



Инструментальная идентификация говорящего в условиях помех

УДК 004.934.2 : 004.934.8*1

Ю.С. РЫСИН, доцент кафедры “Экология, безопасность жизнедеятельности и электропитание” МГУСИ кандидат технических наук, **А.Н. ТЕРЕХОВ**, доцент кафедры “Общая теория связи” кандидат технических наук, **С.Л. ЯБЛОЧНИКОВ**, зав. кафедрой “Экология, безопасность жизнедеятельности и электропитание” доктор педагогических наук, профессор, кандидат технических наук

Инструментальная идентификация говорящего в условиях помех *One of two speakers instrumental recognition under the impact of noise*

Актуальность результатов исследований, представленных в данной статье, связана с инструментальной идентификацией одного из двух говорящих в условиях воздействия помех, которая может, например, обеспечить оператору связи преимущество, определяя (опережая) желания пользователя во время ведения диалога через call-центр. Или стать определяющим фактором в суде при идентификации личности говорящего. При подготовке статьи выполнен большой объем исследований влияния различного рода шумов и помех на инструментальную идентификацию. В предложенном алгоритме — в отличие от известных — учитывается отношение мощностей формант, что позволяет избавиться идентификацию от абсолютных величин (амплитудной зависимости), а следовательно повысить точность измерений. Выработаны рекомендации и предложены реализации, позволяющие определить оптимальный порог, который обеспечивает разбиение фонограммы на временные интервалы, соответствующие паузам и активным участкам в речевом сигнале. Полученные результаты основаны, в том числе, на исследовании разборчивости сложносоставных числительных и их корреляции с длительностью пауз между их элементами.

Relevance of the research results presented in the following article is defined by the one of two speakers instrumental recognition under the impact of noise, which may provide a line operator with an advantage to identify client's needs during a call centre conversation or become a determining factor in personal identification in a court. A large scope of researches was carried out on the impact of different noises and interferences on the instrumental recognition. Unlike the known algorithms, the proposed one takes formant ratio into consideration, which improves measurement accuracy by relieving identification from absolute values (amplitude dependence). Recommendations and implementations were made to determine the optimal threshold value, which allows sound record to be divided into time slots conforming to pauses and active parts of voice signal. The obtained results are also based on intelligibility of compound numerals and it's correlation with the length of pauses between the elements.

Ключевые слова: относительная мощность формант, инструментальная идентификация, идентификация говорящих, разборчивость сложносоставных числительных, воздействие помех.

Keywords: relative power formant, instrumental identification, identification of speakers, intelligibility compound numerals; interference effect.

Целью данной публикации является описание нового метода идентификации, обладающего лучшими характеристиками. Классическая идентификация говорящего основана на трех основных моментах временных и спектральных характеристик звуков речи:

основная информация гласных и некоторых согласных заключается почти исключительно в формантах (в их расположении по частоте и в соотношениях их амплитуд);

в восприятии согласных главными являются временные изменения амплитуд, изменения частот фор-

мант, т. е. основная информация согласных заключается во временных характеристиках;

характерным признаком всех гласных и большей части согласных, за исключением взрывных звуков, является непрерывность временных изменений частот формант и амплитуд составляющих.

Исходя из того, что все виды артикуляции взаимосвязаны, не имеет принципиального значения, каким из них пользоваться для оценки верности восприятия речи. Однако надо помнить, что только фразы содержат законченную

мысль, и поэтому именно они должны рассматриваться как конечная продукция для оценки [1].

Форманта характеризуется амплитудой A_i , частотой F_i и шириной полосы ΔF_i . Различные звуки имеют разное число формант: гласные — до 4, глухие согласные — до 5 — 6. Наиболее информативны первые три форманты: F_1 , F_2 и F_3 . Наиболее вероятные частоты расположения F_1 — 150 — 900 Гц, F_2 — 550 — 2800 Гц и F_3 — 1500 — 3400 Гц.

Изменение положения формант происходит с частотой 10 — 20 Гц, а их интенсивности — с частотой 20