

УДК 004.41

ПРЕОБРАЗОВАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ В ПОСЛЕДОВАТЕЛЬНОСТЬ С ПРИМЕНЕНИЕМ WFST.

Ткаченко М.А. (Университет ИТМО), Асадулаев А.А. (Университет ИТМО)

Научный руководитель – кандидат технических наук, старший научный сотрудник

Гусарова Н.Ф.

(Университет ИТМО)

Рассматривается применение WSFT для потокового распознавания языка жестов с коррекцией ошибок. Детали решения демонстрируются на примере датасетов поверхностных электромиограмм снятых с мышц рук в процессе показа жестов языка глухонемых.

Введение.

Проблемы распознавания жестовых языков привлекают все больше внимания в исследовательском сообществе. На данный момент существует значительное количество работ, посвященных распознаванию жестов с изображений и видеоматериалов. Также, весьма важным направлением является распознавание жестов с использованием записей или потоковых электромиограмм. На эту тему также существует некоторое количество работ, посвященных распознаванию жестов с локального языка жестов на локальный естественный язык. Существуют работы на эту тему для ирландского, итальянского и части американского (10 символов) алфавита.

В данной работе рассматривается преобразование последовательности сигналов электромиограмм жестового языка в естественный язык с возможностью коррекции ошибок.

Основная часть.

Предлагаемое решение базируется на комбинации подходов машинного обучения и применения WFST.

WFST представляет собой вариацию конечного автомата и, в отличии от традиционного конечного автомата занимается не определением формального языка, а установлением отношения между входной и выходной последовательностью.

В WFST между двумя состояниями может быть несколько переходов и в таком случае, с каждым переходом ассоциирован вес, который отражает вероятность появления конкретной подпоследовательности в последовательности. Таким образом, при использовании данного подхода можно получить на выходе языковую модель, которая не только преобразует входную последовательность в выходную, но еще и предоставляет информацию о схожести выходной последовательности на другие.

Этот подход широко используется, например, в распознавании речи и в моделях машинного обучения занимает то же место что и трансформеры.

Выводы.

Предложен подход для потокового распознавания электромиограмм с возможностью коррекции ошибок с использованием WFST. Продемонстрирована эффективность этого подхода на открытых датасетах различных языков и обозначены дальнейшие шаги в данной работе.

Ткаченко М. А (автор)

Подпись

Асадулаев А.А. (автор)

Подпись

Гусарова Н.Ф. (научный руководитель)

Подпись