

АЛГОРИТМ СТАТИСТИЧЕСКОГО АНАЛИЗА ДАННЫХ

студент гр. 914301 Пурькова М. В.

Научный руководитель - канд. техн. наук Ролич О. Ч.

Белорусский государственный университет информатики и радиоэлектроники
Минск, Беларусь

Статистический анализ, наряду со спектральным, корреляционным, кепстральным, мел-кепстральным, вейвлет и другими видами, является одним из наиболее популярных инструментов технического анализа больших данных, в том числе, их потоков в масштабе реального времени [1].

В анализе потока данных формируется выборка, как правило, в виде массива некоторой заданной длины N , которая преобразуется или подвергается первичному анализу. Затем из потока выхватывается очередная выборка, и процесс повторяется. Множество преобразованных периодических выборок образуют аналитический ансамбль, на основании которого проводится окончательный анализ, а технической системой принимаются конкретные решения.

В спектральном анализе в целях повышения разрешения и контрастности спектрограммы, представляющей спектральный аналитический ансамбль, выборка умножается на окно, в качестве которого чаще всего выступают прямоугольное окно, окно Хэмминга, Ханна (или Хеннинга), Блэкмана [2].

Проводя аналогию со спектральным анализом, в статистическом анализе понятия «окна» не существует. Это означает, что на практике применяется только прямоугольное окно. Поэтому, с целью унификации алгоритмов статистического и спектрального анализа как наиболее популярных инструментов в данной работе предлагается использовать «окно» и в статистическом анализе.

С учётом предложения использования «окна» (далее, корректирующего окна) алгоритм вычисления двухмерной гистограммы распределения плотности уровней квантования в потоке данных будет выглядеть так, как показано на рисунке 1. Отдельные операторы алгоритма рисунка 1 и синтаксис выражений ориентированы на Си-подобные языки программирования.

В алгоритме, не снижая общности рассуждений, моделируемый поток данных ограничен и представлен массивом $M[Q]$ достаточно большой, от нескольких десятков до сотен тысяч элементов, длины Q . Кроме массива $M[]$, моделирующего поток данных, в качестве входных параметров алгоритма выступают:

m_x и m_n – максимальное и минимальное значения массива $M[]$, определяющие диапазон изменения уровня квантования;

N – длина массива $f[]$ выборки;

$step$ – шаг выборок, определяющий их степень перекрытия, $0 < step \leq N$;

$win[N]$ – корректирующее окно как массив длиной N ;

QH – количество уровней квантования, определяющее длину массива одномерной гистограммы их плотности;

$R[256]$, $G[256]$, $B[256]$ – цветовая палитра.

Алгоритм возвращает цветное изображение $S[3][K][QH]$ двухмерной гистограммы с учётом стандартной палитры [3]. Три слоя массива $S[][][]$ определяют соответственно красный (первый слой, индекс «0» в 11-м блоке алгоритма), зелёный (второй слой, индекс «1» в 11-м блоке алгоритма) и синий (третий слой, индекс «2» в 11-м блоке алгоритма) компоненты в виде двухмерных матриц беззнаковых однобайтных целых значений в цветовой системе RGB888.

Действие корректирующего окна на входную выборку отражено в блоке 5 через умножение потоковых данных на элементы массива-окна $win[]$.

В блоках 6 и 9 вызываются стандартные функции $hist()$ вычисления классической

одномерной гистограммы плотности уровней квантования и $\max()$ определения максимального значения в массиве $h2D[K][QH]$ двумерной гистограммы без учёта цветовой палитры.

Результаты применения предложенного алгоритма к потокам оцифрованных данных внутренних шумов MEMS-акселерометра отражены на рисунках 2, (а) – 2, (е).

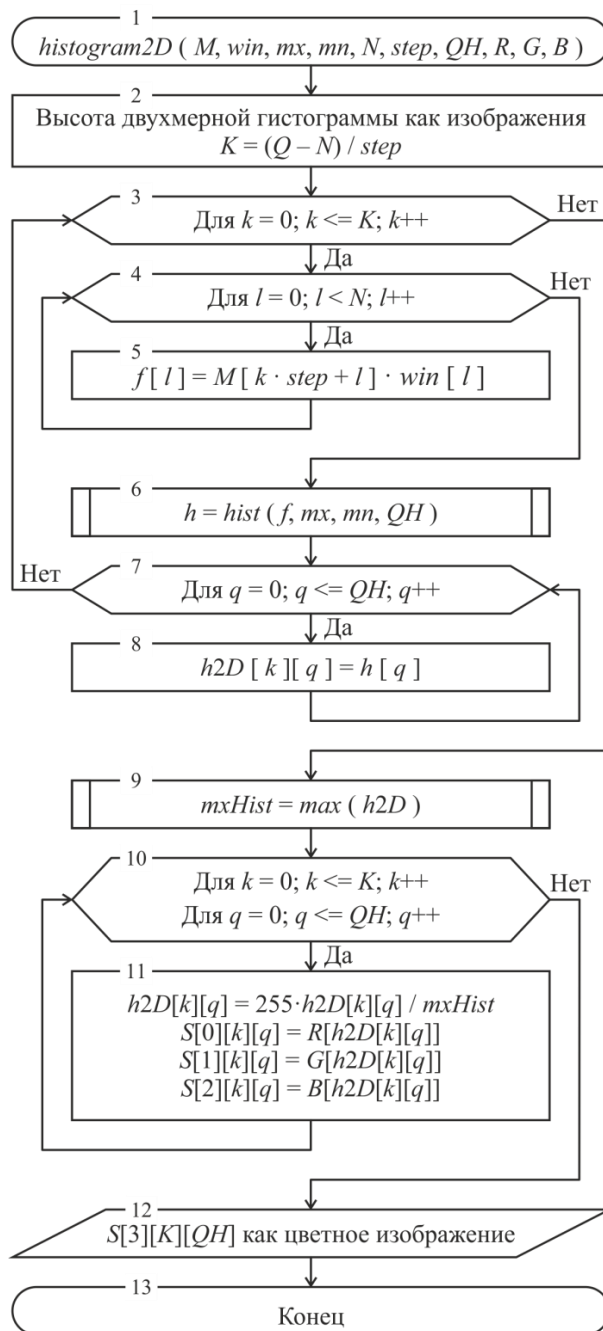


Рис 1. Схема алгоритма вычисления двумерной гистограммы плотности уровней квантования.

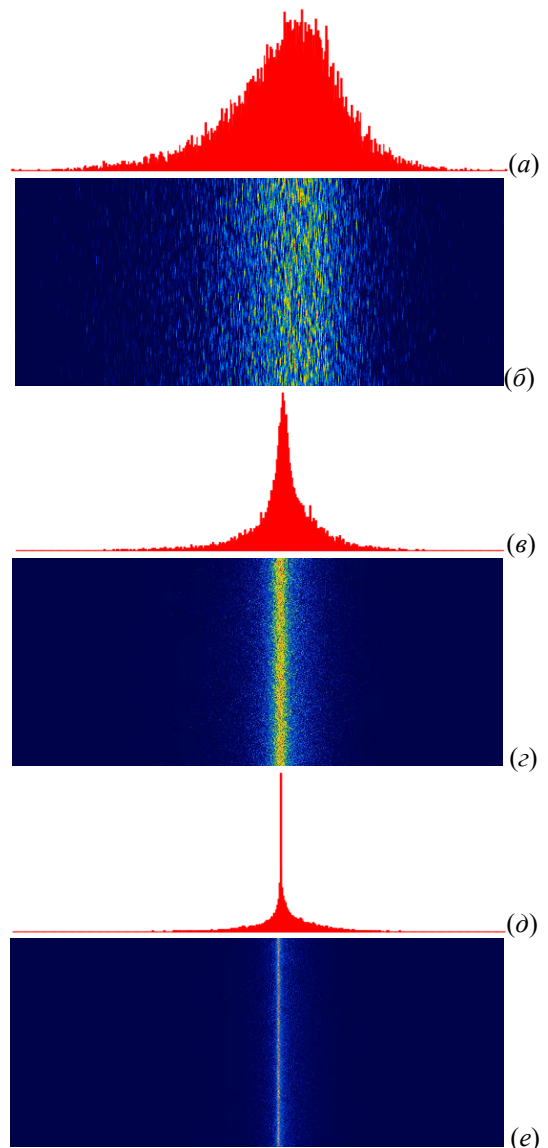


Рис 2. Гистограммы плотности шума акселерометра:

- (а) – одномерная с прямоугольным окном;
- (б) – двумерная с прямоугольным окном;
- (в) – одномерная с окном Хэмминга;
- (г) – двумерная с окном Хэмминга;
- (д) – одномерная с окном Хеннинга;
- (е) – двумерная с окном Хеннинга.

Исследуемый сигнал акселерометра представляет шум с асимметричным законом распределения. Знак асимметрии, в отличие от вида распределения, не зависит от корректирующего окна. Из двумерных гистограмм рисунков 2, (б), (з), (е) видно, что процесс шума стационарный. Но гистограммы 2, (д) и 2, (е), построенные на базе окна Хеннинга, имеют меньшую ширину и большую контрастность, нежели гистограммы с применением иных, представленных на рисунках 2, (а) – 2, (з) корректирующих окон.

В целом, в работе, кроме прямоугольного, исследованы результаты действия наиболее распространённых корректирующих окон, таких как окно Хэмминга $wHm_l = 0.53836 - 0.46164 \cdot \cos\left(\frac{2\pi \cdot l}{N-1}\right)$, Хеннинга $wHn_l = 0.5 \cdot \left(1 - \cos\left(\frac{2\pi \cdot l}{N-1}\right)\right)$ и Блэкмана $wB_l = \frac{1-\alpha}{2} - \frac{1}{2} \cdot \cos\left(\frac{2\pi \cdot l}{N-1}\right) + \frac{\alpha}{2} \cdot \cos\left(\frac{4\pi \cdot l}{N-1}\right)$, $\alpha = 0.16$, $l = 0, 1, \dots, (N-1)$ (см. также блок 4 рисунка 1). Наилучшими показателями по разрешению и контрастности двумерных гистограмм обладают окна Хеннинга и Блэкмана.

Предложенный алгоритм найдёт применение в задачах потокового анализа данных на основе совместного спектрально-статистического инструментария. Он позволит унифицировать и упростить алгоритм анализа с точки зрения программы, где не требуется дополнительной проверки типа анализа (статистического или спектрального) на каждой итерации, особенно при необходимости детектирования нестационарных процессов.

Литература

1. Зуев, В. А. Методика статистического анализа характеристик входных потоков запросов в системах обработки информации / В. А. Зуев, А. Н. Панфилов, А. Н. Скоба // Электронный научный журнал «Инженерный вестник Дона». – 2015. – № 2. – [Электронный ресурс]. – 2020. – Режим доступа: http://www.ivdon.ru/uploads/article/pdf/IVD_44_zuev.pdf_737bf9b28b.pdf.
2. Волосюк, В. К. Использование новых окон Кравченко при гармоническом анализе методом дискретного преобразования Фурье / В. К. Волосюк, В. В. Павликов, В. Ю. Севостьянов // Радіоелектронні і комп'ютерні системи. – 2007. – № 1 (20). – С. 5 – 11. – [Электронный ресурс]. – 2020. – Режим доступа: http://irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?C21COM=2&I21DBN=UJRN&P21DBN=UJRN&IMAGE_FILE_DOWNLOAD=1&Image_file_name=PDF/recs_2007_1_3.pdf.
3. Ролич, О.Ч. Тепловой контроль работы двигателя на основе статистического анализа сигналов болометра / О.Ч. Ролич, В.Е. Тарасенко, В.С. Ивашко // Изобретатель. – 2019. – № 2-3 – С. 40 – 44.

УДК 338.27

ПРИМЕНЕНИЕ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ В БИЗНЕСЕ

студент гр. 914301 Соломкина Ю. С.

Научный руководитель - канд. техн. наук Ролич О. Ч.

Белорусский государственный университет информатики и радиоэлектроники
Минск, Беларусь

Экономика современных предприятий не может полноценно функционировать без информации [1]. На сегодняшний день бизнес-процессы в информационных компаниях часто проходят без участия людей. Интеллектуальные машины самостоятельно следят за исполнением контрактов, контролируют производство, логистику, экономя при этом значительные финансовые суммы. Так работают интеллектуальные компании как