности параметров используемого вейвлета и выбора отсчетов масштаба вейвлет-преобразования.

Наиболее близким по совокупности признаков является способ анализа речи (Патент РФ №2403626, G10L 11/04, 2010), заключающийся в записи голосового сигнала диктора, его интегральном преобразовании в амплитудный спектр, вычислении автокорреляционного колебания при сдвиге полученного амплитудного спектра на частотной оси и вычислении частоты основного тона на основе локального интервала между одним из максимумов и одним из минимумов автокорреляционного колебания с последующим анализом изменения частоты основного тона на основе меры различимости.

В способе не определена мера различимости для сравнения различных эмоций, не используется информация о значениях амплитуд спектральных компонент голосового сигнала (распределении энергии голосового сигнала в значимых для решаемой задачи областях его амплитудного спектра); не учитывается форма частотно-временного распределения спектральных компонент, которые существенным образом определяют эмоциональное состояние человека (Галунов В.И. О возможности определения эмоционального состояния по речи / В.И.Галунов // Речевые технологии. - 2008. - №1. - С.60-66).

Все это приводит к уменьшению достоверности оценок эмоционального состояния диктора и их плохой воспроизводимости для различных дикторов.

Заявляемое изобретение предназначено для повышения вероятности правильных оценок эмоционального состояния человека по его речевому сигналу за счет использования непрерывного вейвлет-преобразования в качестве ядра интегрального преобразования и общей для детектирования различных видов эмоционального состояния диктора меры различимости.

Технический результат заключается в повышении достоверности и воспроизводимости оценок эмоционального состояния диктора.

Технический результат достигается тем, что в известном способе выявления эмоционального состояния человека по голосу, заключающийся в записи речевого сигнала и его последующей обработке, включающей в себя расчет коэффициентов интегрального преобразования путем свертки речевого сигнала с ядром преобразования, и последующем анализе полученных коэффициентов на основе меры различимости, согласно изобретению, коэффициентами интегрального преобразования являются коэффициенты локального вейвлет-спектра непрерывного вейвлет-преобразования речевого сигнала, и мерой различимости является евклидова невязка между локальными спектрами непрерывного вейвлет-преобразования; при этом для вычисления коэффициентов локального вейвлет-спектра непрерывного вейвлет-преобразования речевого сигнала используется вейвлет Морле.

Получаемый при осуществлении изобретения технический результат, а именно, повышение достоверности и воспроизводимости оценок эмоционального состояния человека по речевому сигналу, достигается за счет применения частотно-временного анализа существенных параметров речевого сигнала, адекватно описывающих эмоциональное состояние человека. В основе такого анализа лежит регистрация изменений во времени спектра непрерывного вейвлет-преобразования речевого сигнала, что позволяет получать аналитические выражения для практической оценки существенных параметров.

В непрерывном вейвлет-преобразовании в качестве материнского вейвлета используется вейвлет Морле (Голубинский А.Н. Выявление эмоционального состояния человека по речевому сигналу на основе вейвлет-анализа // Вестник ВИ МВД России. - 2011. - №3. - С.144-153), частотно-временные характеристики которого аналогичны характеристикам базилярной мембраны (Юрков П.Ю. Разработка и исследование методов и средств голосовой аутентификации с динамически изменяемым множеством ключевых слов: автореф. дис. канд. техн. наук: 05.13.19, 05.13.17 / П.Ю.Юрков. - Таганрог, 2006. - 16 с.). Существенные отсчеты масштаба непрерывного вейвлет-преобразования выбираются с использованием показательной аппроксимации функции масштаба (Голубинский А.Н., Асташов Р.А. К вопросу о выборе масштаба непрерывного вейвлет-преобразования для обработки речевых сигналов // Охрана, безопасность, связь - 2011: Материалы международной научно-практической конференции. Часть 1. - Воронеж: Воронежский институт МВД России, 2011. - С.64-68). Для детектирования эмоционального состояния применяется мера различимости между локальными спектрами НВП, что позволяет обеспечить повышение реальной и потенциальной точности оценки эмоционального состояния (Голубинский А.Н. Выявление эмоционального состояния человека по речевому сигналу на основе вейвлет-анализа // Вестник ВИ МВД России. - 2011. - №3. - С.144-153).

Заявляемый способ поясняется фиг.1, где схематически изображены основные блоки, реализующие способ выявления эмоционального состояния человека по голосу.

Блок-схема алгоритма выявления эмоционального состояния человека по голосу (фиг.1) включает источник речевого сигнала в цифровой или аналоговой форме, например микрофон (М) 1 и аналого-цифровой преобразователь (АЦП) 2, блок расчета существенных отсчетов масштаба непрерывного вейвлет-преобразования (БРСОМНВП) 3, блок расчета коэффициентов локального спектра непрерывного вейвлет-преобразования (БРКЛСНВП) 4, коммутатор 5, который работает следующим образом: нижнее положение переключателя - получение эталонных параметров голоса диктора, верхнее положение переключателя - детектирование эмоционального состояния, блок расчета мер различимости между локальными спектрами непрерывного вейвлет-преобразования (БРМРЛСНВП) 6, блок запоминания эталонных параметров голоса диктора (БЗЭПГД) 7, блок принятия решения об эмоциональном состоянии диктора (БПРЭСД) 8.

Предложенная блок-схема показывает работу в двух режимах: режиме обучения и режиме детектирования эмоционального состояния.

В режиме обучения речевой сигнал с микрофона 1 подают через АЦП 2 в блок 3, в котором на основе показательной аппроксимации функции масштаба вычисляются отсчеты масштаба НВП:

http://img.findpatent.ru/img_data/1208/12080158.jpg

Константы в показателях (1) связаны выражением:

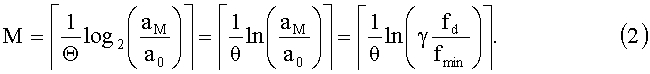
Θ=θ/ln(2).

Минимальное значение масштаба вейвлета:

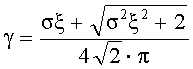
http://img.findpatent.ru/img_data/1208/12080159.jpg,

где Δt - эффективный временной размер материнского вейвлета, для вейвлета Морле равный: http://img.findpatent.ru/img_data/1208/12080160.jpg ; fd - частота дискретизации; σ - параметр масштаба.

Номер наибольшего отсчета масштаба рассчитывается по формуле:

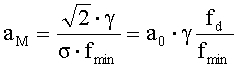


Здесь   - округление до ближайшего целого числа; fmin - минимальная существенная частота в спектре речевого сигнала;

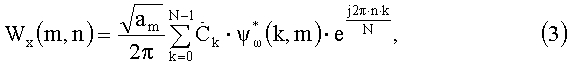
,

где ξ - доминантная частота, принимаемая равной: ξ=5/σ.

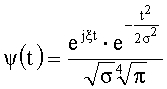
Значение М-го отсчета масштаба рассчитывается по формуле:

.

Затем в блоке 4 осуществляется расчет коэффициентов локального спектра непрерывного вейвлет-преобразования по формуле:

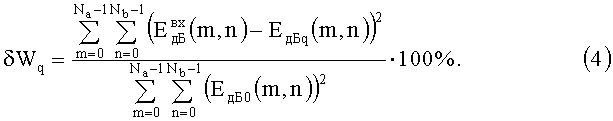


здесь http://img.findpatent.ru/img_data/1208/12080165.jpg - коэффициенты дискретного преобразования Фурье для отсчетов xi, вычисляемые на основе алгоритма быстрого преобразования Фурье (Баскаков С.И. Радиотехнические цепи и сигналы: Учебник для вузов / С.И.Баскаков. - 4-е изд. - М.: Высшая школа, 2005. - С.389-395.); N - количество отсчетов; ψ(k,m) - преобразование Фурье от материнского вейвлета Морле:



После этого в блоке 7 осуществляется запоминание значений коэффициентов локального спектра непрерывного вейвлет-преобразования.

В режиме детектирования эмоционального состояния после процедуры расчета коэффициентов локального спектра непрерывного вейвлет-преобразования в блоке 6 осуществляется расчет меры различимости по формуле:

δ W q = ∑ m = 0 N a − 1 ∑ n = 0 N b − 1 ( E д Б в х ( m , n ) − E д Б q ( m , n ) ) 2 ∑ m = 0 N a − 1 ∑ n = 0 N b − 1 ( E д Б 0 ( m , n ) ) 2 ⋅ 100 % .                      ( 4 )

ЗдесьE д Б в хhttp://img.findpatent.ru/img_data/1208/12080168.jpg - коэффициенты локального спектра непрерывного вейвлет-преобразования (в децибелах) входной реализации речевого сигнала;q = 1, Q ¯http://img.findpatent.ru/img_data/1208/12080169.jpg , где Q - количество видов эмоциональных состояний человека (в базе данных), например: депрессия, тоска, печаль, норма, радость, страх, гнев; ЕдБ0(m,n) - локальный спектр непрерывного вейвлет-преобразования в логарифмическом масштабе для эмоции "норма".

Затем в блоке 8 осуществляется определение наименьшего значения меры различимостиδ W min = min 4 [ δ W q ]http://img.findpatent.ru/img_data/1208/12080170.jpg . Если δWmin не превышает заданного порогового значения W0, то принимается решение, что входному речевому материалу соответствует эмоциональное состояние диктора, при котором наблюдалось δWmin.

Принципом детектирования какого-либо эмоционального состояния человека по его голосу, общим, для всех известных способов, является сравнение текущих значений некоторого набора параметров голосового сигнала, существенных для решения данной задачи, с «эталонными» значениями того же набора, однозначно характеризующими наличие данного эмоционального состояния. Вероятность соответствия текущего эмоционального состояния детектируемому состоянию будет тем выше, чем меньше значение меры различимости текущих и «эталонных» значений существенных параметров.

Достоверность способа детектирования эмоционального состояния диктора в общем случае обеспечивается эффективностью (прецизионностью) решающего критерия, состоящего из правила сравнения (решения) и меры различимости, т.е. в основном определяется прецизионностью меры различимости, которая в свою очередь обеспечивается:

- выбором существенных параметров как аргументов меры различимости;

- способом формирования значений существенных параметров, обеспечивающим высокую контрастность сопоставления текущего и заданного эмоционального состояния;

- способом формирования численного значения или вектора меры различимости.

Повышение вероятности правильного определения эмоционального состояния человека по его голосу в заявляемом способе обеспечивается:

- использованием в качестве существенных параметров коэффициентов локального спектра непрерывного вейвлет-преобразования, характеризующихся большой чувствительностью к изменению базовой функции, т.е. обладающих более существенными взаимными отличиями по сравнению с соответствующими значениями первичных параметров речевого сигнала (интенсивности, темпа, спектральной плотности мощности и др.);

- использованием вейвлета Морле в качестве материнского вейвлета в непрерывном вейвлет-преобразовании речевого сигнала, что повышает различимость существенных параметров;

- применением евклидовой невязки в качестве меры различимости, что позволяет однозначно и прецизионно сопоставлять текущее эмоциональное состояние диктора одновременно со всем набором детектируемых эмоциональных состояний по единственному текущему численному значению.

1. Способ выявления эмоционального состояния человека (депрессия, тоска, печаль, норма, радость, страх, гнев) по голосу, заключающийся в записи речевого сигнала и его последующей обработке, включающей в себя расчет частотно-временных параметров в виде коэффициентов интегрального преобразования путем свертки речевого сигнала с ядром преобразования, и последующем анализе полученных коэффициентов на основе меры различимости, отличающийся тем, что коэффициентами интегрального преобразования являются коэффициенты локального вейвлет-спектра непрерывного вейвлет-преобразования речевого сигнала, и мерой различимости является евклидова невязка между локальными спектрами непрерывного вейвлет-преобразования, эмоциональное состояние выявляется по наименьшему из значений меры различимости для всех эмоциональных состояний человека в базе данных, а для исключения из анализа речевого сигнала с не соответствующим базе данных эмоциональным состоянием диктора проводится сравнение меры различимости с заданным пороговым значением.

2. Способ по п. 1, отличающийся тем, что при вычислении коэффициентов локального вейвлет-спектра непрерывного вейвлет-преобразования речевого сигнала используется вейвлет Морле.

<http://www.findpatent.ru/patent/255/2553413.html>  
© FindPatent.ru - патентный поиск, 2012-2016

Изобретение относится к средствам распознавания эмоций человека по голосу. Технический результат заключается в повышении точности определения эмоционального состояния русскоязычного абонента. Обнаруживают интенсивности голоса и темпа, определяемого скоростью, с которой появляется голос, соответственно, и обнаруживают в виде величины времени, интонацию, которая отражает рисунок изменения интенсивности в каждом слове, выполняемом голосом, на основе введенного голосового сигнала. Получают первую величину изменения, указывающую изменение интенсивности обнаруженного голоса в направлении оси времени, и вторую величину изменения, указывающую изменение темпа голоса в направлении оси времени, и третью величину изменения, указывающую изменение интонации голоса в направлении оси времени. Вводят голосовой сигнал русскоязычного абонента, а затем обнаруживают интенсивности голоса и темпа. После того как получают третью величину измерения обнаруживают частоту основного тона голосового сигнала и получают четвертую величину изменения, указывающую изменение частоты основного тона в направлении оси времени, генерируют сигналы, выражающие эмоциональное состояние гнева, страха, печали и удовольствия, соответственно, на основе указанных первой, второй, третьей и четвертой величин изменения. 3 ил.

Изобретение относится к средствам распознавания эмоций человека по голосу и может быть использовано для обнаружения эмоций в интеллектуальных инфокоммуникационных системах, а также при проведении различного рода психологических исследований.

Расширение поля коммуникативного взаимодействия должностных лиц, а также постоянно растущие психологические нагрузки при принятии управленческих решений, связанные с уменьшением квоты доверия общающихся друг к другу, трансформируют формально-ролевое общение в деловое, при котором наряду с обменом информацией должны учитываться особенности личности абонента, его настроение, физиологическое и эмоциональное состояния. Перспективным в этом смысле может стать отказ от традиционных принципов кодирования и передачи аудио (речевых) сигналов в инфокоммуникационных системах в пользу интеллектуальной обработки сигналов.

Интеллектуальность (совмещение передачи и обработки информации на различных уровнях представления) инфокоммуникационных систем должна закладываться на ранних этапах их жизненного цикла и одной из функций реализовывать возможность определения эмоционального состояния абонента по голосу.

Известны способы определения эмоциональной напряженности (стресса) (патенты RU 2068653 от 10.11.1996 и RU 2073484 от 20.02.1997), согласно которым регистрируют кожно-гальваническую реакцию, частоту сердечных сокращений и частоту дыхания и по их динамике оценивают эмоциональную напряженность. Общим недостатком указанных аналогов является невозможность обнаружения эмоций (эмоциональной напряженности) человека без использования датчиков.

Известен способ определения эмоций по синтезированному речевому сигналу (патент JP 02-236600 от 19.09.1990), согласно которому из оцифрованного речевого сигнала выделяют частоту основного тона и рассчитывают амплитудный спектр, а затем на основе указанных параметров генерируют сигнал, выражающий эмоцию. Недостатком аналога является низкая точность обнаружения эмоциональных состояний.

Наиболее близким по технической сущности к заявляемому способу и выбранным в качестве прототипа является способ обнаружения эмоций (патент RU 2287856 от 20.11.2006), заключающийся в том, что вводят голосовой сигнал; обнаруживают интенсивности голоса и темп, определяемый скоростью, с которой появляется голос, соответственно, и обнаруживают в виде величины времени, интонацию, которая отражает рисунок изменения интенсивности в каждом слове, выполняемом голосом, на основе введенного голосового сигнала; получают первую величину изменения, указывающую изменение интенсивности обнаруженного голоса в направлении оси времени, вторую величину изменения, указывающую изменение темпа голоса в направлении оси времени, и третью величину изменения, указывающую изменение интонации голоса в направлении оси времени, соответственно; и генерируют сигналы, выражающие эмоциональное состояние по меньшей мере гнева, печали и удовольствия, соответственно, на основе указанных первой, второй и третьей величин изменения.

Способ-прототип предусматривает распознавание эмоций на основе изменений интенсивности, темпа и интонации голоса во времени. Однако в большинстве языков эмоционально-различительную функцию выполняет частота основного тона (ЧОТ). В [А.Ф.Хроматиди, И.Б.Старченко. Оценка акустических параметров эмоциональной речи / Первая ежегодная научная конференция студентов и аспирантов базовых кафедр ЮНЦ РАН, 2009. - С.212-214] установлено, что среднее значение ЧОТ поднимается в состоянии удовольствия и понижается в состоянии печали, кроме того существенно меняется динамика изменения ЧОТ: при печали происходит ее плавное уменьшение, при гневе появляются резкие пики в изменении частоты. Таким образом, динамика изменения ЧОТ является важнейшим средством определения эмоциональной информации, а недостатком способа-прототипа является низкая точность обнаружения эмоций, в частности обнаружение эмоций для русского языка.

Задачей изобретения является разработка способа обнаружения эмоций по голосу, позволяющего повысить точность определения эмоционального состояния русскоязычного абонента.

В заявленном способе эта задача решается тем, что в способе обнаружения эмоций по голосу, в котором вводят голосовой сигнал; обнаруживают интенсивности голоса и темп, определяемый скоростью, с которой появляется голос, соответственно, и обнаруживают в виде величины времени, интонацию, которая отражает рисунок изменения интенсивности в каждом слове, выполняемом голосом, на основе введенного голосового сигнала; получают первую величину изменения, указывающую изменение интенсивности обнаруженного голоса в направлении оси времени, вторую величину изменения, указывающую изменение темпа голоса в направлении оси времени, и третью величину изменения, указывающую изменение интонации голоса в направлении оси времени, соответственно; дополнительно обнаруживают частоту основного тона голосового сигнала и получают четвертую величину изменения, указывающую изменение частоты основного тона в направлении оси времени. Затем генерируют сигналы, выражающие эмоциональное состояние гнева, страха, печали и удовольствия, соответственно, на основе указанных первой, второй, третьей и четвертой величин изменения.

Новая совокупность существенных признаков позволяет достичь указанного технического результата за счет обнаружения изменения частоты основного тона и генерации сигналов, выражающих эмоциональное состояние говорящего, на основе четырех величин изменения.

Проведенный анализ уровня техники позволил установить, что аналоги, характеризующиеся совокупностью признаков, тождественных всем признакам заявленного способа обнаружения эмоций, отсутствуют. Следовательно, заявленное изобретение соответствует условию патентоспособности «новизна».

Результаты поиска известных решений в данной и смежных областях техники с целью выявления признаков, совпадающих с отличительными от прототипа признаками заявленного объекта, показали, что они не следуют явным образом из уровня техники. Из уровня техники также не выявлена известность влияния предусматриваемых существенными признаками заявленного изобретения преобразований на достижение указанного технического результата. Следовательно, заявленное изобретение соответствует условию патентоспособности «изобретательский уровень».

Заявленное изобретение поясняется следующими фигурами:

на фиг.1 - вариант реализации системы обнаружения эмоций по голосу согласно предлагаемому способу;

на фиг.2 - решающие правила определения эмоций согласно предлагаемому способу;

на фиг.3 - результаты оценки точности определения эмоционального состояния.

Реализация заявленного способа заключается в следующем (фиг.1).

Голосовой сигнал, введенный через микрофон 101, квантуется с помощью аналого-цифрового преобразователя 102, а затем преобразуется в цифровой сигнал. Цифровой голосовой сигнал, полученный на выходе аналого-цифрового преобразователя, подается в блок 103 обработки сигналов, блок 104 обнаружения фонем, блок 105 обнаружения слов и блок 106 обнаружения частоты основного тона.

Блок 103 обработки сигналов извлекает частотные составляющие, необходимые для обнаружения интенсивности голоса. Блок 107 обнаружения интенсивности обнаруживает интенсивность сигнала, извлеченного блоком 103 обработки сигналов. Например, в качестве интенсивности можно использовать результат, полученный путем усреднения величины амплитуды голосового сигнала или его динамического диапазона D.

Блок 104 обнаружения фонем реализует сегментацию каждой фонемы голосового сигнала, введенного в него. Блок 108 обнаружения темпа принимает сигнал сегментации каждой фонемы, выданный блоком 104 обнаружения фонем, и обнаруживает число фонем F, которые появляются в единицу времени. В качестве цикла обнаружения темпа устанавливается время, равное, например, 10 с. Однако, если обнаружена сегментация фразы, то отсчет фонем останавливается до момента времени обнаружения сегментации фразы, даже если сегментация фразы обнаружена внутри 10 с, и вычисляется величина темпа. В частности, темп определяется для каждой фразы.

Блок 105 обнаружения слов реализует сегментацию каждого слова голосового сигнала, введенного в него. Блок 109 обнаружения интонации принимает сигнал сегментации каждого слова, выданный блоком 105 обнаружения слов, и обнаруживает интонацию, выражающую рисунок изменения интенсивности голоса в слове. Таким образом, блок 109 обнаружения интонации обнаруживает характеристический рисунок интенсивности в сегментации. Как показано в прототипе, в блоке 109 обнаружения интонации предусмотрены полосовой фильтр, блок преобразования абсолютной величины, блок сравнения, блок обнаружения центра зоны и блок обнаружения интервала зон. В качестве величины интонации I на выходе блока 109 обнаружения интонации выступает результат усреднения значений интервалов, между зонами в спектре мощности сигналов, для которых характерно превышение некоторого порогового значения.

Блок 106 обнаружения частоты основного тона реализует определение частоты основного тона введенного в него голосового сигнала. Блок 106 обнаружения частоты основного тона FОТ может быть реализован, например, в соответствии с известным решением (патент №78977 от 10.12.2008).

Эмоциональное состояние человека изменяется, поэтому для правильного определения эмоций, включающих гнев, страх, печаль и удовольствие, обязательно необходимо обнаруживать изменение характеристических величин, таких как интенсивность D, темп F, интонация I и частота основного тона FОТ.

В системе обнаружения эмоций, показанной на фиг.1, с целью обеспечения возможности опоры на величины характеристик в прошлом, величину интенсивности D, выдаваемую блоком 107 обнаружения интенсивности, величину темпа F, выдаваемую блоком 108 обнаружения темпа, величину интонации I, выдаваемую блоком 109 обнаружения интонации, и величину частоты основного тона FОТ, выдаваемую блоком 106 обнаружения частоты основного тона, временно сохраняют в блоке 110 временного хранения данных.

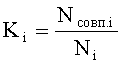
Кроме того, блок 111 обнаружения изменения эмоций принимает имеющуюся величину интенсивности D, выдаваемую блоком 107 обнаружения интенсивности, имеющуюся величину темпа F, выдаваемую блоком 108 обнаружения темпа, имеющуюся величину интонации I, выдаваемую блоком 109 обнаружения интонации, и имеющуюся величину частоты основного тона FОТ, выдаваемую блоком 106 обнаружения частоты основного тона. Блок 111 обнаружения изменения эмоций также принимает прошлые величины интенсивности, темпа, интонации и частоты основного тона, которые хранятся в блоке 110 временного хранения данных. Таким образом, блок 111 обнаружения изменения эмоций обнаруживает изменения в интенсивности, темпе, интонации и частоте основного тона голоса, соответственно. Блок 112 обнаружения эмоций по голосу принимает изменения интенсивности ΔD, темпа ΔF, интонации ΔI и частоты основного тона ΔFОТ голоса, которые выдает блок 111 обнаружения изменения эмоций, оценивает текущее эмоциональное состояние и генерирует сигналы, выражающие эмоциональное состояние гнева, страха, печали и удовольствия, в этом варианте реализации системы.

Заявленный способ обнаружения эмоций по голосу обеспечивает повышение точности определения эмоционального состояния русскоязычного абонента. Для доказательства достижения заявленного технического результата приведены следующие экспериментальные исследования.

Для определения эмоционального состояния использовались записи эмоциональной речи 80 профессиональных актеров - мужчин и женщин в возрасте от 28 до 32 лет. Каждым из них были произнесены 4 слова (картон, тихо, молоко, посуда) с выражением четырех эмоциональных состояний: гнева, страха, печали и удовольствия.

Указанные записи обрабатывались с использованием варианта выполнения системы обнаружения эмоций согласно способу-прототипу и варианту реализации системы обнаружения эмоций по голосу (фиг.1) согласно предлагаемому способу. При этом блок 112 обнаружения эмоций по голосу оценивал текущее эмоциональное состояние и генерировал сигналы, выражающие эмоциональное состояние гнева, страха, печали и удовольствия, согласно решающим правилам определения эмоций, представленным на фиг.2.

Для оценки точности определения эмоционального состояния русскоязычного абонента использовался коэффициент совпадений

K i = N с о в п . i N i ,

где Nсовп.i - число правильно определенных записей с выражением i-го эмоционального состояния; Ni - общее число записей с выражением г-го эмоционального состояния; i=1, 2, 3, 4 - номер эмоционально состояния - гнева, страха, печали и удовольствия соответственно.

Результаты оценивания согласно способу-прототипу и предлагаемому способу (фиг.3) указывают на повышение точности определения эмоционального состояния в заявляемом способе и на возможность решения поставленной задачи изобретения.

Способ обнаружения эмоций по голосу, заключающийся в том, что обнаруживают интенсивности голоса и темпа, определяемого скоростью, с которой появляется голос, соответственно, и обнаруживают в виде величины времени интонацию, которая отражает рисунок изменения интенсивности в каждом слове, выполняемом голосом, на основе введенного голосового сигнала; получают первую величину изменения, указывающую изменение интенсивности обнаруженного голоса в направлении оси времени, и вторую величину изменения, указывающую изменение темпа голоса в направлении оси времени, и третью величину изменения, указывающую изменение интонации голоса в направлении оси времени, отличающийся тем, что вводят голосовой сигнал русскоязычного абонента, а затем обнаруживают интенсивности голоса и темпа; после того как получают третью величину измерения, обнаруживают частоту основного тона голосового сигнала и получают четвертую величину изменения, указывающую изменение частоты основного тона в направлении оси времени; генерируют сигналы, выражающие эмоциональное состояние гнева, страха, печали и удовольствия, соответственно, на основе указанных первой, второй, третьей и четвертой величин изменения.

<http://www.findpatent.ru/patent/251/2510955.html>  
© FindPatent.ru - патентный поиск, 2012-2016