

ИССЛЕДОВАНИЕ ОЦЕНКИ КАЧЕСТВА РЕЧЕВЫХ СИГНАЛОВ ДЛЯ АВТОМАТИЧЕСКОЙ СИСТЕМЫ ИДЕНТИФИКАЦИИ ДИКТОРОВ

Инкин М.Е. (Университет ИТМО)

Научный руководитель – к.т.н., Столбов М.Б. (Университет ИТМО)

Введение. Работа посвящена исследованию корреляции мер качества речевого сигнала STOI (Short Time Objective Intelligibility) – кратковременная объективная разборчивость и STI (Speech Transmission Index) – индекс передачи речи с оценкой сходства дикторских признаков для одинаковых дикторов (надежность), получаемой с помощью системы автоматической идентификации дикторов (АИД). Основной задачей работы является подготовка методических рекомендаций пользователям системы АИД.

На надежность идентификации влияют условия записи тестовых сигналов: тип устройства звукозаписи, вариативность речи диктора, шумы окружения, каналы связи, акустические свойства помещений и ряд других факторов. В работе исследуется влияние двух факторов – отношения сигнал-шум (ОСШ) и уровня реверберации.

Основная гипотеза – возможность использования объективных мер качества речевых сигналов STOI и STI для предсказания надежности идентификации в различной акустической обстановке.

В работе ставится задача исследования корреляции оценок качества STOI и STI с надёжностью решения АИД в различных условиях проведения тестовой звукозаписи.

Основная часть. Рассмотрены публикации по методам оценки качества и разборчивости речевых сигналов (РС). Исследованы объективные меры качества РС. Для дальнейших исследований выбраны меры STOI и STI. Мера STOI характеризует корреляцию двух речевых сигналов (образцового и тестового), вычисляемую в частотно-временной области [1-3]. Образцовый сигнал считается 100% разборчивым, а для тестового сигнала рассчитывается оценка качества STOI. Мера STI характеризует степень модуляции огибающих сигнала в октавных полосах частот, зависящую как от уровня шума, так и от реверберации [4].

Проведены эксперименты по оценке корреляции меры STOI и оценки сходства, вычисляемой в системе АИД. Выполнены звукозаписи для разных дикторов в разных условиях: помещение, расстояние диктор-микрофон, наличие и уровень шума (речеподобный шум ресторана). Эксперименты показали высокую (до 0,99) корреляцию меры STOI и оценки сходства в зависимости от расстояния до микрофона, ОСШ и уровня реверберации.

Результаты и выводы

Если пользователю АИД известно место будущей записи тестового сигнала, то с помощью STOI пользователь может определить условия (дистанцию и пр.), при которых тестовый сигнал будет обеспечивать надежную идентификацию.

Предполагается создание базы данных оценки STOI в разных помещениях – это позволит предоставить пользователю, которому неизвестно место будущей записи тестового сигнала, рекомендации о типе помещений, дистанции диктор-микрофон, при которых тестовый сигнал будет обеспечивать надежную идентификацию.

В дальнейшем предполагается разработка методических рекомендаций для пользователей АИД: использование мер STOI и STI для проведения предварительной оценки места и условий записи тестового сигнала.

Литература.

1. Таал С. Н. et al. A short-time objective intelligibility measure for time-frequency weighted noisy speech – 2010 IEEE international conference on acoustics, speech and signal processing – IEEE, 2010 – pp. 4214-4217.
2. Таал С. Н. et al. An algorithm for intelligibility prediction of time–frequency weighted noisy speech – IEEE Transactions on Audio, Speech, and Language Processing – 2011 – №. 7. – pp. 2125-2136.
3. Jensen J., Таал С. Н. An algorithm for predicting the intelligibility of speech masked by modulated noise maskers – IEEE/ACM Transactions on Audio, Speech, and Language Processing. – 2016 – №. 11 – pp. 2009-2022.
4. Houtgast T. and Steeneken M. A Review of the MTF Concept in Room Acoustics and Its Use for Estimating Speech Intelligibility in Auditoria – J. Acoust. Soc. Am., vol. 77, No. 3 (1985).