ФЕДЕРАЛЬНОЕ АГЕНТСТВО

ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

ГОСТ Р 59879— 2021

Эргономика. Проектирование и применение испытаний речевых технологий

МЕТОДИКА ОПРЕДЕЛЕНИЯ ПОКАЗАТЕЛЕЙ КАЧЕСТВА РАСПОЗНАВАНИЯ ГОЛОСОВЫХ КОМАНД УПРАВЛЕНИЯ

Издание официальное

Москва
Российский институт стандартизации
2021

gost.e-ecolog.ru

Предисловие

1 РАЗРАБОТАН Федеральным государственным унитарным предприятием «Научно-исследовательский институт «Квант» («ФГУП «НИИ «Квант»)

2 ВНЕСЕН Техническим комитетом по стандартизации ТК 164 «Искусственный интеллект»

3 УТВЕРЖДЕН И ВВЕДЕН В ДЕЙСТВИЕ Приказом Федерального агентства по техническому регулированию и метрологии от 25 ноября 2021 г. № 1582-ст

4 ВВЕДЕН ВПЕРВЫЕ

*Правила применения настоящего стандарта установлены в статье 26 Федерального закона от 29 июня 2015 г. № 162-ФЗ «О стандартизации в Российской Федерации». Информация об изменениях к настоящему стандарту публикуется в ежегодном (по состоянию на 1 января текущего года) информационном указателе «Национальные стандарты», а официальный текст изменений и поправок — в ежемесячном информационном указателе «Национальные стандарты». В случае пересмотра (замены) или отмены настоящего стандарта соответствующее уведомление будет опубликовано в ближайшем выпуске ежемесячного информационного указателя «Национальные стандарты». Соответствующая информация, уведомление и тексты размещаются также в информационной системе общего пользования — на официальном сайте Федерального агентства по техническому регулированию и метрологии в сети Интернет (*gost.e-ecolog.ru

© Оформление. ФГБУ «РСТ», 2021

Настоящий стандарт не может быть полностью или частично воспроизведен, тиражирован и распространен в качестве официального издания без разрешения Федерального агентства по техническому регулированию и метрологии

Содержание

1 Область применения 1
2 Термины и определения 1
3 Общее описание системы 2
4 Показатели качества системы распознавания голосовых команд управления 2
5 Методика оценки показателей качества распознавания голосовых команд управления 2
5.1 Общие требования 2
5.2 Виды испытаний 3
5.3 Методика оценки показателя полноты словаря команд управления 3
5.4 Методика оценки показателя ошибки распознавания голосовых команд управления 4
5.5 Методика оценки показателя реального времени распознавания 5
6 Итоговая оценка качества системы распознавания голосовых команд управления 5
Приложение А (обязательное) Требования к тестовым данным 1 6
Приложение Б (обязательное) Требования к тестовым данным 2 7
Приложение В (обязательное) Требования к тестовым данным 3 8
Приложение Г (обязательное) Требования к грамматике системы распознавания голосовых команд управления 9
Приложение Д (обязательное) Пример текстового файла с результатом распознавания голосовой команды 10
Приложение Е (обязательное) Форма протокола испытаний системы распознавания голосовых команд управления 11
Библиография 13

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

Эргономика. Проектирование и применение испытаний речевых технологий

МЕТОДИКА ОПРЕДЕЛЕНИЯ ПОКАЗАТЕЛЕЙ КАЧЕСТВА РАСПОЗНАВАНИЯ ГОЛОСОВЫХ КОМАНД УПРАВЛЕНИЯ

Ergonomics. Design and application of speech technology tests.
Methodology for determining quality indicators of recognition of voice control commands

Дата введения — 2022—03—01

1 Область применения

1.1 В настоящем стандарте установлены методика определения показателей качества распознавания голосовых команд, перечень основных параметров и методы их оценки для определения качества работы систем распознавания голосовых команд управления. Данная методика определяет виды испытаний и последовательность их проведения.

1.2 Настоящий стандарт распространяется на системы распознавания голосовых команд управления.

1.3 Оценки, установленные настоящим стандартом, могут быть получены в автоматизированном (программном) режиме и использованы для сравнения различных систем распознавания голосовых команд управления, а также для формирования рекомендаций по их совершенствованию.

2 Термины и определения

В настоящем стандарте применены следующие термины с соответствующими определениями:

2.1 **голосовая команда:** Речевое сообщение, которое может быть распознано в качестве команды для функционального устройства.

2.2 **словарь (системы распознавания голосовых команд управления):** Перечень всех неповторяющихся команд системы распознавания голосовых команд управления.

2.3 **грамматика (системы распознавания голосовых команд управления):** Набор правил, заданный в определенном формате, определяющий структуру возможных команд в виде линейного перечисления слов или словосочетаний с возможностью подстановки переменных значений.

2.4 **результат распознавания:** Ответ от системы распознавания команд голосового управления в виде текстового представления распознанной команды и при наличии соответствующего ему значения достоверности.

2.5 **достоверность:** Один из выходных параметров системы распознавания голосовых команд управления, который представляет собой число в диапазоне от 0 до 1, характеризующее апостериорную оценку соответствия голосовой команды ответу от системы распознавания.

2.6 **ложный отказ:** Тип ошибки, при которой результат распознавания голосовой команды из словаря содержит истинное значение голосовой команды при значении достоверности, не превышающем заданное пороговое значение, либо результат распознавания не содержит ни одной команды, приведенной в словаре.

2.7 **ложное срабатывание:** Тип ошибки, при которой результат распознавания голосовой команды, отсутствующей в словаре, содержит значение достоверности, превышающее заданное пороговое значение.

2.8 **ошибка спутывания:** Тип ошибки, при которой результат распознавания голосовой команды из словаря содержит ложное значение при значении достоверности, превышающем заданное пороговое значение.

Издание официальное

2.9 **верный результат распознавания:** Результат распознавания голосовой команды из словаря, содержащий истинное значение при значении достоверности, превышающем заданное пороговое значение.

2.10 **показатель реального времени распознавания:** Коэффициент, который определяется как отношение времени, затраченного на распознавание голосовой команды, к длительности анализируемого аудиосигнала, содержащего произнесение данной команды.

2.11 **задержка распознавания голосовой команды управления:** Разность во времени между завершением произнесения команды и получением результата распознавания.

3 Общее описание системы

3.1 Система распознавания голосовых команд управления представляет собой программное обеспечение, реализующее функцию определения соответствия голосовой команды ее текстовому представлению.

3.2 Для решения задачи распознавания голосовых команд управления могут быть использованы системы распознавания речи двух типов: системы, позволяющие распознавать заданные последовательности элементов, и системы, позволяющие распознавать произвольные последовательности элементов. Первый тип — это системы распознавания с заданным словарем; второй тип — системы распознавания слитной речи.

3.3 Вне зависимости от типа системы распознавания голосовых команд управления должна быть обеспечена полнота словаря команд управления.

3.4 Система распознавания голосовых команд управления может обеспечивать возможность изменения словаря за счет использования внешней грамматики в формате EBNF (см. [1]).

3.5 Важными характеристиками системы распознавания голосовых команд управления являются количество ошибок, которое оценивают в соответствии с 5.4, и быстродействие, которое может быть оценено через показатель реального времени распознавания согласно 5.5.

4 Показатели качества системы распознавания голосовых команд управления

Настоящим стандартом нормированы следующие показатели качества системы распознавания голосовых команд:

- полнота словаря команд управления;
- ошибка распознавания голосовых команд;
- показатель реального времени распознавания.

5 Методика оценки показателей качества распознавания голосовых команд управления

5.1 Общие требования

5.1.1 Все испытания проводят с использованием заданного списка команд и тестовых данных, содержание которых не должно быть использовано разработчиками при подготовке проверяемой системы распознавания голосовых команд для исключения возможности фальсификации результатов проведения оценки качества по настоящему стандарту.

5.1.2 Система проходит проверку на тестовых данных трех типов:

- тестовые данные 1 — нормальные условия (речь с нормативным произнесением, без артефактов, все команды из заданного списка команд);
- тестовые данные 2 — усложненные условия (наличие речевых артефактов, все команды из заданного списка команд);
- тестовые данные 3 — условия повышенной сложности (возможно наличие речевых артефактов, все команды отсутствуют в заданном списке команд).

Подробное описание требований к техническим и качественно-содержательным характеристикам тестовых данных приведено в приложениях А—В.

5.1.3 Тестовые данные готовят по заданному списку команд, составленному в виде простого перечисления, или на основе грамматики, подготовленной в соответствии с приложением Г.

5.1.4 Для прохождения испытаний система распознавания голосовых команд управления в качестве обязательного входного параметра должна принимать аудиофайл, формат которого определен в приложении А.

5.1.5 Для систем распознавания, имеющих возможность указания или расширения словаря через указание грамматики или иным способом, такие данные о словаре должны быть загружены до начала испытаний либо переданы в систему в качестве входного параметра в процессе испытаний.

5.1.6 Формой представления результатов распознавания голосовых команд должны быть текстовые файлы, имена которых соответствуют именам аудиофайлов из тестовых данных 1—3, описанных в приложениях А—В. Каждый текстовый файл должен состоять из двух строк, содержащих следующую информацию:

- текстовый результат распознавания;
- достоверность результата распознавания.

Для систем распознавания, не формирующих показатель достоверности, следует записывать значение 1.

При наличии показателей достоверности для каждого отдельного слова команды следует сохранять эти данные в файлах с результатами распознавания и отражать их в протоколе испытаний в виде дополнительных сведений о системе распознавания голосовых команд управления.

В приложении Д приведен пример структуры файла, содержащего результат распознавания голосовой команды управления.

5.1.7 Для корректного сравнения результатов распознавания с эталонными значениями из тестовых данных следует предварительно выполнить нормализацию текстов (преобразование чисел в числительные, расшифровка аббревиатур, учет регистра символов, знаков препинания, пробельных символов и пр.).

5.1.8 Результаты испытаний должны быть оформлены протоколом, содержащим:

- объект испытаний;
- цель испытаний;
- дату проведения испытаний;
- место проведения испытаний;
- материально-техническое обеспечение;
- условия и методику проведения испытаний;
- результаты испытаний;
- выводы и рекомендации.

Пример протокола испытаний приведен в приложении Е.

5.2 Виды испытаний

5.2.1 Испытания выполняются в полном объеме и включают в себя оценку:

- полноты словаря команд управления;
- ошибки распознавания голосовых команд управления;
- показателя реального времени распознавания.

5.2.2 Для показателя полноты словаря команд управления пороговым значением является 1, что соответствует способности системы распознавания голосовых команд достоверно распознавать все голосовые команды управления.

5.2.3 Оценку ошибки распознавания голосовых команд управления для систем распознавания слитной речи осуществляют в соответствии с 5.4.1.

5.2.4 Оценку ошибки распознавания голосовых команд управления для систем распознавания с заданным словарем осуществляют в соответствии с 5.4.2.

5.3 Методика оценки показателя полноты словаря команд управления

5.3.1 Оценку полноты словаря осуществляют на тестовых данных 1, описание которых приведено в приложении А.

5.3.2 Для оценки полноты словаря команд управления необходимо вычислить отношение количества неповторяющихся достоверно распознанных команд к общему количеству неповторяющихся заданных команд. При наличии команд, в которых присутствуют параметры, общее количество команд определяют без учета величины диапазона значений данных параметров. Таким образом, команды «измени громкость радио до 1» и «измени громкость радио до 10» учитывают как одну команду. Полным словарь считают в том случае, если отношение равно 1.

5.3.3 Результаты проверки заносят в строку «Полнота словаря голосовых команд управления» таблицы Е.1 (приложение Е). При фиксировании полного словаря, в строку таблицы заносят: «Полный словарь голосовых команд»; в противном случае: «Отсутствие полного словаря голосовых команд».

5.4 Методика оценки показателя ошибки распознавания голосовых команд управления

Выбор показателя, используемого для оценки ошибки распознавания голосовых команд управления, определен типом системы распознавания. При испытании системы распознавания голосовых команд управления на основе распознавания слитной речи следует применять показатель спутывания, а для системы распознавания с заданным словарем — показатель целевой функции по формуле (2) в 5.4.2. Когда результат распознавания не содержит значение достоверности, следует полагать его равным 1.

5.4.1 Методика оценки показателя ошибки распознавания голосовых команд управления для систем распознавания слитной речи

5.4.1.1 Для измерения ошибки распознавания голосовых команд используют объединенные тестовые данные 1—3.

5.4.1.2 Для получения показателя спутывания следует вычислить редакционное расстояние — метрику, определяющую разность между двумя последовательностями элементов (в данном случае — слов). Такая метрика выражается в минимальном количестве удалений, вставок и замен, необходимых для того, чтобы две последовательности элементов стали равными. В качестве последовательностей элементов используют текстовый результат распознавания, сохраненный в соответствии с 5.1.6, и текст произнесенной голосовой команды, который содержится в тестовых данных. Веса для удалений, вставок и замен принимаются равными 1, для корректных подстановок принимается вес, равный 0.

5.4.1.3 Для оценки ошибки распознавания голосовых команд управления используют показатель спутывания *WER*, вычисляемый по формуле

$$WER = \frac{\sum_{i}(N_{Del} + N_{Ins} + N_{Subst})}{N_{Ref}}, \quad (1)$$

где i — количество файлов в тестовых данных;

N_{Del} — количество удалений;

N_{Ins} — количество вставок;

N_{Subst} — количество замен;

N_{Ref} — общее количество слов в текстах голосовых команд.

5.4.1.4 Полученное значение показателя спутывания заносят в строку «Ошибка распознавания голосовых команд» таблицы Е.1 (приложение Е).

5.4.2 Методика оценки показателя ошибки распознавания голосовых команд управления для систем с заданным словарем

5.4.2.1 Для вычисления целевой функции используют тестовые данные 1—3.

5.4.2.2 Для всех аудиофайлов из тестовых данных 1 и 2 проверяют наличие текстового файла, содержащего результат распознавания. Если для какого-либо аудиофайла отсутствует текстовый файл с результатом, то для него создают одноименный текстовый файл, на первой строке которого содержится пробел, на второй — значение «0».

5.4.2.3 Тестовые данные 1 и 2 используют для определения количества достоверных результатов распознавания, вероятности ошибок спутывания и ложных отказов. Для определения данных показателей следует сравнить текстовые результаты распознавания с текстами произнесенных голосовых команд и полученное в результате распознавания значение достоверности с заданным пороговым значением.

5.4.2.4 Для тех аудиофайлов, для которых отсутствует текстовый файл с результатом распознавания либо результат содержит значение, соответствующее неопределенному результату (такие значения определяет разработчик системы распознавания голосовых команд управления), следует полагать значение результата распознавания команды равным пустой строке со значением достоверности, равным «1».

5.4.2.5 Тестовые данные 3 используют для определения вероятности ложных срабатываний.

5.4.2.6 Для оценки ошибки распознавания голосовых команд управления используют целевую функцию — метрику, которую определяют как взвешенную сумму вероятности ложных отказов, ошибок спутывания и вероятности ложных срабатываний $C_{Primary}$ для установленного порога доверия, вычисляемую по формуле

$$C_{Primary} = \frac{2P_{EMiss}(\theta) + \beta_1 \cdot P_{FalseAlarm}(\theta) + \beta_2 \cdot P_{FalseAlarm}(\theta)}{2}, \quad (2)$$

где $P_{EMiss}(\theta)$ — вероятность ложного отказа, оценивается по тестовым данным 1 и 2, вычисляется как сумма всех верных результатов распознавания, представленных в соответствии с 5.1.6, в которых содержится значение достоверности меньше θ, деленная на суммарное количество аудиофайлов из тестовым данным 1 и 2;

β_1 и β_2 — коэффициенты, вычисляемые по формуле

$$\beta_i = \frac{C_{FalseAlarm}}{C_{Miss}} \cdot \frac{1 - P_{Target-i}}{P_{Target-i}}, \quad (3)$$

$C_{FalseAlarm}$ — вес ложного срабатывания;

C_{Miss} — вес ложного отказа;

$P_{Target-i}$ — априорная вероятность команд;

$P_{FalseAlarm}(\theta)$ — вероятность ложного срабатывания, оцениваемая по тестовым данным 3, вычисляемая как сумма всех результатов распознавания, представленных в соответствии с 5.1.6, в которых содержится значение достоверности больше θ и всех ошибок спутывания, деленная на суммарное количество аудиофайлов из тестовых данных 1 и 2, а также на суммарное количество аудиофайлов из тестовых данных 3.

Веса должны быть посчитаны для β_1 с учетом $P_{Target1} = 0{,}95$ и β_2 с учетом $P_{Target2} = 0{,}6$, а также такие веса, где C_{Miss} и $C_{FalseAlarm}$ могут быть выбраны в диапазоне [0,1]. Рекомендуемое значение весов равно 1. Значение θ выбирают таким образом, чтобы минимизировать значение целевой функции $C_{Primary}$.

5.4.2.7 Полученное значение заносят в строку «Ошибка распознавания голосовых команд» таблицы Е.1 (приложение Е).

5.5 Методика оценки показателя реального времени распознавания

5.5.1 Для вычисления значения показателя реального времени распознавания необходимо использовать тестовые данные 1—3.

5.5.2 Для вычисления значения показателя реального времени распознавания необходимо измерить время работы системы распознавания. За время работы системы распознавания принимают время, прошедшее от момента запуска распознавания первого аудиофайла до момента получения результата распознавания по последнему аудиофайлу из тестового множества. Не допускается параллельная обработка нескольких аудиофайлов, разрешается только последовательная обработка. Время работы системы распознавания голосовых команд должно быть измерено в миллисекундах.

5.5.3 Показатель реального времени распознавания RT вычисляют по формуле

$$RT = \frac{T}{L}, \quad (4)$$

где T — время работы системы распознавания;

L — суммарная длительность аудиофайлов.

5.5.4 Результаты вычисления заносят в строку «Показатель реального времени распознавания» таблицы Е.1 (приложение Е).

6 Итоговая оценка качества системы распознавания голосовых команд управления

6.1 Итоговую оценку качества распознавания голосовых команд управления формируют исходя из оценки полноты словаря команд управления, ошибки распознавания голосовых команд управления и показателя реального времени распознавания.

6.2 Испытания считают пройденными в полном объеме в том случае, если получена итоговая оценка качества распознавания голосовых команд управления и заполнен протокол испытаний по форме, приведенной в приложении Е.

Приложение А
(обязательное)

Требования к тестовым данным 1

Тестовые данные формируют на основании заданного списка команд управления и включают пары аудио- и текстовых файлов, содержащие произнесение и текст голосовой команды соответственно.

Объем тестовых данных определен количеством команд в заданном списке команд управления и количеством дикторов (каждая команда должна быть произнесена каждым диктором).

При формировании тестовых данных должны быть учтены следующие требования:

- все голосовые команды должны быть исполнены на одном языке носителями данного языка;
- количество дикторов должно быть не менее 20: одинаковое количество мужчин и женщин в возрасте от 20 до 50 лет;
- темп речи должен соответствовать средним значениям для выбранного языка (например, для русского языка темп должен составлять от 80 до 120 слов в минуту);
- в речи дикторов должны отсутствовать выраженные дефекты произнесения;
- все аудиозаписи должны быть приведены к единому формату [рекомендуемый формат — Windows PCM (*.wav), 16 кГц, 16 бит];
- длительность пауз не должна превышать 500 мс;
- отношение сигнал/шум аудиофайлов должно соответствовать одному из условий таблицы А.1 и должно быть выбрано исходя из наиболее похожих условий, предполагаемых в качестве стандартных для работы системы распознавания голосовых команд.

Таблица А.1 — Акустические условия записи

Условие	Описание	ОСШ[1], дБ
Офисное акустическое окружение	Офис, т. е. комната, в которой люди работают за письменным столом, обычно или возможно с компьютером	Более 20
Домашнее акустическое окружение	Жилое помещение, т. е. комната с мебелью, местами. Могут присутствовать телевизор или аудиоустройства	От 10 до 20
Общественное место	Внушительного размера помещение (холл) или улица, где обычно присутствует большое количество людей	От 10 до 15
Автомобиль	Салон автомобиля	5—15
Сложные условия	Высокий уровень внешних шумов, например военный полигон, техника, производящая сильный шум	Менее 7
[1]Отношение сигнал/шум.		

Для той команды, которая содержит один или несколько параметров, каждый диктор должен записать по одной реализации. Если диапазон значений параметров команды превышает количество дикторов, каждый диктор фиксирует уникальную реализацию данной команды, в противном случае записывают все возможные реализации, часть которых может быть произнесена несколькими дикторами.

Приложение Б
(обязательное)

Требования к тестовым данным 2

Тестовые данные формируют на основании заданного списка команд управления и включают пары аудио- и текстовых файлов, содержащие произнесение и текст голосовой команды соответственно.

Объем тестовых данных определен количеством команд в заданном списке команд управления и количеством дикторов (каждая команда должна быть произнесена каждым диктором).

При формировании тестовых данных должны быть учтены следующие требования:

- все голосовые команды должны быть исполнены на одном языке (допускается акцентное произнесение);
- количество дикторов должно быть равно 20: 10 мужчин и 10 женщин в возрасте от 20 до 50 лет;
- все аудиозаписи должны быть приведены к единому формату [рекомендуемый формат — Windows PCM (*.wav), 16 кГц, 16 бит];
- отношение сигнал/шум должно составлять от 5 до 25 дБ, но не превышать отношение сигнал/шум тестовых данных 1;
- длительность пауз не должна превышать 500 мс.

В каждой аудиозаписи должно присутствовать не более трех речевых и/или технических артефактов следующих типов:

- клиппирование;
- реверберация;
- крик;
- шепотная речь (отсутствие вокализованных отрезков речи);
- заикание;
- неполное произнесение слов;
- смех;
- речевые заполнители («эээ», «ммм» и т. п.);
- кашель;
- темп речи за пределами диапазона, определенного требованиями к тестовым данным 1 (пониженный либо повышенный).

Некоторые аудиозаписи могут содержать вставки слов, отсутствующих в заданном списке команд управления. Для этого следует предварительно составить перечень таких команд, чтобы дикторы могли их прочесть.

Необходимо обеспечить сбалансированность типов артефактов и их количества в пределах записей одного диктора.

При формировании тестовых данных рекомендуется выбирать условия записи, технические и речевые артефакты с учетом особенностей применения тестируемой системы распознавания голосовых команд управления.

Для последующего анализа результатов и определения типов артефактов, влияющих на качество работы системы распознавания голосовых команд управления, рекомендуется сохранять сведения о соответствии типов артефактов аудиофайлам.

Для той команды, которая содержит один или несколько параметров, каждый диктор должен записать по одной реализации. Если диапазон значений параметров команды превышает количество дикторов, каждый диктор фиксирует уникальную реализацию данной команды, в противном случае записывают все возможные реализации, часть которых может быть произнесена несколькими дикторами.

Приложение В
(обязательное)

Требования к тестовым данным 3

Тестовые данные формируются на основании голосовых команд, отсутствующих в заданном списке команд управления, и включают пары аудио- и текстовых файлов, содержащие произнесение и текст голосовой команды соответственно. Текст команд должен быть заранее подготовлен для предъявления дикторам. Количество команд должно быть равно количеству аудиофайлов в тестовых данных 1, при этом количество слов в команде должно быть равно количеству слов в соответствующей команде из заданного списка команд управления.

При формировании тестовых данных должны быть учтены следующие требования:

- все голосовые команды должны быть исполнены на одном языке (допускается акцентное произнесение);
- необходимо использовать слова, фонетически близкие к словам из заданного списка команд управления, в равном соотношении со словами, фонетически далекими от слов из заданного списка команд управления. Пример фонетически близких слов: «проход» — «пароход», далеких слов: «проход» — «магазин»;
- количество дикторов должно быть равно 20: 10 мужчин и 10 женщин в возрасте от 20 до 50 лет;
- все аудиозаписи должны быть приведены к единому формату [рекомендуемый формат — Windows PCM (*.wav), 16 кГц, 16 бит];
- отношение сигнал/шум должно составлять от 5 до 25 дБ, но не превышать отношение сигнал/шум тестовых данных 1;
- длительность пауз не должна превышать 500 мс.

В речи дикторов могут присутствовать артефакты, описанные в требованиях к тестовым данным 2.

Приложение Г
(обязательное)

Требования к грамматике системы распознавания голосовых команд управления

Грамматика для описания словаря системы распознавания голосовых команд задана в формате EBNF, описанном в [1].

Пример грамматики, определяющий одну команду:

level = "0" | "1" | "2" | "3" | "4" | "5" | "6" | "7" | "8" | "9";

rule = измени громкость радио до level;

grammar = { rule }.

Для проведения испытаний следует составлять грамматику, содержащую команды с параметрами. Если установлена предполагаемая область применения системы распознавания голосовых команд (медицина, спорт, военная или бытовая область и т. д.), то семантическое наполнение грамматики следует выполнять с учетом этой информации.

Приложение Д
(обязательное)

Пример текстового файла с результатом распознавания голосовой команды

Строка 1: уменьши громкость радио.
Строка 2: 0.82 [0.33 0.89 0.99]

Строка 1 содержит текст результата распознавания.

Строка 2 содержит достоверность результата распознавания. При наличии достоверностей распознавания каждого слова их следует указывать в скобках после значения достоверности всего результата распознавания.

Приложение Е
(обязательное)

Форма протокола испытаний системы распознавания голосовых команд управления

Е.1 Объект испытаний:

Система распознавания голосовых команд управления _____
наименование системы

_____

Е.2 Цель испытаний:

Испытания проводились с целью установления работоспособности и качественных характеристик системы распознавания голосовых команд.

Е.3 Дата проведения испытаний: _____
число, месяц, год

Е.4 Место проведения испытаний:_____
наименование организации, ответственной за испытания

Е.5 Материально-техническое обеспечение:

Для проведения испытаний системы распознавания голосовых команд использовались вычислительные средства со следующими характеристиками:

_____
наименование, модель и основные характеристики процессора

_____
объем оперативной памяти и ее основные характеристики

_____
графические ускорители (если использовались)

Е.6 Условия и методика проведения испытаний:

_____
описание тестовых данных и грамматик, использованных в ходе испытаний

_____

_____

_____

_____

_____

Е.7 Результаты испытаний:

Таблица Е.1 — Результаты испытаний системы распознавания голосовых команд

Показатель качества	Полученное значение показателя
Полнота словаря голосовых команд управления	
Ошибка распознавания голосовых команд	
Показатель реального времени распознавания	

Е.8 Дополнительные сведения о системе распознавания голосовых команд управления

_____

_____

Е.9 Выводы и рекомендации

_____

_____

_____

Испытания проводили:

Наименование организации	инициалы, фамилия	_____
Наименование организации	инициалы, фамилия	_____
Наименование организации	инициалы, фамилия	_____
Наименование организации	инициалы, фамилия	_____
Наименование организации	инициалы, фамилия	_____

Библиография

[1] ИСО/МЭК 14977:1996 (ISO/IEC 14977:1996) Информационные технологии. Синтаксический метаязык. Расширенная БНФ [Information technology — Syntactic metalanguage — Extended Backus-Naur Form (BNF)]

УДК 004.896:006.354 ОКС 35.240.99

Ключевые слова: эргономика, испытания речевых технологий, качество распознавания голосовых команд управления

Редактор *Л.С. Зимилова*
Технический редактор *И.Е. Черепкова*
Корректор *Р.А. Ментова*
Компьютерная верстка *Г.Д. Мухиной*

Сдано в набор 29.11.2021. Подписано в печать 28.12.2021. Формат 60×84⅛. Гарнитура Ариал.
Усл. печ. л. 2,32. Уч.-изд. л. 1,90.

Подготовлено на основе электронной версии, предоставленной разработчиком стандарта

Создано в единичном исполнении в ФГБУ «РСТ»
для комплектования Федерального информационного фонда стандартов,
117418 Москва, Нахимовский пр-т, д. 31, к. 2.
gost.e-ecolog.ru gost.e-ecolog.ru