

НЕЙРОСЕТЕВОЙ КВАНТОВАТЕЛЬ АУДИОКОДЕРА

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Аврамов В.В.

Герасимович В.Ю. – асс. каф. ЭВС, Петровский А.А. – д.т.н., профессор

Одной из задач современных аппаратных и программных мультимедийных комплексов является доставка пользователю максимально сжатого медиа-контента высокого качества, и вместе с этим встает вопрос о хорошо сжатом, качественном аудио. Несмотря на существующее многообразие методов компрессии и кодирования аудио по-прежнему актуальны исследования, по результатам которых возможно повысить эффективность известных подходов к компрессии цифровых аудио-сигналов или получить новые алгоритмы компрессии и кодирования. В настоящее время одним из наиболее перспективных направлений таких исследований являются нейронные сети.

Для таких задач как кодирование данных и уменьшения их размерности в теории искусственных нейронных сетей существует специальная архитектура – нейронная сеть автоассоциатора (автокодировщик) [1,2]. Архитектура такой сети представляет собой нейронную сеть прямого распространения с одним скрытым слоем, размерность входного слоя которой равна размерности выходного слоя. Основной задачей обучения нейронной сети автоассоциатора является получение на выходе значений наиболее близких ко входным.

Как показано на рисунке 1, нейросетевой квантователь для коэффициентов преобразователя аудиоречевого кодера [4] представлен многослойным автокодировщиком. Данная модель получена путем объединения трех заранее обученных RBM (Restricted Boltzmann machine) [3]. После объединения всех слоев, нейронная сеть дообучалась по методу обратного распространения ошибки [1].

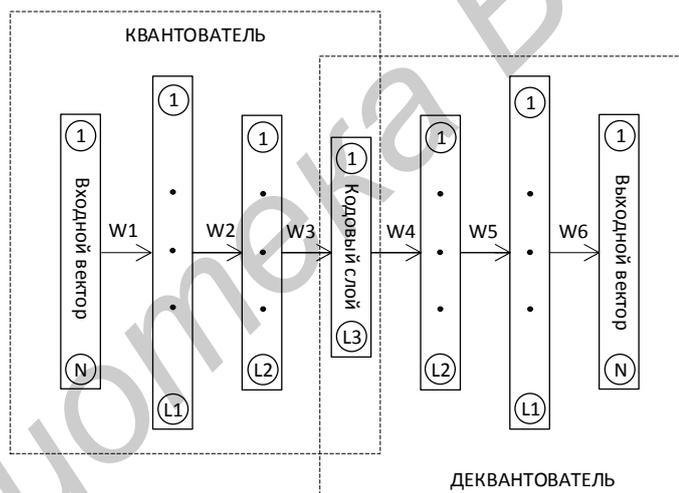


Рис. 1 - Структура нейросетевого квантователя коэффициентов аудиокодера

Как показано на схеме, квантователь представлен трехслойной нейронной сетью прямого распространения. На вход квантователя поступают векторы коэффициентов аудиокодера, на выходе кодового слоя формируется двоичный код. Деквантователь представлен симметричной квантователю нейронной сетью. На вход деквантователя поступают двоичные коды с выходов квантователя, которые в процессе обработки преобразуются обратно в векторы коэффициентов аудиокодера.

Основным преимуществом нейросетевого подхода к задаче квантования коэффициентов аудиокодера является высокая степень компрессии без существенной потери качества сигнала. Основным недостатком нейросетевого подхода – высокая вычислительная сложность.

Несмотря на высокую вычислительную сложность, нейронные сети остаются одним из наиболее перспективных направлений исследований в кодировании и компрессии информации.

Список использованных источников:

1. Хайкин С. Нейронные сети: полный курс. 2 издание. – М.: Издательский дом «Вильямс», 2006 г. – 1104 с.
2. Кохонен Т. Ассоциативные запоминающие устройства. – М.: Мир, 1982. – 384 с.
3. Hinton G. E., Salakhutdinov R. R. Reducing the Dimensionality of Data with Neural Networks // Science, 313(5786). – 2006. – pp. 504–507.
4. Петровский Ал.А., Петровский А.А. Масштабируемые аудиоречевые кодеры на основе адаптивного частотно-временного анализа звуковых сигналов //Труды СПИИРАН, 1(50). – 2017. – с. 56-92.