

КЛАССИФИКАЦИЯ ЭЛЕМЕНТОВ РЕЧИ ЕСТЕСТВЕННОГО ЯЗЫКА СЕТЬЮ КОХОНЕНА

А.В. Леонов

Нахождение эффективного способа классификации фонетических единиц, являющихся характерными элементами для конкретного естественного языка, при машинном распознавании речи имеет решающее значение при последующем формировании письменной формы. Работа существующих систем распознавания речи для русского языка критична при смене диктора. Поиск методов, обеспечивающих надежное обобщение параметров речевого сигнала, сохраняет свою актуальность. С этой целью было проведено исследование варианта классификации элементов речи с применением сети нейронов Кохонена. В качестве учебного материала использована 15 часовая запись радиоспектакля «Мастер и Маргарита» М.Булгакова. Файлы аудиозаписи были предварительно обработаны утилитой sox (ОС Linux) для получения одноканального сигнала с частотой дискретизации 16000 Гц. В записи спектакля музыкальное сопровождение и специальные эффекты занимают незначительную часть, текст читается разными актерами. С учетом частотного диапазона человеческого голоса и присутствия в записи нескольких женских, и мужских голосов, материал можно считать приемлемым для цели исследования.

Сигнал записи предварительно обрабатывался скользящим окном для выявления наиболее значимых участков и нормализованные значения преобразования Хаара [1] подавались на вход сети Кохонена, состоящей из 500 нейронов. Размер скользящего окна был скорректирован в сторону уменьшения с 1024 до 128 отсчетов, что позволило повысить чувствительность системы к высокочастотным составляющим сигнала. Количество нейронов сети Кохонена подбиралось, начиная с большего значения по соображениям ограниченной вычислительной мощности компьютера, чтобы обеспечить приемлемое время обучения. При использовании ПК с процессором AMD Sempron 2500 суммарное время обучения составило менее 3 часов, общее количество эпох не превысило 8 миллионов.

После завершения обучения осуществлялась проверка обученной сети для отдельных звуков речи русского языка одним диктором на способность к классификации. Запись звуков производилась с использованием микрофона Naiko M22PC и Realtek AC'97 Audio. Звуки произносились отдельно так, как они звучат в слове, а не как при произнесении алфавита. Номера выигрывавших нейронов для рассмотренных звуков приведены в таблицах 1-4.

Открытые гласные получили наибольшее количество признаков в качестве выигравших нейронов, что позволяет не только выполнить классификацию звуков, но при наличии образцов разных дикторов можно выявлять принадлежность голоса. Согласные звуки определяются с меньшим количеством признаков. Это объясняется меньшей чувствительностью испытываемой системы к верхней полосе частот и меньшей мощностью звукового сигнала (уровень записи для всех звуков не менялся).

Таблица 1 – Открытые гласные и начинающиеся с «Й»

| Номера выигравших нейронов для отдельных звуков | | | | | | | | | | |
|---|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| А | О | У | Ы | И | Э | Я | Ё | Ю | Е | Й |
| 3 | 43 | 43 | 107 | 51 | 21 | 51 | 21 | 198 | 51 | 6 |
| 22 | 51 | 124 | 126 | 66 | 51 | 66 | 22 | 258 | 66 | 33 |
| 36 | 66 | 162 | 159 | 107 | 56 | 68 | 51 | 33 | 107 | 107 |
| 51 | 72 | 194 | 162 | 162 | 66 | 93 | 52 | 43 | 124 | 162 |
| 60 | 107 | 198 | 182 | 173 | 72 | 104 | 66 | 51 | 126 | 182 |
| 66 | 109 | 213 | 194 | 182 | 87 | 107 | 100 | 66 | 159 | 194 |
| 68 | 124 | 299 | 198 | 194 | 100 | 124 | 104 | 107 | 162 | 198 |
| 78 | 126 | 307 | 213 | 198 | 107 | 126 | 107 | 159 | 173 | 258 |
| 88 | 154 | 320 | 223 | 204 | 124 | 154 | 109 | 162 | 182 | 299 |
| 89 | 159 | 348 | 258 | 223 | 126 | 158 | 126 | 182 | 198 | 303 |
| 93 | 162 | 350 | 299 | 232 | 133 | 159 | 159 | 194 | 204 | 307 |
| 94 | 173 | 367 | 303 | 258 | 154 | 162 | 162 | 257 | 223 | 314 |
| 100 | 182 | 390 | 348 | 299 | 159 | 172 | 173 | 299 | 258 | 348 |
| 104 | 194 | 407 | 350 | 303 | 162 | 173 | 182 | 303 | 261 | 369 |
| 107 | 198 | 477 | 367 | 320 | 173 | 182 | 191 | 307 | 268 | 373 |
| 128 | 204 | 481 | 373 | 348 | 182 | 198 | 194 | 316 | 272 | 413 |
| 158 | 214 | 496 | 390 | 386 | 198 | 199 | 198 | 320 | 299 | 471 |
| 159 | 223 | 182 | 407 | 407 | 204 | 223 | 204 | 332 | 303 | 481 |
| 162 | 232 | 258 | 471 | 471 | 209 | 227 | 214 | 348 | 307 | 488 |
| 173 | 246 | 488 | 480 | 481 | 223 | 232 | 222 | 350 | 314 | 496 |
| 174 | 258 | | 481 | 496 | 227 | 254 | 223 | 390 | 318 | |
| 182 | 261 | | 496 | | 232 | 258 | 232 | 407 | 327 | |
| 191 | 268 | | | | 246 | 261 | 233 | 453 | 348 | |
| 198 | 269 | | | | 254 | 272 | 252 | 480 | 349 | |
| 204 | 272 | | | | 258 | 283 | 258 | 496 | 386 | |
| 210 | 294 | | | | 261 | 286 | 261 | | 407 | |
| 218 | 294 | | | | 268 | 294 | 268 | | 413 | |
| 222 | 318 | | | | 269 | 299 | 269 | | 454 | |
| 246 | 320 | | | | 272 | 308 | 274 | | 471 | |
| 252 | 348 | | | | 294 | 314 | 281 | | 472 | |
| 254 | 349 | | | | 299 | 318 | 294 | | 481 | |
| 258 | 413 | | | | 300 | 322 | 299 | | 488 | |
| 261 | 453 | | | | 301 | 327 | 308 | | 496 | |
| 268 | 454 | | | | 308 | 348 | 318 | | | |
| 272 | 471 | | | | 314 | 349 | 322 | | | |
| 280 | 472 | | | | 318 | 359 | 348 | | | |
| 281 | 481 | | | | 322 | 360 | 407 | | | |
| 294 | 488 | | | | 324 | 412 | 453 | | | |
| 306 | 497 | | | | 327 | 453 | 454 | | | |
| 314 | | | | | 348 | 454 | 471 | | | |
| 326 | | | | | 349 | 471 | 472 | | | |
| 327 | | | | | 386 | 472 | 488 | | | |
| 348 | | | | | 396 | 481 | 496 | | | |
| 357 | | | | | 400 | 488 | | | | |
| 360 | | | | | 407 | 496 | | | | |
| 374 | | | | | 454 | | | | | |
| 382 | | | | | 466 | | | | | |
| 386 | | | | | 471 | | | | | |
| 397 | | | | | 472 | | | | | |
| 406 | | | | | 473 | | | | | |
| 407 | | | | | 478 | | | | | |

| Номера выигравших нейронов для отдельных звуков | | | | | | | | | | |
|---|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| А | О | У | Ы | И | Э | Я | Ё | Ю | Е | И |
| 412 | | | | | 481 | | | | | |
| 462 | | | | | 488 | | | | | |
| 471 | | | | | 496 | | | | | |
| 472 | | | | | | | | | | |
| 477 | | | | | | | | | | |
| 488 | | | | | | | | | | |
| 496 | | | | | | | | | | |

Таблица 2 – Глухие и звонкие согласные

| Номера выигравших нейронов для отдельных звуков | | | | | | | | | | |
|---|----------|----------|----------|----------|----------|----------|----------|----------|----------|--|
| Б | В | Г | Д | П | Ф | К | Р | М | Н | |
| 24 | 36 | 159 | 162 | 43 | 94 | 94 | 9 | 38 | 162 | |
| 43 | 51 | 162 | 198 | 66 | 107 | 104 | 24 | 51 | 36 | |
| 51 | 107 | 182 | 107 | 194 | 194 | 107 | 36 | 66 | 38 | |
| 68 | 109 | 198 | 159 | 198 | 198 | 158 | 43 | 94 | 51 | |
| 93 | 159 | 258 | 181 | 208 | 208 | 159 | 51 | 154 | 94 | |
| 107 | 162 | 43 | 194 | 246 | 257 | 162 | 56 | 159 | 107 | |
| 126 | 182 | 107 | 208 | 258 | 258 | 182 | 107 | 162 | 173 | |
| 159 | 194 | 173 | 223 | 299 | 350 | 194 | 124 | 173 | 182 | |
| 162 | 198 | 194 | 258 | 307 | 367 | 198 | 159 | 181 | 198 | |
| 194 | 223 | 208 | 299 | 367 | 373 | 208 | 162 | 198 | 208 | |
| 198 | 252 | 257 | 307 | 390 | 390 | 222 | 182 | 204 | 224 | |
| 204 | 258 | 299 | 350 | 454 | 480 | 258 | 191 | 208 | 249 | |
| 249 | 263 | 303 | 367 | 480 | 487 | 299 | 194 | 213 | 257 | |
| 258 | 268 | 316 | 390 | 496 | 488 | 303 | 198 | 232 | 258 | |
| 263 | 272 | 320 | 407 | | 496 | 327 | 204 | 257 | 274 | |
| 281 | 281 | 348 | 480 | Л | | 359 | 223 | 258 | 299 | |
| 299 | 299 | 367 | 487 | 33 | | 367 | 232 | 274 | 303 | |
| 303 | 306 | 407 | 496 | 51 | | 390 | 246 | 299 | 316 | |
| 316 | 316 | 418 | | 162 | | 407 | 257 | 303 | 325 | |
| 318 | 318 | 453 | | 348 | | 412 | 258 | 314 | 327 | |
| 350 | 407 | 471 | | 413 | | 477 | 268 | 325 | 367 | |
| 404 | 436 | 480 | | 496 | | 480 | 272 | 327 | 373 | |
| 488 | 453 | 488 | | 107 | | 488 | 281 | 367 | 407 | |
| 496 | 471 | 496 | | 159 | | 496 | 292 | 373 | 413 | |
| | 472 | | | 182 | | | 306 | 407 | 418 | |
| Т | 480 | | | 194 | | | 307 | 413 | 471 | |
| 107 | 481 | | | 198 | | | 332 | 418 | 488 | |
| 182 | 488 | | | 223 | | | 350 | 453 | 496 | |
| 488 | 496 | | | 258 | | | 367 | 471 | | |
| | | | | 268 | | | 368 | 496 | | |
| | | | | 322 | | | 390 | | | |
| | | | | 436 | | | 393 | | | |
| | | | | 486 | | | 407 | | | |
| | | | | 488 | | | 471 | | | |
| | | | | | | | 477 | | | |
| | | | | | | | 480 | | | |
| | | | | | | | 481 | | | |
| | | | | | | | 487 | | | |
| | | | | | | | 488 | | | |
| | | | | | | | 496 | | | |

Однако это не мешает выполнить их классификацию за исключением звука «Т». Заметно больше признаков обнаружилось для звонких согласных в сравнении с глухими. Малое количество признаков обнаружилось для «Ц», «Ч», «С». Остальные шипящие согласные могут быть классифицированы.

Таблица 3 – Шипящие согласные звонкие и глухие

| Номера выигравших нейронов для отдельных звуков | | | | |
|---|----------|----------|----------|----------|
| З | Ж | Ц | Ш | Щ |
| 107 | 66 | 107 | 63 | 107 |
| 159 | 107 | 488 | 91 | 182 |
| 173 | 159 | | 107 | 327 |
| 182 | 162 | Ч | 110 | 386 |
| 194 | 173 | 107 | 159 | 488 |
| 198 | 182 | 488 | 173 | |
| 258 | 198 | | 182 | Х |
| 299 | 204 | С | 186 | 107 |
| 303 | 223 | 107 | 327 | 182 |
| 488 | 227 | 488 | 359 | 198 |
| 496 | 232 | | 386 | 246 |
| | 233 | | 488 | 471 |
| | 258 | | | 481 |
| | 261 | | | 488 |
| | 269 | | | |
| | 299 | | | |
| | 314 | | | |
| | 348 | | | |
| | 471 | | | |
| | 472 | | | |
| | 488 | | | |

Для мягких согласных звуков «ТЬ», «СЬ», «КЬ», «ХЬ» обнаружено малое количество признаков классификации. В некоторых случаях недостаток признаков может быть преодолен при проверке слов в словаре, но словарь существенно увеличивает объем кода для системы распознавания.

Таблица 4 – Мягкие согласные

| Номера выигравших нейронов для отдельных звуков | | | | | | | |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| Бь | Вь | Дь | Нь | Мь | Фь | Рь | ТЬ |
| 43 | 159 | 159 | 43 | 51 | 107 | 66 | 488 |
| 159 | 162 | 162 | 44 | 107 | 162 | 154 | 107 |
| 162 | 198 | 173 | 51 | 154 | 182 | 162 | 198 |
| 194 | 258 | 182 | 173 | 159 | 191 | 191 | |
| 198 | 272 | 194 | 198 | 162 | 194 | 198 | Зь |
| 258 | 413 | 198 | 213 | 173 | 198 | 258 | 107 |
| 299 | 453 | 258 | 258 | 182 | 204 | 314 | 159 |
| 303 | 496 | 299 | 299 | 194 | 208 | 349 | 162 |
| 307 | | 320 | 303 | 198 | 258 | 488 | 173 |
| 316 | Гь | 488 | 320 | 213 | 299 | 496 | 198 |
| 348 | 107 | 496 | 407 | 232 | 303 | | 258 |
| 413 | 162 | | 413 | 258 | 350 | Ль | 488 |
| 480 | 258 | Пь | 496 | 272 | 386 | 51 | |
| 488 | 51 | 173 | 107 | 299 | 390 | 66 | Сь |
| 496 | 159 | 182 | 159 | 303 | 454 | 162 | 198 |
| | 198 | 198 | 162 | 316 | 480 | 299 | 258 |

| Номера выигравших нейронов для отдельных звуков | | | | | | | |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| Бь | Вь | Дь | Нь | Мь | Фь | Рь | Ть |
| Кь | 299 | 258 | 182 | 318 | 488 | 496 | 471 |
| 107 | 316 | 327 | 194 | 407 | 496 | 107 | 488 |
| | 348 | 386 | 272 | 413 | | 198 | |
| Хь | 496 | 407 | 318 | 471 | | 258 | |
| 107 | | 488 | 472 | 481 | | 471 | |
| 488 | | | | 496 | | 488 | |

В целом некоторые имеющиеся недостатки классификации преодолимы. Полученная в результате обучения система классификации элементов речи может быть использована при создании программного обеспечения распознавания речевого сигнала, в том числе может являться прототипом для систем независимых от голоса диктора.

Список использованных источников

1. Леонов А.В. Оконный метод сегментации речевого сигнала // Вестник ВГТУ, 2005, вып. 7, с. 122-124.
2. Rabiner L.R., Schafer R.W. Digital processing of speech signals. – Prentice Hall Inc., 1978. – 265 p.

SUMMARY

The are offered model and the creation technology speech elements classification of natural language, including independent of an voice announcers.