

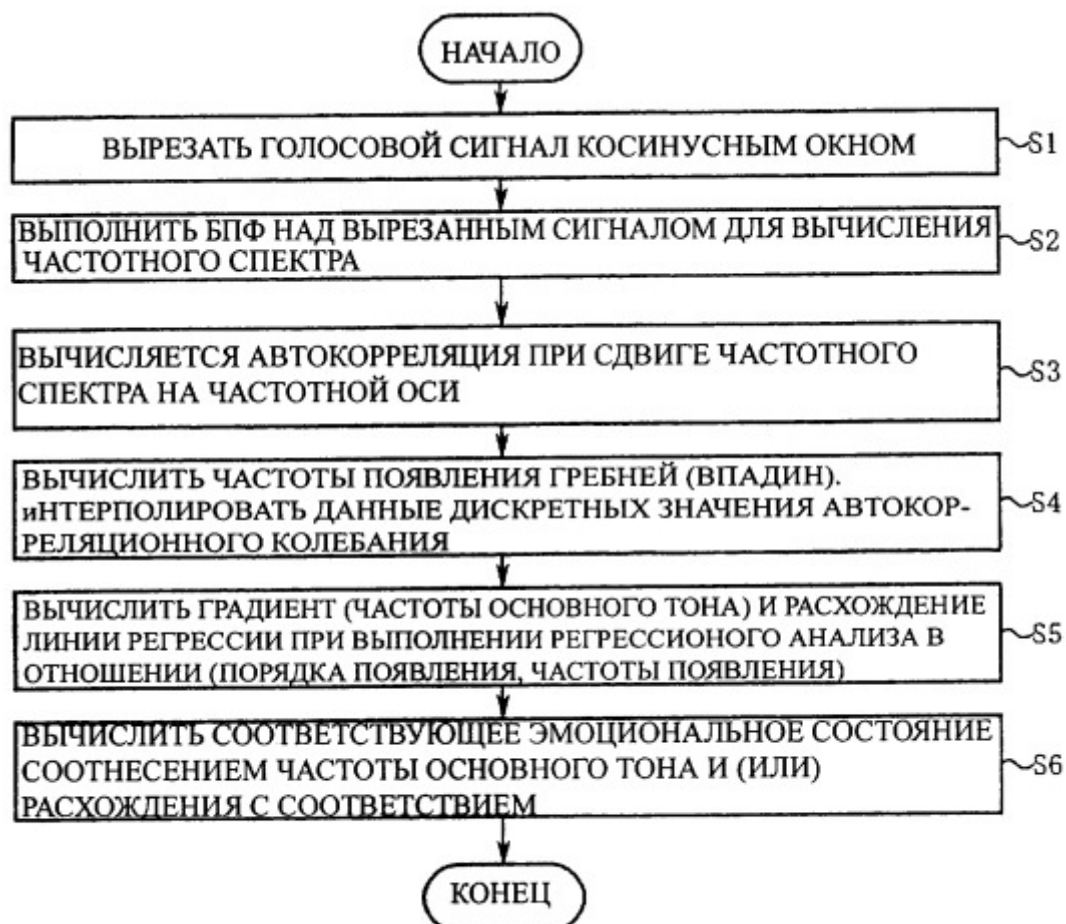
FindPatent.ru

Патентный поиск

найти

Найти

Анализатор речи, обнаруживающий частоту основного тона, способ анализа речи и программа анализа речи



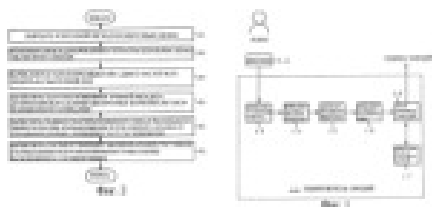
Фиг. 2

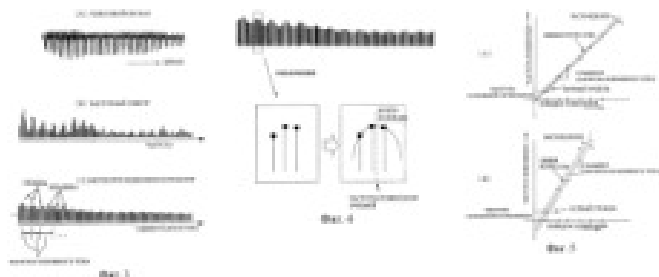
Авторы патента:

МОНМА Фумиаки (JP)

МИЦУЁСИ Сундзи (JP)

ОГАТА Каору (JP)





G10L11/04 - определение пика речевых сигналов

Владельцы патента RU 2403626:

А.Г.И. ИНК. (JP)

МИЦУЁСИ Сундзи (JP)

Изобретение относится к методу анализа речи, обнаруживающему частоту основного тона голоса, а также к методу обнаружения эмоций, оценивающему эмоцию на основании частоты основного тона голоса. Технический результат - повышение точного и уверенного обнаружения голосовой частоты и повышение точности оценки эмоций. Анализатор речи включает в себя блок поступления голосового сигнала, блок частотного преобразования, автокорреляционный блок и блок обнаружения основного тона. Блок частотного преобразования преобразует голосовой сигнал, полученный в блоке поступления голосового сигнала, в частотный спектр. Автокорреляционный блок вычисляет автокорреляционное колебание при сдвиге частотного спектра на частотной оси. Блок обнаружения основного тона вычисляет частоту на основе локального интервала между гребнями или впадинами автокорреляционного колебания. 3 н. и 6 з.п. ф-лы, 5 ил.

Область техники, к которой относится изобретение

Настоящее изобретение относится к методу анализа речи, обнаруживающему частоту основного тона голоса.

Изобретение также относится к методу обнаружения эмоций, оценивающему эмоцию на основании частоты основного тона голоса.

Уровень техники

Из уровня техники известны методы, оценивающие эмоцию экзаменуемого путем анализа голосового сигнала этого экзаменуемого.

Например, в патентном документе 1 раскрывается метод, в котором вычисляется основная частота поющего голоса, и эмоция певца оценивается из поднимающихся и падающих изменений основной частоты в конце пения.

Патентный документ 1: заявка Японии №10-187178 (1998).

Сущность изобретения

Проблема, подлежащая решению изобретением

Основная частота ясно возникает в музыкальном инструментальном звуке, и основную частоту легко обнаружить.

Однако поскольку голос, в общем, включает в себя хриплый голос, вибрирующий голос и тому подобное, основная частота флуктуирует. Кроме того, компоненты гармонического тона будут нерегулярными.

Поэтому эффективный способ, с уверенностью обнаруживающий основную частоту из этого вида голоса, не установлен.

Соответственно цель изобретения состоит в обеспечении метода точного и уверенного обнаружения голосовой частоты.

Другая цель изобретения состоит в обеспечении нового метода оценки эмоций на основе обработки речи.

Средства для решения этой проблемы

(1) Анализатор речи согласно изобретению включает в себя блок поступления голосового сигнала, блок частотного преобразования, автокорреляционный блок и блок обнаружения основного тона.

Блок поступления голосового сигнала получает голосовой сигнал от проверяемого.

Блок частотного преобразования преобразует голосовой сигнал в частотный спектр.

Корреляционный блок вычисляет автокорреляционное колебание при сдвиге частотного спектра на частотной оси.

Блок обнаружения основного тона вычисляет частоту на основе локального интервала между гребнями или впадинами автокорреляционного колебания.

(2) Автокорреляционный блок предпочтительно вычисляет дискретные данные автокорреляционного колебания при дискретном сдвиге частотного спектра на частотной оси. Блок обнаружения основного тона интерполирует дискретные данные автокорреляционного колебания и вычисляет частоты появления локальных гребней или впадин по линии интерполяции. Блок обнаружения основного тона вычисляет частоту основного тона на основе интервала частот появления, вычисленного, как указано выше.

(3) Блок обнаружения основного тона предпочтительно вычисляет множество (порядок появления, частоту появления) по отношению к по меньшей мере одному из гребней или впадин автокорреляционного колебания. Блок обнаружения основного тона выполняет регрессионный анализ порядков появления и частот появления и вычисляет частоту основного тона на основе градиента полученной линии регрессии.

(4) Блок обнаружения основного тона предпочтительно исключает отсчеты, флуктуация уровня которых в автокорреляционном колебании мала, из распределения вычисленного множества (порядок появления, частота появления). Блок обнаружения основного тона выполняет регрессионный анализ по отношению к остальному распределению и вычисляет частоту основного тона на основе градиента полученной линии регрессии.

(5) Блок обнаружения основного тона предпочтительно включает в себя блок выделения и блок вычитания.

Блок выделения выделяет «компоненты, зависящие от формант», включенных в автокорреляционное колебание, путем выполнения аппроксимации кривой к автокорреляционному колебанию.

Блок вычитания вычисляет автокорреляционное колебание, в котором эффект формант ослаблен за счет исключения компонент из автокорреляционного колебания.

Согласно этой конфигурации блок обнаружения основного тона вычисляет частоту основного тона на основе автокорреляционного колебания, в котором эффект формант ослаблен.

(6) Вышеуказанный анализатор речи предпочтительно включает в себя блок хранения соответствия и блок оценки эмоций.

Блок хранения соответствия сохраняет по меньшей мере соответствие между «частотой основного тона» и «эмоциональным состоянием».

Блок оценки эмоций оценивает эмоциональное состояние проверяемого путем соотнесения с соответствием для частоты основного тона, обнаруженной блоком обнаружения основного тона.

(7) В вышеуказанном анализаторе речи по (3) блок обнаружения основного тона предпочтительно вычисляет по меньшей мере одно из «степени изменения (порядка появления, частоты появления) по

отношению к линии регрессии» и «отклонение между линией регрессии и исходными точками» в качестве нерегулярности частоты основного тона. Анализатор речи обеспечивается блоком хранения соответствия и блоком оценки эмоций.

Блок хранения соответствия сохраняет по меньшей мере соответствие между «частотой основного тона», а также «нерегулярностью частоты основного тона» и «эмоциональным состоянием».

Блок оценки эмоций оценивает эмоциональное состояние проверяемого путем соотнесения соответствия для «частоты основного тона» и «нерегулярности частоты основного тона», вычисленных в блоке обнаружения основного тона.

(8) Способ анализа речи в изобретении включает в себя следующие этапы:

- 1) получают голосовой сигнал от проверяемого,
- 2) преобразуют голосовой сигнал в частотный спектр,
- 3) вычисляют автокорреляционное колебание при сдвиге частотного спектра на частотной оси,
- 4) вычисляют частоту основного тона на основе локального интервала между гребнями или впадинами автокорреляционного колебания.

(9) Программа анализа речи по изобретению представляет собой программу, позволяющую компьютеру функционировать в качестве анализатора речи по любому из вышеуказанных пунктов (1)-(7).

Преимущество изобретения

[1] В изобретении голосовой сигнал преобразуется в частотный спектр единойжды. Частотный спектр включает в себя флюктуацию частоты основной гармоники и нерегулярность гармонических тональных компонент в качестве шума. Поэтому трудно считывать частоту основной гармоники из частотного спектра.

В изобретении автокорреляционное колебание вычисляется при сдвиге частотного спектра на частотной оси. В автокорреляционном колебании спектральный шум с низкой периодичностью подавляется. В результате в автокорреляционном колебании гармонические тональные компоненты с сильной периодичностью появляются как периодические гребни.

В изобретении частота основного тона вычисляется точно путем вычисления локального интервала между гребнями или впадинами, появляющимися периодически, на основе автокорреляционного колебания, шум которого делают низким.

Частота основного тона, вычисленная, как указано выше, иногда походит на частоту основной гармоники, однако она не всегда соответствует частоте основной гармоники, потому что частота основного тона не вычисляется из максимального пика или первого пика автокорреляционного колебания. Можно вычислять частоту основного тона стабильно и точно даже из голосового сигнала, у которого частота основной гармоники нечеткая, путем вычисления частоты основного тона из интервала между гребнями (впадинами).

[2] В изобретении предпочтительно вычислять дискретные данные автокорреляционного колебания при дискретном сдвиге частотного спектра на частотной оси. Согласно дискретной обработке число вычислений можно снизить, а время обработки можно укоротить. Однако частота для дискретного сдвига становится большой, разрешение автокорреляционного колебания становится низким и точность обнаружения частоты основного тона снижается. Соответственно можно вычислять частоту основного тона более точно, чем разрешение дискретных данных, путем интерполяции дискретных данных автокорреляционного колебания и точного вычисления частот появления локальных гребней (или впадин).

[3] Имеется случай, в котором локальные интервалы гребней (или впадин), периодически появляющихся в автокорреляционном колебании, не равны в зависимости от голоса. В это время трудно вычислить точную частоту основного тона, если о частоте основного тона принимается решение по отнесению только к одному определенному интервалу. Соответственно предпочтительно вычислять множество (порядок появления, частота появления) по отношению к по меньшей мере одному из гребней или

впадин автокорреляционного колебания. Возможно вычислять частоту основного тона, в которой расхождения неравных интервалов усредняются аппроксимацией этих (порядка появления, частоты появления) посредством линии регрессии.

Возможно вычислять частоту основного тона точно даже из чрезвычайно слабого голосового сигнала согласно такому способу вычисления частоты основного тона. В результате успешная степень оценки эмоций может увеличиться по отношению к голосовому сигналу, анализ частоты основного тона которого затруднен.

[4] Трудно точно вычислять частоты появления гребней и впадин, потому что точка, где флюктуация уровня мала, становится пологим гребнем (или впадиной). Соответственно предпочтительно, чтобы отсчеты, флюктуация уровня которых в автокорреляционном колебании мала, исключались из распределения (порядка появлений, частоты появлений), вычисленных, как указано выше. Возможно вычислять частоту основного тона более стабильно и точно за счет выполнения регрессионного анализа по отношению к распределению, ограниченному таким образом.

[5] Конкретные пики, перемещающиеся со временем, появляются в частотных компонентах голосового сигнала. Эти пики называются формантами. Компоненты, отражающие форманты, появляются в автокорреляционном колебании в дополнение к гребням и впадинам этого колебания. Соответственно автокорреляционное колебание аппроксимируется кривой, подходящей к флюктуациям автокорреляционного колебания. Расценивается, что кривая представляет собой «компоненты, зависящие от формант», включенных в автокорреляционное колебание». Возможно вычислять автокорреляционное колебание, в котором эффект от формант ослабляется, путем вычитания этих компонент из автокорреляционного колебания. В автокорреляционном колебании, в котором выполняется такая обработка, искажения, вызванные формантами, снижаются. Соответственно, можно вычислять частоту основного тона более точно и уверенно.

[6] Частота основного тона, полученная вышеуказанным образом, является параметром, представляющим характеристики, такие как высота голоса или тембр голоса, который ощутимо меняется согласно эмоциям во время речи. Поэтому возможно выполнять оценку эмоций уверенно даже в голосе, в котором трудно обнаружить частоту основной гармоникой, с помощью частоты основного тона в качестве оценки эмоции.

[7] Помимо того, предпочтительно обнаруживать нерегулярность интервалов между периодическими гребнями (или впадинами) в качестве новой характеристики голоса. Например, степень расхождения (порядка появления, частоты появления) по отношению к линии регрессии вычисляется статистически. Кроме того, например, вычисляется отклонение между линией регрессии и исходными точками.

Нерегулярность, вычисленная, как указано выше, показывает качество улавливающей голос среды, а также представляет мгновенное изменение голоса. Соответственно возможно увеличить виды эмоций для оценки и увеличить степень успеха в оценке мгновенной эмоции путем добавления нерегулярности к частоте основного тона в качестве элемента для оценки эмоций.

Вышеуказанная цель и иные цели в изобретении будут конкретно показаны в нижеследующем пояснении и приложенных чертежах.

Краткое описание чертежей

Фиг.1 представляет собой блок-схему, показывающую детектор 11 эмоций (включающий в себя анализатор речи);

фиг.2 является блок-схемой алгоритма, поясняющей работу детектора 11 эмоций;

фиг.3(А-С) представляют собой виды, поясняющие обработку для голосового сигнала;

фиг.4 представляет собой вид, поясняющий интерполяционную обработку автокорреляционного колебания;

фиг.5(А, В) являются графиками, поясняющими соотношение между линией регрессии и частотой основного тона.

Подробное описание изобретения

Конфигурация варианта осуществления

Фиг.1 представляет собой блок-схему, показывающую детектор 11 эмоций (включающий в себя анализатор речи).

На фиг.1 детектор 11 эмоций включает в себя следующее:

- (1) микрофон 12 - голос проверяемого преобразуется в голосовой сигнал.
- (2) Блок 13 поступления голосового сигнала - поступает голосовой сигнал.
- (3) Блок 14 частотного преобразования - над поступившим голосовым сигналом осуществляют частотное преобразование для вычисления частотного спектра.
- (4) Автокорреляционный блок 15 - автокорреляция частотного спектра вычисляется на частотной оси, и частотная компонента, периодически появляющаяся на частотной оси, вычисляется как автокорреляционное колебание.
- (5) Блок 16 обнаружения основного тона - частотный интервал между гребнями (или впадинами) в автокорреляционном колебании вычисляется в качестве частоты основного тона.
- (6) Блок 17 хранения соответствия - сохраняет соответствие между информацией решения, такой как частота или изменение основного тона, и эмоциональным состоянием проверяемого. Это соответствие может быть создано посредством связывания экспериментальных данных, таких как частота или изменение основного тона, с эмоциональным состоянием, декларируемым проверяемым (злость, радость, напряжение, печаль и т.д.). Формой описания этого соответствия предпочтительно является таблица соответствия, решающая логика или нейронная сеть.
- (7) Блок 18 оценки эмоций - частота основного тона, вычисленная в блоке 16 обнаружения основного тона, соотносится с соответствием в блоке 17 хранения соответствия, чтобы принять решение о соответствующем эмоциональном состоянии. Эмоциональное состояние, по которому принято решение, выводится в качестве оцененной эмоции.

Часть или все из вышеуказанных блоков 13-18 могут выполняться в аппаратном виде. Предпочтительно также реализовать часть или все из вышеуказанных блоков 13-18 с помощью программного обеспечения за счет исполнения программы обнаружения эмоций (программы анализатора речи) в компьютере.

Пояснение работы детектора 11 эмоций

Фиг.2 является блок-схемой алгоритма, поясняющей работу детектора 11 эмоций.

Здесь и далее конкретная операция будет поясняться вместе с номерами этапов, показанных на фиг.2.

Этап S1. Блок 14 частотного преобразования вырезает голосовой сигнал необходимого сечения для вычисления БПФ (быстрого преобразования Фурье) из блока 13 поступления голосового сигнала (см. фиг.3(A)). В это время вырезающая функция, такая как косинусное окно, работает для вырезания участка, чтобы ослабить эффект на обоих концах вырезанного участка.

Этап S2. Блок 14 частотного преобразования выполняет вычисление БПФ над голосовым сигналом, обработанным вырезающей функцией, для вычисления частотного спектра (см. фиг.3(B)).

Поскольку отрицательное значение генерируется, когда обработка подавления уровня посредством общего логарифмического вычисления выполняется в отношении частотного спектра, описанное далее вычисление автокорреляции будет сложным и трудным. Поэтому для частотного спектра предпочтительно выполнять обработку подавления уровня, такую как вычисление корней, в ходе которой можно получить положительное значение, а не обработку подавления уровня путем вычисления логарифма.

Когда усиливается расхождение частотного спектра, может выполняться обработка усиления, такая как вычисление четвертой степени для значения частотного спектра.

Этап S3. В частотном спектре спектр, соответствующий гармоническому тону, такому как звук музыкального инструмента, появляется периодически. Однако поскольку частотный спектр речевого голоса включает в себя усложненные компоненты, как показано на фиг.3(B), ясно определить периодический спектр затруднительно. Соответственно автокорреляционный блок 15 последовательно вычисляет значение автокорреляции при сдвиге частотного спектра на заранее заданную величину в направлении по частотной оси. Дискретные данные значений автокорреляции, полученные посредством этого вычисления, строятся в соответствии со сдвинутой частотой, благодаря чему получаются автокорреляционные колебания (см. фиг.3(C)).

Частотный спектр включает в себя ненужные компоненты, иные, нежели включаются в голосовой диапазон (компоненты постоянного тока и чрезвычайно низкочастотные компоненты). Эти ненужные компоненты ухудшают вычисление автокорреляции. Поэтому предпочтительно, чтобы блок 14 частотного преобразования подавил или удалил эти ненужные компоненты из частотного спектра до вычисления автокорреляции.

Например, предпочтительно вырезать компоненты постоянного тока (например, 60 Гц или меньше) из частотного спектра.

Помимо этого, например, предпочтительно вырезать мгновенные частотные компоненты в качестве шума путем установки заданного более низкого стационарного уровня (например, средний уровень частотного спектра) и выполнения отсечки (более низкий стационарный предел) частотного спектра.

Согласно такой обработке искажение колебания, происходящее при вычислении автокорреляции, можно предотвратить до его появления.

Этап S4. Автокорреляционное колебание представляет собой дискретные данные, как показано на фиг.4. Соответственно блок 16 обнаружения основного тона вычисляет частоты появления в отношении множества гребней и (или) впадин путем интерполяции дискретных данных. Например, в качестве способа интерполяции в этом случае предпочтителен способ интерполяции дискретных данных вблизи гребней или впадин посредством линейной интерполяции или функции кривой, потому что он прост. Когда интервалы дискретных данных достаточно узкие, возможно опустить интерполяционную обработку дискретных данных. Соответственно вычисляется множество данных отсчетов от (порядка появления, частоты появления).

Вычислить точно частоты появления гребней или впадин трудно, потому что точка, где флюктуация уровня автокорреляционного колебания мала, становится пологим гребнем (или впадиной). Поэтому неточные частоты появления исключаются в качестве отсчетов как таковых, при этом снижается точность обнаруженной потом частоты основного тона. Следовательно, по данным отсчетов, у которых флюктуация уровня автокорреляционного колебания мала, принимается решение в распределении от (порядка появления, частоты появления), вычисленных, как указано ранее. Затем распределение, пригодное для анализа частоты основного тона, получается путем отсека от распределения тех данных отсчетов, о которых принято решение вышеуказанным образом.

Этап S5. Блок 16 обнаружения основного тона вычитает данные отсчетов соответственно из распределения, полученного на этапе S4, размещая частоты появления согласно порядку появления. В это время порядок появления, который был отсека, потому что флюктуация уровня автокорреляционного колебания мала, будет пропущенным числом.

Блок 16 обнаружения основного тона выполняет регрессионный анализ в координатном пространстве, в котором размещены данные отсчетов, вычисляя градиент линии регрессии. Частота основного тона, из которой вырезана флюктуация частоты появления, может быть вычислена на основе этого градиента.

При выполнении регрессионного анализа блок 16 обнаружения основного тона статистически вычисляет изменение частот появления в отношении линии регрессии в качестве расхождения частоты основного тона.

Помимо этого, вычисляется отклонение между линией регрессии и исходными точками (например, пересечение линии регрессии), и в случае, когда это отклонение больше заранее заданного допустимого предела, можно принять решение, что участок голосового сигнала не пригоден для обнаружения

основного тона (шум и тому подобное). В этом случае предпочтительно обнаруживать частоту основного тона в отношении остальных участков голосового сигнала, иных, нежели этот участок голосового сигнала.

Этап S6. Блок 18 оценки эмоций принимает решение о соответствующем эмоциональном состоянии (злость, радость, напряжение, печаль и т.п.) путем ссылки на соответствие в блоке 17 хранения соответствия для данных (частота основного тона, расхождение), вычисленных на этапе S5.

Преимущество этого варианта осуществления и тому подобное

Сначала со ссылкой на фиг.5(A, B) будет поясняться различие между настоящим изобретением и прототипом.

Частота основного тона этого варианта осуществления соответствует интервалу между гребнями (или впадинами) автокорреляционного колебания, который соответствует градиенту линии регрессии на фиг.5(A, B). С другой стороны, традиционная частота основной гармоники соответствует частоте появления первого гребня, показанного на фиг.5(A, B).

На фиг.5(A) линия регрессии проходит вблизи исходных точек, а потому расхождение мало. В этом случае в автокорреляционном колебании гребни появляются регулярно на почти равных интервалах. Поэтому частоту основной гармоники можно ясно обнаружить даже в прототипе.

С другой стороны, на фиг.5(B) линия регрессии широко отклоняется от исходных точек, то есть расхождение велико. В этом случае гребни автокорреляционного колебания появляются на нерегулярных интервалах. Поэтому частота основной гармоники представляет собой неясную речь и определить частоту основной гармоники трудно. В прототипе частота основной гармоники вычисляется из частоты появления на первом гребне, поэтому в таком случае вычисляется неверная частота основной гармоники.

В изобретении же в этом случае надежность частоты основного тона можно найти на основе того, проходит ли линия регрессии, найденная из частот появления гребней, вблизи исходных точек, или же расхождение частоты основного тона мало или отсутствует. Поэтому в данном варианте осуществления определяют, что надежность частоты основного тона в отношении голосового сигнала по фиг.5(B) низка, и сигнал можно отсечь от информации для оценки эмоций. Соответственно можно использовать только частоту основного тона, имеющую высокую надежность, что позволяет более успешно оценивать эмоции.

В случае по фиг.5(B) возможно вычислить степень градиента в качестве частоты основного тона в более широком смысле. Предпочтительно взять явную частоту основного тона в качестве информации для оценки эмоций. Далее можно также вычислить «степень расхождения» и (или) «отклонение между линией регрессии и исходными точками» в качестве нерегулярности частоты основного тона.

Предпочтительно взять эту нерегулярность, вычисленную таким образом, в качестве информации для оценки эмоций. Предпочтительно также, как само собой разумеющееся, что явная частота основного тона и ее нерегулярность, вычисленная таким образом, используются для оценки эмоций. В этих процессах будет реализована оценка эмоций, в которой не только частота основного тона в узком смысле, но также характеристики или расхождение частоты голоса отражаются всесторонним образом.

В данном варианте осуществления, кроме того, локальные интервалы гребней (или впадин) вычисляются путем интерполяции дискретных данных автокорреляционного колебания. Поэтому возможно вычислять частоту основного тона с большим разрешением. В результате расхождение частоты основного тона можно обнаруживать более тонко, и становится возможной более точная оценка эмоций.

Кроме того, в этом варианте осуществления степень расхождения частоты основного тона (расхождение, стандартное отклонение и тому подобное) добавляется в качестве информации оценки эмоций. Степень расхождения частоты основного тона показывает уникальную информацию, такую как нестабильность или степень негармонического тона голосового сигнала, что пригодно для обнаружения эмоций, таких как отсутствие уверенности или степень напряжения говорящего. Помимо этого, можно реализовать детектор лжи, обнаруживающий типичные эмоции при произнесении лжи согласно степени напряженности и тому подобного.

Дополнительные предметы варианта осуществления

В приведенном выше варианте осуществления частоты появления гребней или впадин вычисляются как они есть из автокорреляционного колебания. Однако этим изобретение не ограничивается.

Например, конкретные пики (форманты), перемещающиеся со временем, появляются в частотных компонентах голосового сигнала. Кроме того, в автокорреляционном колебании компоненты, отражающие форманты, появляются в дополнение к частоте основного тона. Поэтому предпочтительно, чтобы «компоненты, зависящие от формант», включенных в автокорреляционное колебание, оценивались путем аппроксимации автокорреляционного колебания функцией кривой в положении, не подходящем к мгновенному изменению гребней и впадин. Компоненты (аппроксимированная кривая), оцененные таким образом, вычитаются из автокорреляционного колебания, чтобы вычислить автокорреляционное колебание, в котором ослаблен эффект от формант. За счет такой обработки искажение колебания формантами может быть вырезано из автокорреляционного колебания, посредством чего частота основного тона вычисляется точно и уверенно.

Помимо этого, например, небольшой гребень появляется между гребнем и гребнем автокорреляционного колебания в конкретном голосовом сигнале. Когда этот небольшой гребень ошибочно распознается как гребень автокорреляционного колебания, вычисляется частота половинного основного тона. В этом случае предпочтительно сравнивать высоту гребней в автокорреляционном колебании и рассматривать небольшие гребни как впадина в этом колебании. Согласно этой обработке возможно вычислять точную частоту основного тона.

Предпочтительно также, чтобы регрессионный анализ выполнялся для автокорреляционного колебания, чтобы вычислить линию регрессии, и пиковые точки выше, чем линия регрессии в автокорреляционном колебании, обнаруживаются как гребни автокорреляционного колебания.

В вышеприведенном варианте осуществления оценка эмоций выполняется за счет использования (частоты основного тона, расхождения) в качестве решающей информации. Однако данный вариант осуществления не ограничивается этим. Например, предпочтительно выполнять оценку эмоций с помощью по меньшей мере частоты основного тона в качестве решающей информации. Предпочтительно также выполнять оценку эмоций за счет использования в качестве решающей информации последовательных во времени данных, в которых решающая информация собирается во временных последовательностях. Помимо этого, предпочтительно выполнять оценку эмоций, к которой в качестве решающей информации добавляется тенденция изменения эмоции путем добавления эмоции, оцененной в прошлом. Предпочтительно также реализовать оценку эмоций, к которой добавляется содержание речевого общения путем добавления в качестве решающей информации содержательной информации, полученной путем распознавания речи.

В вышеприведенном варианте осуществления частота основного тона вычисляется посредством регрессионного анализа. Однако данный вариант осуществления не ограничивается этим. Например, интервал между гребнями (или впадинами) автокорреляционного колебания вычисляется как частота основного тона. Или, например, частоты основного тона вычисляются на соответствующих интервалах гребней (или впадин), и выполняется статистическая обработка, принимая это множество частот основного тона в качестве распределения, чтобы принять решение о частоте основного тона и степени ее расхождения.

В вышеприведенном варианте осуществления предпочтительно вычислять частоту основного тона в отношении голоса говорящего и создавать соответствие для оценки эмоций на основе временного расхождения (интонационного расхождения) частоты основного тона.

Изобретатели по настоящему изобретению провели эксперименты по оценке эмоций в отношении к музыкальным композициям, таким как поющий голос или инструментальное исполнение (вид голосового сигнала), за счет использования соответствия, экспериментально созданного из голоса говорящего.

Конкретно, можно получить интонационную информацию, которая отличается от простого изменения тона, путем дискретизации временного изменения частоты основного тона на временных интервалах короче, чем музыкальные ноты. (Участок голосового сигнала для вычисления одной частоты основного тона может быть короче или длиннее, чем музыкальные ноты.)

В качестве другого способа можно получить интонационную информацию, в которой отражается множество нот, за счет выполнения дискретизации на длинном участке голосового сигнала, включающем в себя множество нот, таких как отдельные пункты, для вычисления частоты основного тона.

При оценке эмоций посредством музыкальных композиций обнаружено, что эмоциональная отдача имеет ту же самую тенденцию, что и эмоция, воспринимаемая человеком при прослушивании музыкальной композиции (или эмоция, которую композитор предполагал вложить в музыкальную композицию).

Например, можно обнаружить эмоцию радости / грусти согласно различию ключа, такого как мажорный ключ / минорный ключ. Можно также обнаружить сильную радость в хоровой части при оживлении хорошего темпа. Можно далее обнаружить злость из сильного барабанного боя.

В этом случае соответствие, созданное из речевого голоса, используется как оно есть, и естественно, можно экспериментально создать соответствие, специализированное для музыкальных композиций, при использовании детектора эмоций, который является эксклюзивным для музыкальных композиций.

Соответственно можно оценивать эмоции, представленные в музыкальных композициях, за счет использования детектора эмоций согласно варианту осуществления. При практическом использовании этого детектора можно создать устройство, моделирующее состояние восприятия музыки человеком, или робота, реагирующего соответственно на восторг, злость, печаль или удовольствие, демонстрируемые музыкальными композициями, и тому подобное.

В вышеприведенном варианте осуществления соответствующее эмоциональное состояние оценивается на основе частоты основного тона. Однако изобретение этим не ограничивается. Например, эмоциональное состояние можно оценивать путем добавления по меньшей мере одного указанного ниже параметра:

- 1) изменение частотного спектра в единицу времени;
- 2) цикл флюктуации, время нарастания, время выдержки или время убывания частоты основного тона;
- 3) разность между частотой основного тона, вычисленной из гребней (впадин) в диапазоне нижних частот, и средней частотой основного тона;
- 4) разность между частотой основного тона, вычисленной из гребней (впадин) в диапазоне верхних частот, и средней частотой основного тона;
- 5) разность между частотой основного тона, вычисленной из гребней (впадин) в диапазоне нижних частот, и частотой основного тона, вычисленной из гребней (впадин) в диапазоне верхних частот, или тенденция ее увеличения и уменьшения;
- 6) максимальное значение или минимальное значение интервалов гребней (или впадин);
- 7) число следующих друг за другом гребней (впадин);
- 8) скорость речи;
- 9) значение мощности голосового сигнала или ее изменение во времени;
- 10) состояние частотного диапазона, выделенного в голосовом сигнале из человеческого звукового диапазона.

Соответствие для оценки эмоций можно создать заранее путем связывания частоты основного тона с экспериментальными данными вышеприведенного параметра и эмоционального состояния (злость, радость, напряжение, печаль и тому подобное), декларированного проверяемым. Блок 17 хранения соответствия сохраняет это соответствие. С другой стороны, блок 18 оценки эмоций оценивает эмоциональное состояние путем ссылки на соответствие блока 17 хранения соответствия для частоты основного тона и вышеприведенных параметров, вычисленных из голосового сигнала.

Применения частоты основного тона

Согласно выделению из частоты основного тона эмоциональных элементов из голоса или слуховых явлений (настоящий вариант осуществления) вычисляются частотные характеристики и основные тоны. Помимо этого, можно легко вычислить формантную информацию или информацию мощности на основе изменения на временной оси. Кроме того, можно сделать эту информацию видимой.

- (1) Поскольку состояния флюктуации голоса, слуховых явлений, музыки и тому подобного согласно временным изменениям проявляются за счет выделения частоты основного тона, становятся возможными плавный анализ эмоций и ритма восприятия и анализ тембра голоса или музыки.
- (2) Информация характера изменения во временных изменениях информации, полученных путем анализа основного тона в варианте осуществления, может применяться к видео, действию (выражению или движению), музыке, синтаксису и тому подобному в дополнение к чувствительному общению.
- (3) Можно также выполнять анализ основного тона путем рассмотрения информации, имеющей ритм (именуемой ритмической информацией), такой как видео, действие (выражение или движение), музыка, синтаксис, в качестве голосового сигнала. Помимо этого, возможен анализ характера изменения, касающийся ритмической информации на временной оси. Можно также преобразовать ритмическую информацию в информацию иной формы выражения, делая ритмическую информацию видимой или слышимой на основе результатов этих анализов.
- (4) Можно также применять характер изменения и тому подобное, полученные средствами анализа эмоций, чувствительности, ритмической информации, тембра и т.п., к характеристическому анализу эмоций, чувствительности, психологии и т.п. За счет использования этого результата можно найти характер изменений чувствительности, параметра, порога или тому подобного, который будет общим или взаимосвязанным.
- (5) В качестве вторичного использования можно оценивать психологическое состояние путем оценки психологической информации, такой как духовная сторона, из степени изменения эмоциональных элементов или одновременно обнаруженного состояния различных эмоций. В результате возможны применения к системе управления анализом потребителей товаров, анализу аутентичности личности и тому подобному в финансах или на телефонном узле согласно психологическому состоянию абонентов, пользователей или других групп.
- (6) При вынесении решения об эмоциональных элементах согласно частоте основного тона можно получить элементы для построения модели путем анализа психологических характеристик (эмоций, направленности, предпочтений, мыслей (психологических желаний)), которыми обладают люди. Психологические характеристики людей можно применять к существующим системам, предметам торговли, услугам и моделям бизнеса.
- (7) Как описано выше, в анализе речи по изобретению частоту основного тона можно обнаруживать стабильно и уверенно даже из неразличимого поющего голоса, мурлыкающего пения, инструментального звука и т.п. При применении вышеуказанного можно реализовать систему караоке, в которой можно оценивать точность пения и принимать определенные решения в отношении неразличимого поющего голоса, который трудно было оценивать в прошлом.

Помимо этого, становится возможным сделать видимыми основной тон, интонацию и изменение основного тона путем отображения частоты основного тона или ее изменения на экране. Можно эстетически усваивать точный основной тон, интонацию и изменение основного тона в более короткий период времени при обращении к визуализированному основному тону, интонации или изменению основного тона поющего голоса. Кроме того, можно эстетически усваивать точный основной тон, интонацию и изменение основного тона умелого певца, давая возможность сделать видимыми и имитировать основной тон, интонацию и изменение основного тона умелого певца.

- (8) Поскольку можно обнаружить частоту основного тона из неразличимого мурлыкающего пения или музыки а капелла, которую трудно было обнаружить в прошлом, за счет выполнения анализа речи согласно изобретению можно автоматически формировать стабильно и уверенно музыкальные оценки.
- (9) Анализ речи согласно изобретению можно применить к системе обучения языкам. Конкретно, частоту основного тона можно обнаруживать стабильно и уверенно даже из речевого голоса на незнакомых

иностранных языках, литературном языке и диалекте за счет использования анализа речи согласно изобретению. Система обучения языкам, задающая правильный ритм и произношение иностранных языков, литературного языка и диалекта, может быть создана на базе частоты основного тона.

(10) Помимо этого, анализ речи согласно изобретению можно применять к системе руководства сценарными ролями. То есть частоту основного тона в незнакомых сценарных ролях можно обнаруживать стабильно и уверенно при использовании анализа речи по изобретению. Частота основного тона сравнивается с частотой основного тона умелого актера и благодаря этому устанавливается система руководства сценарными ролями, выполняющая не только руководство сценарными ролями, но также авторскими ремарками.

(11) Можно также применить анализ речи согласно настоящему изобретению к системе тренировки голоса. Конкретно, нестабильность основного тона и неточный способ вокализации обнаруживаются из частоты основного тона голоса, и можно давать совет и тому подобное, благодаря чему система тренировки голоса указывает правильный способ вокализации.

Применения психологического состояния, полученного оценкой эмоций

(1) В общем, результаты оценки психологического состояния можно использовать для продуктов, которые изменяются при обработке в зависимости от психологического состояния. Например, можно устанавливать виртуальные личности (такие как агенты, фигуры) на компьютере, которые изменяют отклики (фигуры, характеристики общения, психологические характеристики, чувствительность, эмоциональный рисунок, рисунок ответвлений в беседе и т.п.) согласно психологическому состоянию другой стороны. Помимо этого, например, можно применять системы, реализующие поиск товарных продуктов, обработку заявок на товарные продукты, работу телефонных узлов, приемные системы, анализ чувствительности потребителей, управление потребителями, игры, пачинко, пачисло, распространение содержания, создание содержания, сетевой поиск, услуги сотовых телефонов, пояснение товарных продуктов, презентационная и обучающая поддержка, в зависимости от гибкости психологического состояния потребителя.

(2) Результаты оценки психологического состояния можно также использовать для продуктов, в общем, увеличивающих точность обработки путем обеспечения психологического состояния в качестве точной информации пользователей. Например, в системе распознавания речи точность распознавания речи можно увеличить за счет выбора словаря, имеющего высокую степень близости в отношении психологического состояния говорящего среди распознанных словарей-кандидатов.

(3) Результаты оценки психологического состояния можно также использовать для продуктов, в общем, увеличивающих защищенность путем оценки незаконных намерений пользователей из психологического состояния. Например, в системе аутентификации пользователей можно увеличить защищенность путем отказа в аутентификации или требования дополнительной аутентификации для пользователей, вызывающих такое психологическое состояние, как тревога или притворство. Далее можно создать повсеместную систему на основе метода аутентификации с высокой защищенностью.

(4) Результаты оценки психологического состояния можно также использовать для продуктов, в общем, в которых психологическое состояние имеет дело с рабочим вводом, например, система, в которой обработка (управление, обработка речи, обработка изображений, тестовая обработка или тому подобное) выполняется путем принятия психологического состояния в качестве рабочего ввода. Помимо этого, можно реализовать систему, поддерживающую создание рассказа, в которой рассказ развивается путем принятия психологического состояния в качестве рабочего ввода и управления движением фигур. Кроме того, можно реализовать систему, поддерживающую создание музыки, выполняющую создание или адаптацию музыки, соответствующие психологическому состоянию, путем принятия психологического состояния в качестве рабочего ввода и изменения конфигурации темперамента, ключей или инструментальной конфигурации. Далее можно реализовать устройство авторских ремарок путем принятия психологического состояния в качестве рабочего ввода и управления окружающей средой, такой как освещение, BGM и тому подобное.

(5) Результаты оценки психологического состояния можно также использовать для устройств, в общем, помогающих в психоанализе, анализе эмоций, анализе чувствительности, анализе характеров или

психологическом анализе.

(6) Результаты оценки психологического состояния можно также использовать для устройств, в общем, выводящих психологическое состояние вовне за счет использования таких выразительных средств, как звук, голос, музыка, запах, цвет, видео, фигуры, вибрация или свет. С помощью такого устройства можно помогать людям в психологической связи.

(7) Результаты оценки психологического состояния можно также использовать для систем связи, в общем, выполняющих передачу информации о психологическом состоянии. Например, можно применять их для чувствительной связи или для чувствительной и эмоциональной резонансной связи.

(8) Результаты оценки психологического состояния можно также использовать для устройств, в общем, принимающих решение о психологическом эффекте (оценивающих психологический эффект), вызываемом у человека содержанием, таким как видео или музыка. Помимо этого, можно создать систему базы данных, в которой можно искать содержание на основе психологического эффекта путем сортировки содержания, касающегося психологического эффекта как предмета.

Возможно также обнаруживать степень возбуждения голоса или эмоциональную тенденцию исполнителя в содержании или инструментального исполнителя путем анализа самого содержания, такого как видео и музыка, тем же самым способом, что и для голосового сигнала. Помимо этого, можно также обнаруживать характеристики содержания путем выполнения распознавания голоса или распознавания сегментации фоном в отношении голоса в содержании. Содержание сортируется согласно таким результатам обнаружения, которые позволяют осуществлять поиск содержания на основе характеристик содержания.

(9) Далее результаты оценки психологического состояния можно также использовать для устройств, в общем, объективно принимающих решение о степени удовлетворенности пользователей при использовании товарного продукта согласно психологическому состоянию. Разработку продукта и создание описаний, которые доступны пользователям, можно легко выполнить с помощью такого устройства.

(10) Помимо этого, результаты оценки психологического состояния можно также применять в следующих областях:

системы поддержки санитарного ухода, консультационные системы, вождение автомобилей, управление транспортными средствами, слежение за состоянием водителя, пользовательский интерфейс, операционная система, робот, аватар, место сетевой торговли, система обучения соответствию, электронное обучение, обучающая система, тренировка манер, система обучения ноу-хау, определение способностей, решение о значащей информации, область искусственного интеллекта, применение к нейронным сетям (в том числе нейрон), стандарты решений и стандарты ветвлений для моделирования или система, требующая вероятностной модели, ввод психологических элементов в моделирование рынка, такое как экономическое или финансовое, анкетирование, анализ эмоций или чувствительности артистов, проверка финансовых кредитов, система управления кредитами, содержание, такое как гадание, носимый компьютер, повсеместные сетевые товары, поддержка интуитивных решений людей, рекламный бизнес, управление домами и залами, фильтрация, поддержка решений для пользователей, управление на кухне, в ванной, туалете и тому подобном, социальные приборы, одежда, взаимосвязанная с волокнами, которые меняют мягкость и способность дышать, виртуальное любимое животное или робот, помогающие в лечении и связи, система планирования, координационная система, система управления поддержкой графика, система поддержки в приготовлении пищи, поддержка музыкального исполнения, видеоэффект DJ, устройство караоке, системы видеопланирования, аутентификация личности, конструирование, устройство моделирования конструирования, система для стимулирования покупательских наклонностей, система управления человеческими ресурсами, аудит, коммерческий поиск виртуальной группы покупателей, система моделирования присяжных / суда, изобразительная тренировка для спорта, искусство, бизнес, стратегия и т.п., поддержка создания памятного содержания об умерших и предках, система или услуга сохранения эмоционального или чувственного рисунка в жизни, служба ориентирования / консьержек, поддержка создания сетевых страниц регистрации, услуга сообщений, будильник, приборы здоровья, массажные инструменты, зубная щетка, медицинские приборы, биоприбор, метод переключения, метод управления, концентратор,

отраслевая система, охлаждающая система, молекулярный компьютер, квантовый компьютер, компьютер фон-Неймановского типа, биочиповый компьютер, система Больцмана, AI управление и нечеткое управление.

Замечание: относительно получения голосового сигнала в шумовой среде

Изобретатели по настоящему изобретению построили измерительную среду с помощью звуконепроницаемой маски, описанной ниже, для того чтобы обнаруживать частоту основного тона голоса в хорошем состоянии даже в шумовой среде.

Сначала в качестве базового материала для звуконепроницаемой маски получают противогаз (SAFETY No. 1880-1, изготовленный TOYOSAFETY). Противогаз делается из резины в части, касающейся и покрывающей рот. Поскольку резина вибрирует согласно окружающему шуму, окружающий шум входит внутрь противогаса. Затем резиновая часть наполняется силиконом (QUICK SILICON, светло серый в жидком виде плотностью 1,3, изготовленный NISSIN RESIN Co, Ltd.), чтобы сделать противогаз тяжелее. Затем пять или более кухонных бумаг и губок наслаиваются друг на друга в воздушном фильтре противогаса, чтобы увеличить уплотняющую способность. В центральной части камеры противогаса в этом состоянии предусматривается путем вталкивания небольшой микрофон. Подготовленная таким образом звуконепроницаемая маска может эффективно подавлять вибрацию от окружающего шума за счет собственной массы силикона и многоуровневой структуры несвязанного материала. В результате небольшое звуконепроницаемое помещение, имеющее форму маски, успешно формируется вблизи рта проверяемого и может подавлять эффект окружающего шума, а также собирать голос проверяемого в хорошем состоянии.

Помимо этого, возможно иметь общение с проверяемым, на которое не влияет так сильно окружающий шум, путем ношения на ушах проверяемого наушников, к которым приняты те же самые меры защиты от звука.

Вышеприведенная звуконепроницаемая маска эффективна для обнаружения частоты основного тона. Однако поскольку уплотнительное пространство звукопоглощающей маски узкое, голос стремится заглушаться. Поэтому она непригодна для частотного анализа или анализа тембров, иного, нежели частота основного тона. Для таких применений предпочтительно, чтобы через звукопоглощающую маску обеспечивался трубопровод, принимающий такую же звукопоглощающую обработку, что и маска, чтобы вентилировать маску извне (воздушная камера) звукопоглощающей среды. В этом случае проверяемый может дышать без каких-либо проблем, и не только рот, но и нос можно закрыть этой маской. Согласно добавлению этого вентиляционного оборудования можно снизить глушение голоса в звукопоглощающей маске. Помимо этого, имеется небольшое неудобство, такое как ощущение затрудненного дыхания для проверяемого, поэтому возможно собирать голос в более естественном состоянии.

Изобретение можно реализовать в разных других формах без отхода от сути или его основных характеристик. Поэтому вышеприведенный вариант осуществления является просто примером в разнообразных объектах, которые не следует интерпретировать в ограничительном смысле. Диапазон изобретения показан в формуле изобретения и не совсем связан с описанием. Помимо этого, различные модификации или изменения, принадлежащие к эквивалентным пределам формулы изобретения, находятся в рамках изобретения.

Промышленная применимость

Как описано выше, изобретение представляет собой метод, который можно использовать для анализатора речи и т.п.

1. Анализатор речи, содержащий:

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.

2. Анализатор речи по п.1, в котором упомянутый автокорреляционный блок вычисляет дискретные данные упомянутого автокорреляционного колебания при дискретном сдвиге упомянутого частотного спектра на частотной оси, и при этом упомянутый блок обнаружения основного тона интерполирует дискретные данные упомянутого автокорреляционного колебания, вычисляет частоты появления одного из локальных гребней и впадин и вычисляет частоту основного тона на основе интервала упомянутых частот появления.

3. Анализатор речи по п.1, в котором упомянутый блок обнаружения основного тона вычисляет множество (порядок появления, частоту появления) по отношению к, по меньшей мере, одному из гребней и впадин автокорреляционного колебания, выполняет регрессионный анализ для упомянутого порядка появления и упомянутых частот появления и вычисляет частоту основного тона на основе градиента линии регрессии.

4. Анализатор речи по п.1, в котором упомянутый блок обнаружения основного тона вычисляет множество (порядок появления, частоту появления) по отношению к, по меньшей мере, одному из гребней и впадин автокорреляционного колебания, исключает отсчеты, флуктуация уровня которых в автокорреляционном колебании мала, из распределения (упомянутого порядка появления, упомянутой частоты появления), выполняет регрессионный анализ по отношению к упомянутому остальному распределению и вычисляет частоту основного тона на основе градиента линии регрессии.

5. Анализатор речи по п.1, в котором упомянутый блок обнаружения основного тона включает в себя: блок выделения, выделяющий «компоненты, зависящие от формант», включенных в упомянутое автокорреляционное колебание путем выполнения аппроксимации кривой к упомянутому автокорреляционному колебанию, и

блок вычитания, вычисляющий автокорреляционное колебание, в котором эффект формант ослаблен за счет исключения упомянутых компонент из упомянутого автокорреляционного колебания, и вычисляет частоту основного тона на основе упомянутого автокорреляционного колебания, в котором эффект формант ослаблен.

6. Анализатор речи по п.1, дополнительно содержащий:

блок хранения соответствия, сохраняющий, по меньшей мере, соответствие между «частотой основного тона» и «эмоциональным состоянием»; и

блок оценки эмоций, оценивающий эмоциональное состояние проверяемого путем соотнесения с упомянутым соответствием для упомянутой частоты основного тона, обнаруженной упомянутым блоком обнаружения основного тона.

7. Анализатор речи по п.3, в котором упомянутый блок обнаружения основного тона вычисляет, по меньшей мере, одно из «степени изменения (упомянутого порядка появления, упомянутой частоты появления) по отношению к упомянутой линии регрессии» и «отклонение между упомянутой линией регрессии и исходными точками» в качестве нерегулярности упомянутой частоты основного тона, и содержит далее:

блок хранения соответствия, сохраняющий по меньшей мере соответствие между «частотой основного тона», а также «нерегулярностью частоты основного тона» и «эмоциональным состоянием»; и

блок оценки эмоций, оценивающий эмоциональное состояние упомянутого проверяемого путем соотнесения соответствия для «частоты основного тона» и «нерегулярности частоты основного тона», вычисленных в упомянутом блоке обнаружения основного тона.

8. Способ анализа речи, содержащий этапы, на которых:

получают голосовой сигнал от проверяемого;

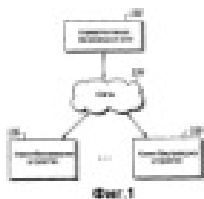
преобразуют упомянутый голосовой сигнал в частотный спектр;

вычисляют автокорреляционное колебание при сдвиге упомянутого частотного спектра на частотной оси; и

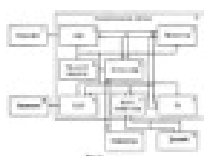
вычисляют частоту основного тона на основе локального интервала между одним из гребней и впадин упомянутого автокорреляционного колебания.

9. Машиночитаемый носитель, содержащий считываемый компьютером код для воплощения функции в качестве анализатора речи по п.1.

Похожие патенты:

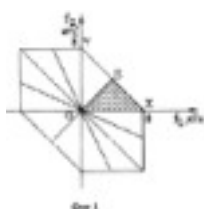


Квантование основного тона для распределенного распознавания речи // 2331932



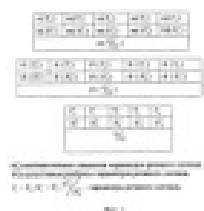
Способ распознавания слов речи // 2296376

Изобретение относится к области анализа и распознавания речевых сигналов. .



Способ распознавания фонем речи и устройство для реализации способа // 2268504

Изобретение относится к области анализа и распознавания речевых сигналов и касается способа распознавания фонем речи. .



Способ автоматической идентификации личности // 2161826

Изобретение относится к обработке информации и может быть использовано в телекоммуникационных системах. .

$$C(d) = \sum_{n=0}^{N-1} S_{\text{avg}}(\ln S_{\text{avg}}(n - d))$$

Оценка периода основного тона // 2421826

Изобретение относится к оценке периода основного тона в аудиосигналах

Оказать финансовую помощь
проекту FindPatent.ru

