

К. В. Рубанов

(ГГТУ им. П. О. Сухого, Гомель)

МЕТОДОЛОГИЯ СОСТАВЛЕНИЯ ОБУЧАЮЩЕЙ ВЫБОРКИ ДЛЯ АВТОМАТИЧЕСКОГО ИЗОЛИРОВАНИЯ ВОКАЛА ИЗ МУЗЫКАЛЬНОЙ КОМПОЗИЦИИ

Для обучения любой нейронной сети необходимо подготовить обучающую выборку, данные в которой послужат основой для принятия последующих решений обученной моделью. При обучении с учителем, исходный датасет должен включать в себя массив групп признаков для каждого этапа обучения и маску для каждой такой группы.

Составление набора пар из аудиозаписи и её отдельного вокального аудиоряда требует больших трудозатрат, поскольку отдельный от аудиозаписи вокальный аудиоряд найти практически невозможно, но есть возможность выделить а капеллу следующими способами: использование программ iZotope RX7 Editor или Audacity; отделение вокала из песни, при наличии её отдельного инструментального аудиоряда, с помощью программы utagoe; использование аудиозаписей только с вокальным или инструментальным рядом, для первого варианта маской будет сама аудиозапись, для второго – одномерный тензор из нулевых элементов; создание собственных аудиозаписей, случайные комбинируя а капеллы и случайный инструментальный ряд.

Каждый подход имеет свои недостатки, которые можно компенсировать, смешав результаты каждого подхода. Полученные аудиозаписи разбиваются на фреймы, каждый из которых состоит из 513 сэмплов. Фреймы необходимо собрать во фрагменты, каждый из которых содержит 25 фреймов [1]. Технология составления фрагментов состоит в выделении STFT окна, состоящего из 25 фреймов, каждый фрагмент образуется смещением STFT окна на один фрейм, начиная с нулевого и заканчивая n-25-ым фреймом. Таким образом, из n-фреймов можно получить n-24 фрагмента, поскольку значимыми для каждого фрагмента являются только средние фреймы. В качестве ожидаемого ответа сети, следует использовать фрейм из аудиозаписи-маски, соответствующий временному интервалу среднего фрейма исследуемого фрагмента.

Литература

1 Реализация метода разделения речи с помощью глубоких нейронных сетей в режиме реального времени: заявка 14/536,114 Соединённые Штаты Америки / Ш. Кампбелл; заявитель Шэннон Кампбелл; патент. проверенный Вонг и др. – № 2017/0061978; заявл. 07.04.14; опубл. 02.03.17; приоритет 07.11.14, № 14/536,114 (США).