КОМПРЕССИЯ АУДИОДАННЫХ С ПОМОЩЬЮ ГРАФИЧЕСКОГО ПРЕОБРАЗОВАНИЯ

А.В. Соловьёва, Ю. А. Шейко
Научный руководитель – Ролич О.Ч.
к-т техн. наук, доцент
Белорусский государственный университет информатики и радиоэлектроники

Графическое преобразование. Получив блок аудиосигнала с размером кадра N, мы можем создать график G={V,E,s}, где V и E - это вершины и края графика, и $s\epsilon R^{N\times 1}$ это аудиосигнал, для которого матрица графика определяется как $K\epsilon R^{N\times N}$. Для данного графика элементы матрицы примыкания W получены следующим образом [1],

$$W_{ij} = \begin{cases} w_{ij}.\ if(i,j) \in E \\ 0.\ otherwise \end{cases} \quad (1)$$

где w_{ij} – вес ребра между i и j на графике. Матрица степеней $D\epsilon R^{N\times N}$ - диагональная матрица, для которой элементы определены следующим образом,

$$D_{ij} = \begin{cases} \sum w_{ij}\ .\ if\ i = j \\ 0.\ otherwise \end{cases} \quad (2)$$

Тогда матрица Кирхгофа K будет определена следующим образом,

$$K = D - W \quad (3)$$

где оператор K также известен как оператор Кирхгофа, как дань уважения Густаву Кирхгофа за эти исследования и достижения в электрических сетях. Кирхгофа назвал матрицу примыкания W матрицей проводимости. Матрица K была бы настоящей симметричной, и, основываясь на спектральной теории, собственное значение разложения (EVD) этой матрицы приводило бы к набору реальных неотрицательных собственных значений, обозначаемых Λ = $\{\lambda_1 \dots \lambda_2\}$, и набор соответствующих независимых (далее - ортогональных) собственных векторов V = $\{\nu_1 \dots \nu_2\}$, полученных как

$$K = V\Lambda V^T \quad (4)$$

Затем мы можем использовать эти ортогональные собственные векторы для декорреляции сигнала, заданного на графике, т.е.

$$c = V^T s \quad (5)$$

где $c\epsilon R^{N\times 1}$ – приблизительная матрица коэффициентов разреженного преобразования.

Аудио компрессия с помощью графического преобразования (ГТ). Чтобы применить ГТ к входному аудиосигналу, мы ввели две графические структуры, как показано на рисунках 1 и 2. Учитывая тот факт, что близлежащие образцы аудиосигналов будут сильно коррелировать, мы предполагаем, что любой образец в кадре будет одним узлом на графике, который соединяется с соседними образцами через края. Для простоты объяснения объясняемая структура показана на рисунке 1, например, фрейм с 8 образцами длины.

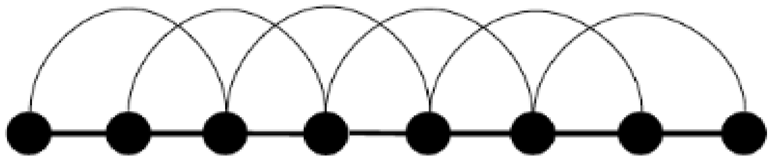

Рисунок 1 - Графическая структура 1

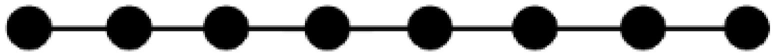

Рисунок 2 - Графическая структура 2

Матрицы сопряженности могут быть определены следующим образом:

$$W_1 = \begin{bmatrix} 0 & 1 & 0.1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0.1 & 0 & 0 & 0 & 0 \\ 0.1 & 1 & 0 & 1 & 0.1 & 0 & 0 & 0 \\ 0 & 0.1 & 1 & 0 & 1 & 0.1 & 0 & 0 \\ 0 & 0 & 0.1 & 1 & 0 & 1 & 0.1 & 0 \\ 0 & 0 & 0 & 0.1 & 1 & 0 & 1 & 0.1 \\ 0 & 0 & 0 & 0 & 0.1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0.1 & 1 & 0 \end{bmatrix}$$

$$W_2 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Рисунок 3 - Матрицы сопряженности

Мы использовали первую структуру для ГТ-1 и вторую для ГТ-2 метода. Как показано на рисунке 1, вес ребер между образцом и его первым соседом больше, чем вес второго соседа.

Для получения базисной матрицы преобразования на основе графов мы предложили две графические структуры для аудиосигналов. Затем мы использовали базовую матрицу ГТ для декорреляции данных и сжатия звука. Наш метод имеет лучшую производительность, чем другие методы, основанные на трансформации, такие как DCT и FWHT.

Библиографический список

1. Rao K. R. The transform and data compression handbook / K. R. Rao, P. C. Yip // CRC press, 2000. — Vol.35-59.