

УДК 004.852

МОДЕЛЬ МАШИННОГО ПЕРЕВОДА С ЯЗЫКА ЖЕСТОВ В ТЕКСТОВЫЙ ФОРМАТ

Данг Хань (Университет ИТМО)

Научный руководитель – профессор, д.т.н Бессмертный Игорь Александрович
(Университет ИТМО)

Введение. Исследования по машинному переводу с помощью нейронных сетей из одного текстового языка на другой, с преобразования аудио дорожки в текст или наоборот, дали много положительных результатов. Актуальность задачи машинного перевода усиливается, когда речь заходит о видео у людей с нарушениями слуха [1], входными данными для которых являются видео без звуковой дорожки, что усложняет задачу.

Основная часть. Для построения модели машинного перевода с языка жестов на текстовый язык необходимо выполнить следующие задачи:

1. Определить точку завершения предложения на языке жестов в видео, которая будет эквивалентна точке в конце предложения в текстовом виде.
2. Вычислить и построить матрицы распределения пауз между словами в тексте [2].
3. Распознать жесты в видео, заполнить пропущенные слова в предложениях, которые не были определены.

Из словаря языка жестов через набор входных данных необходимо построить матрицу распределения вероятностей как отношение между словами в словаре. При выполнении тестирования перевода с видео без звука на языке жестов в текстовый язык, была выдана точность 90%

Выводы. Результаты показали, что предложенный метод машинного перевода позволяет переводить видео без звука с точностью 90% в перевод на текстовый формат. Предложенный метод может быть применен для автоматического перевода видео-материалов на язык жестов у людей с нарушениями слуха, а также для создания систем, позволяющих синхронизировать тексты и видео-информацию у людей с нарушениями слуха. Подобные задачи применяются в жизни, чтобы помочь нам понять фильмы или рассказы, которые были созданы людьми с нарушениями слуха.

Список использованных источников:

1. Bragg D., Koller O., Bellard M., Berke L., Boudreault P., Braffort A., Caselli N., Huenerfauth M., Kacorri H., Verhoef T., et al. Sign language recognition, generation, and translation: An interdisciplinary perspective // In Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility, Pittsburgh, PA, USA, 28–30 October 2019. - pp. 16–31.
2. Wang H., Chai X., Chen X. A Novel Sign Language Recognition Framework Using Hierarchical Grassmann Covariance Matrix // IEEE Transactions on Multimedia. - 2019, - № 21, - pp. 2806–2814.