

УДК 004.93+57.087.1

**ВЫБОР ПРИЗНАКОВ В ЗАДАЧЕ РАСПОЗНАВАНИЯ ЭМОЦИЙ ДИКТОРА**

**А.В. Шолохов**

Исследуется подход к выбору наиболее информативных признаков в задаче автоматического определения эмоционального состояния человека по записям устной речи. Для получения оптимального подмножества признаков по заданному критерию качества распознавания использовался генетический алгоритм.

**Ключевые слова:** распознавание эмоций, речевые признаки, генетические алгоритмы.

Важность проблемы автоматического распознавания эмоций объясняется тем, что их выражение в значительной степени улучшает взаимопонимание у людей при общении друг с другом. Знания эмоциональной окраски высказывания позволяют извлечь более полную информацию, содержащуюся в речевом сигнале. Одна из основных проблем в распознавании эмоций заключается в том, что неизвестно, какие именно речевые характеристики и их подмножества отвечают за проявление той или иной эмоции. В связи с этим важным этапом предобработки является процесс извлечения из большого объема тех данных, которые действительно содержат в себе необходимую информацию. Кроме того, наличие шумовых нерелевантных признаков может снизить устойчивость классификатора. Таким образом, из множества всех извлеченных признаков необходимо выбрать подмножество, обучившись на котором, классификатор покажет наилучшую оценку качества распознавания.

Так как количество возможных решений экспоненциально зависит от размерности пространства признаков, было решено использовать генетический алгоритм для поиска оптимального решения. Целевая функция представляла собой оценку вероятности правильного распознавания со слагаемым, штрафующим наборы из большого количества признаков.

Эксперименты проводились на Берлинской базе записей эмоциональной речи EMO-DB [1]. Для каждой аудиозаписи формировались 400-мерные вектора признаков, которые и составили обучающую выборку. В качестве классификатора использовался метод опорных векторов (Support Vector Machine) [2]. В таблице представлены полученные в результате эксперимента оценки вероятности безошибочной классификации эмоций. В левой части ячеек (столбцы 1) показаны оценки для исходного набора признаков, а в правой (столбцы 2) – для отобранных генетическим алгоритмом, количество которых в 10 раз меньше. Значение в каждой ячейке соответствует отдельно обученному бинарному классификатору. Как видно, в некоторых случаях качество распознавания даже повысилось на сокращенных наборах признаков.

Эмоции	Скука		Отвращение		Страх		Радость		Грусть		Нейтральное	
	1	2	1	2	1	2	1	2	1	2	1	2
Гнев	0,95	0,96	0,87	0,89	0,89	0,87	0,75	0,79	0,98	0,97	0,98	0,99
Скука	–		0,85	0,79	0,86	0,79	0,88	0,88	0,78	0,73	0,72	0,67
Отвращение	–		–		0,72	0,68	0,76	0,76	0,89	0,91	0,88	0,88
Страх	–		–		–		0,76	0,77	0,90	0,88	0,89	0,85
Радость	–		–		–		–		0,94	0,95	0,92	0,88
Грусть	–		–		–		–		–		0,83	0,84

Таблица. Оценка вероятности безошибочной классификации эмоций для исходного набора речевых признаков (столбцы с индексом 1) и набора признаков, отобранных генетическим алгоритмом (столбцы с индексом 2)

Эксперименты показали, что генетический алгоритм является эффективным средством для решения задачи выбора подмножества информативных признаков. В дальнейшем планируется исследовать адаптивные генетические алгоритмы, параметры которых могут подстраиваться в ходе работы, тем самым увеличивая скорость поиска решения и увеличивая вероятность выхода из локальных оптимумов.

1. Берлинская база эмоциональной речи [Электронный ресурс]. – Режим доступа: <http://pascal.kgw.tu-berlin.de/emodb/>, свободный. Яз. англ. (дата обращения 09.02.2012).
2. Cortes C., Vapnik V. Support vector networks // Machine Learning. – 1995. – V. 20. – P. 1–25.

**Шолохов Алексей Владимирович** – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, студент, sholokhovalexey@gmail.com