ЦОС ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ

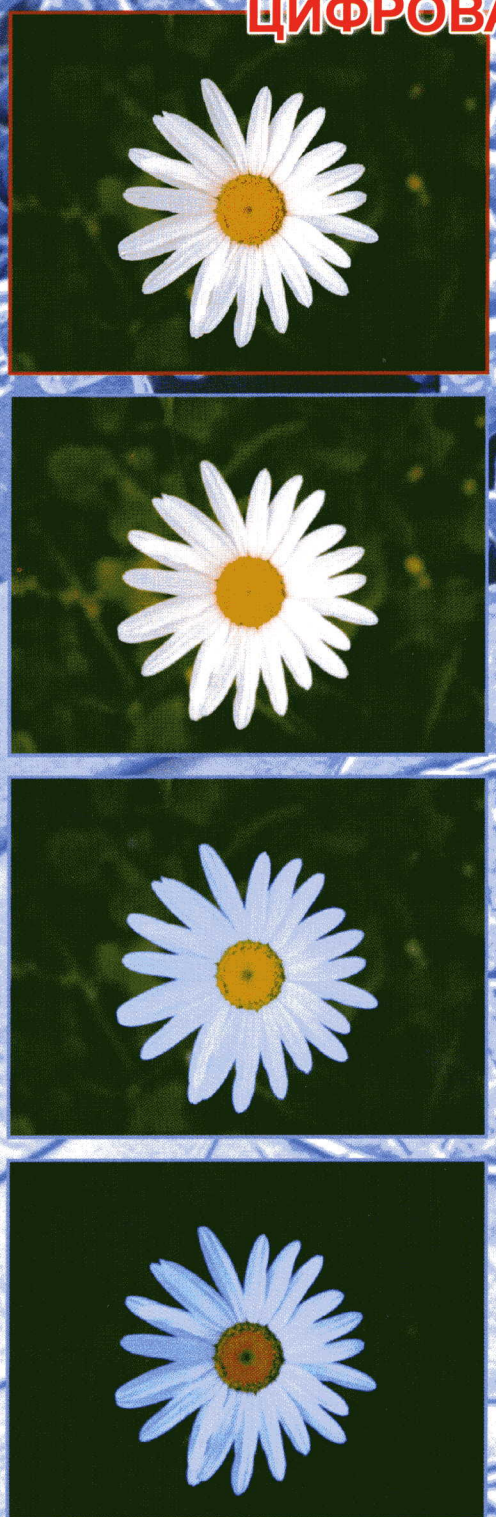
4/2008

НАУЧНО - ТЕХНИЧЕСКИЙ ЖУРНАЛ

ЦИФРОВАЯ ОБРАБОТКА ИЗОБРАЖЕНИЙ

тематический выпуск

- ТЕХНИЧЕСКОЕ ЗРЕНИЕ
- ОБРАБОТКА ВИДЕО
- КОДИРОВАНИЕ И СЖАТИЕ
- ДВУМЕРНАЯ ФИЛЬТРАЦИЯ
- ФРАКТАЛЬНАЯ ОБРАБОТКА



ISSN 1684-2634

9 771684 263005 >

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ

Научно-технический журнал
№ 4/2008

Издается с 1999 года
Выходит четыре раза в год

ГЛАВНЫЙ РЕДАКТОР
Ю.Б. ЗУБАРЕВ

ЗАМЕСТИТЕЛИ ГЛАВНОГО РЕДАКТОРА:
В.В. ВИТЯЗЕВ, В.П. ДВОРКОВИЧ

РЕДАКЦИОННАЯ КОЛЛЕГИЯ:
Б.А. Бабаян, Ю.И. Борисов,
С.А. Грибачев, Г.В. Зайцев, Р.В. Зубарев,
А.П. Кирпичников, М.И. Кривошеев,
Н.А. Кузнецов, М.С. Куприянов, А.А. Ланнэ,
В.Г. Мистюков, С.Л. Мишенков,
А.А.Петровский, Ю.Н. Прохоров,
А.Н. Соловьев, Ю.Г. Сосулин,
В.В. Шахгильдян, Ю.С. Шинаков

Адрес редакции:
107031 Москва, Рождественка, 6/9/20, стр. 1
Тел.: (495) 621-71-08, (495) 621-16-39
Факс: (495) 621-16-39
E-mail: nto.popov@mtu-net.ru
tor@rgrta.ryazan.ru
http://www.dspa.ru

Для писем:
129090 Москва, а/я 48.

Издатель:
Российское научно-техническое общество радиотехники, электроники и связи им. А.С. Попова
Ответственный редактор: П.И. Зудков
Компьютерная верстка: В.В. Андреяхина
Дизайн: Н.С. Осипова

Подписной индекс по каталогу
ОАО «Роспечать» – **82185**

Подписано в печать 30.09.08
Формат 60x90/8.
Гарнитура «Arial». Печать офсетная.
Бумага офсетная. Печ.л. 6,5. Тираж 700 экз.

Заказ № 5385. Отпечатано в
ООО НПЦ «Информационные технологии»
Рязань, ул. Островского, д. 21/1
тел.: (4912) 98-69-84

Издание зарегистрировано в Министерстве Российской Федерации по делам печати, телерадиовещания и средств массовых коммуникаций.
Свидетельство о регистрации ПИ № 77-1488 от 14.01.2000

УЧРЕДИТЕЛИ:
ЗАО «Инструментальные системы»
ФГУП «НИИ радио»

В НОМЕРЕ:

Гришин В.А.
Оценка точности установления соответствия в системах технического зрения 2

Слынько Ю.В.
Решение задачи одновременного сопровождения и оконтуривания методом максимального правдоподобия 7

Радченко Ю.С., Радченко Т.А., Булыгин А.В.
Сравнительный анализ модификаций индекса структурного подобия 11

Умняшкин С.В.
Анализ эффективности применения ортогональных преобразований для кодирования дискретных сигналов с коррелированными отсчетами 15

Торчигин А.В.
Цифровая обработка сигналов в распределенном проекторе для формирования ярких изображений на широкоформатных экранах 19

Григоров И.Г., Строшков В.П., Зайнулин Ю.Г.
Способ формирования из 2D 3D изображения в сканирующей микроскопии 23

Приоров А.Л.
Обработка изображений двумерными нерекурсивными цифровыми фильтрами второго порядка 25

Перегуда Е.С.
Фрактальная обработка изображений 29

Минко А.А., Витязев В.В.
Обобщенная идентификация линейного объекта по сигналам вход-выход 34

Просочкин А.С.
Исследование спектра полиномиальных сплайнов 41

Давыдочкин В.М., Давыдочкина С.В.
Весовые функции для адаптивного гармонического анализа сигналов с многомодовым спектром 44

Костров Б.В., Саблина В.А.
Адаптивная фильтрация изображений со структурными искажениями 49

Бартенев В.Г.
Способ подавления дискретных коррелированных помех адаптивным решетчатым фильтром 54

Аминев Д.А.
Дисковые подсистемы: достижение максимальной скорости при наименьшем количестве дисков 57

Новости университетской программы Texas Instruments Inc. 60

Подписной индекс по каталогу
ОАО «Роспечать» – **82185**

УДК 519.219: 519.234.2

ОЦЕНКА ТОЧНОСТИ УСТАНОВЛЕНИЯ СООТВЕТСТВИЯ В СИСТЕМАХ ТЕХНИЧЕСКОГО ЗРЕНИЯ

Гришин В.А.

Точностные параметры систем технического зрения, используемых для определения параметров движения летательных аппаратов относительно визируемой поверхности, в значительной степени определяются точностью установления соответствия изображений одних и тех же опорных точек на последовательности кадров. В статье предлагается оценивать ошибки установления соответствия методом эллипсоидов по результатам локального анализа критериальной функции соответствия (близости) изображений в окрестности экстремума. Метод пригоден для реализации в реальном времени на борту летательного аппарата.

Введение

Оценка точности установления соответствия необходима для построения модели ошибок измерений системы технического зрения. Такие модели используются для комплексирования измерительных систем различной природы.

Задача оценивания точности установления соответствия является достаточно сложной и производится, как правило, при использовании ряда упрощающих предположений. Чаще всего в качестве модели изображения используются нормальные случайные поля [1-5]. Такое допущение является достаточно спорным в связи с тем, что статистика изображений очень вариабельна и даже в пределах одного кадра может сильно меняться по его площади. Особенно это относится к опорным точкам, которые используются в системах технического зрения для определения параметров относительного движения. Такие точки выбираются при наличии заметных перепадов яркости, конфигурация которых обеспечивает точную их локализацию на разных изображениях. Возможности практического использования таких оценок для реальных изображений очень ограничены, особенно в режиме реального времени.

Другим подходом является использование матрицы Гессе (матрицы вторых производных) яркости изображения для аппроксимации корреляционной матрицы ошибок установления соответствия [6-8]. Однако корректное вычисление матрицы Гессе должно выполняться поточечно для каждой точки окрестности, для которой устанавливается соответствие. Такие расчеты требуют заметных вычислительных ресурсов. В рассматриваемом случае также неявно предполагается, что возмущения, действующие на процесс установления соответствия, изотропны (т.е. не зависят от направления) и некоррелированы по обеим координатам. Сама же оценка определяется с точностью до постоянного множителя, как это явно указано в [6].

Здесь необходимо сделать два замечания. Во-первых, ниоткуда не следует, что возмущения, действующие на процесс установления соответствия, являются изотропными. Например, возмущения, обусловленные искажениями перспективы при полете летательного аппарата, на борту которого установлена камера, будут заведомо не изотропными. Во-вторых, представляется более целесообразным оценивать точность привязки исходя не из матрицы Гессе изображения, а исходя из анализа критериальной функции соответствия изображений. Действительно, точность локализации экстремума все-таки определяется локальной структурой использованной критериальной функцией соответствия, а не производными яркости изображения.

Представляется, что анализ локальной структуры критериальной функции в окрестности экстремума будет более точен и даст дополнительную информацию о характере возмущений, порождающих ошибки процесса установления соответствия. Примеру такого анализа и посвящена данная статья.

Вывод оценки точности установления соответствия

Рассмотрим задачу оценивания точности установления соответствия точек на стереопарах либо на последовательности изображений с учетом характеристик реального изображения. Необходимую информацию можно извлекать из локальной структуры критериальной функции соответствия окрестностей точки, которая получается при поиске соответствия.

Проблема состоит в том, что невозможно обычным способом получить информацию о законе распределения ошибок установления соответствия. Обычный способ подразумевает, что имеется ряд испытаний, результаты которых используются для определения статистических характеристик (закона распределения, моментов). В данном случае каждое испытание является выборкой из своей генеральной совокупности и нельзя производить оценки по совокупности испытаний, принадлежащих разным генеральным совокупностям, как это сделано, например в [9]. По этой причине в статье используется не закон распределения ошибок, а вероятность попадания ошибки в эллипс рассеивания.

Пусть соответствие ищется в двумерной области. Этот случай характерен для обработки последовательности кадров, снятых одной камерой, либо стереопарам, снятым не калиброванными камерами (когда нельзя использовать эпиполярные ограничения и осуществлять поиск вдоль этих ограничений – т.е. поиск по одномерной области).

В этих случаях, как правило, первоначальный поиск осуществляется по равномерной сетке, часто с шагом, равным шагу пикселей. Рассмотрим случай, когда для получения субпиксельного разрешения используется локальная аппроксимация критериальной функции соответствия полиномом второго порядка по 9 точкам (окрестность 3×3). Поместим начало локальной системы координат в точку, соответствующую максимальному значению критериальной функции. Шаг сетки считаем равным 1 – рис. 1.

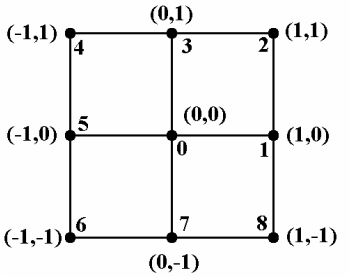


Рис. 1. Локальная система координат по дискретной сетке

Приблизим в смысле минимума среднего квадрата ошибок критериальную функцию соответствия квадратичным полиномом. Квадратичный полином имеет вид:

$$f = ax^2 + bx^2 + cxy + dx + ey + g$$

Коэффициенты квадратичного полинома выражаются через значения критериальной функции соответствия U_i, $i = 0, 1, 2 \ldots 8$ в точках 0, 1, 2 … 8. Использование для локальной аппроксимации экстремума критериальной функции полинома минимальной степени обусловлено тем, что повышение порядка аппроксимирующего полинома заметно увеличивает вычислительные затраты, но при этом совершенно не обязательно приводит к повышению точности установления соответствия. Анализ причин этого эффекта требует детального анализа статистики критериальной функции. Указанный анализ весьма не тривиален и выходит за рамки данной статьи.

Полученные координаты экстремума $x_э$ и $y_э$ позволяют совместить изображения с субпиксельным разрешением. До этого момента никаких дополнительных операций для получения оценок точности установления соответствия не делалось. Теперь предположим, что те причины, которые приводят к появлению ошибок установления соответствия, приводят также и к локальному отклонению отсчетов функции соответствия от квадратичной функции. Оценив эти отклонения, мы получим косвенную оценку указанных возмущений. Для этого перейдем к новой системе координат $(\tilde{x}, \tilde{y})$, начало которой соответствует координатам найденного экстремума $(x_э, y_э)$.

Запишем уравнения, определяющие положение экстремума при малых возмущениях $\varepsilon(\tilde{x}, \tilde{y})$ по координатам $\tilde{x}$ и $\tilde{y}$:

$$\frac{df}{d\tilde{x}} = 2a\tilde{x} + c\tilde{y} + \varepsilon'_x = 0$$
$$\frac{df}{d\tilde{y}} = c\tilde{x} + 2b\tilde{y} + \varepsilon'_y = 0$$

где ε'_x и ε'_y - производные возмущений $\varepsilon(\tilde{x}, \tilde{y})$ по координатам $\tilde{x}$ и $\tilde{y}$ соответственно. Возмущенное решение определяется из уравнений:

$$2a\tilde{x} + c\tilde{y} = -\varepsilon'_x$$
$$c\tilde{x} + 2b\tilde{y} = -\varepsilon'_y$$

Используя матричную форму записи

$A = \begin{pmatrix} 2a & c \\ c & 2b \end{pmatrix}$, получим ошибки установления соответствия при наличии возмущений:

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = A^{-1} \begin{pmatrix} -\varepsilon'_1 \\ -\varepsilon'_2 \end{pmatrix}.$$

Корреляционная матрица ошибок определения экстремума, обусловленных возмущениями:

$$M\left\{\begin{pmatrix} \tilde{x}^2 & \tilde{x}\tilde{y} \\ \tilde{x}\tilde{y} & \tilde{y}^2 \end{pmatrix}\right\} = M\left\{A^{-1}\begin{pmatrix} -\varepsilon'_x \\ -\varepsilon'_y \end{pmatrix}\begin{pmatrix} -\varepsilon'_x \\ -\varepsilon'_y \end{pmatrix}^T \left(A^{-1}\right)^T\right\} =$$

$$A^{-1} M\left\{\begin{pmatrix} -\varepsilon'_x \\ -\varepsilon'_y \end{pmatrix}\begin{pmatrix} -\varepsilon'_x \\ -\varepsilon'_y \end{pmatrix}^T\right\}\left(A^{-1}\right)^T =$$

$$A^{-1}\begin{pmatrix} M\{\varepsilon'_x\,\varepsilon'_x\} & M\{\varepsilon'_x\,\varepsilon'_y\} \\ M\{\varepsilon'_x\,\varepsilon'_y\} & M\{\varepsilon'_y\,\varepsilon'_y\} \end{pmatrix}\left(A^{-1}\right)^T =$$

$$A^{-1}\begin{pmatrix} M\{\varepsilon'_x\,\varepsilon'_x\} & M\{\varepsilon'_x\,\varepsilon'_y\} \\ M\{\varepsilon'_x\,\varepsilon'_y\} & M\{\varepsilon'_y\,\varepsilon'_y\} \end{pmatrix}A^{-1}$$

Здесь $M\{...\}$ обозначает математическое ожидание. В процессе преобразований использовано свойство симметрической матрицы: $\left(A^{-1}\right)^T = A^{-1}$.

При предположении малости ошибки установления соответствия может возникнуть только небольшое смещение положения экстремума. Реально эти ошибки можно оценить лишь косвенно. Представим, что вычислена “идеальная” критериальная функция соответствия. В окрестности экстремума эта функция локально аппроксимируется квадратичным полиномом. Экстремум этого квадратичного полинома соответствует экстремуму критериальной функции соответствия. Реальная критериальная функция соответствия искажена из-за наличия шумов, проективных искажений, и др. Эта реальная функция известна на дискретной сетке и задана своими отсчетами (рис. 2).

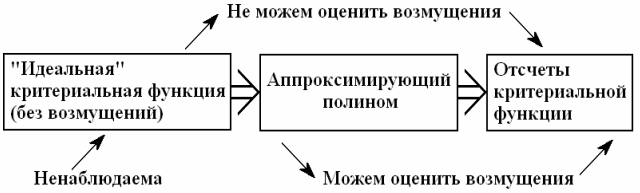


Рис. 2. Приближения критериальной функции.

Аппроксимация полиномом второго порядка локальной окрестности экстремума одновременно фильтрует (т.е. подавляет в определенной степени эти искажения). Причем трудно оценить как спектральные характеристики возмущений, так и фильтрующие свойства такой аппроксимации. Однако, эти искажения не могут быть подавлены (отфильтрованы) полностью. Остаточные искажения вызывают ошибки установления соответствия. Можно предположить, что структура возмущений аппроксимирующего полинома относительно “идеальной” критериальной функции приблизительно соответствует структуре возмущений отсчетов критериальной функции относительно аппроксимирующего полинома с некоторым коэффициентом. Данный коэффициент можно

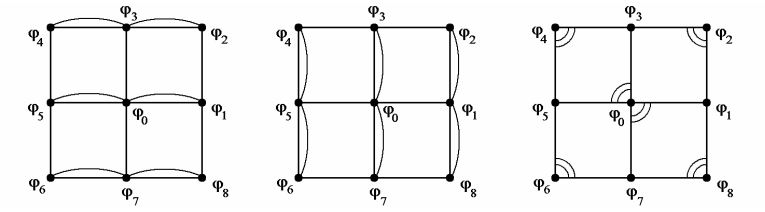


Рис. 3. Шаблоны для вычисления среднего квадратического отклонения и ковариации.

определить экспериментально для изображений разных классов, а затем адаптивно подстраивать с помощью специальной петли регулирования по невязкам процесса определения параметров движения.

Для этого вычислим значения аппроксимирующего полинома во всех 9 точках окрестности: $f_i,\ i = 0, 1, 2 \ldots 8$. Вычтем эти значения из отсчетов функции соответствия $U_i,\ i = 0, 1, 2 \ldots 8$ и вычислим значения ошибок аппроксимации:

$$\varphi_i = U_i - f_i,\ \ i = 0, 1, 2 \ldots 8$$

Получим оценки среднего квадратического отклонения и ковариации производных ошибок аппроксимации по обеим координатам:

$$\sigma_x^2 = \frac{1}{6}((\varphi_2 - \varphi_3)^2 + (\varphi_3 - \varphi_4)^2 + (\varphi_1 - \varphi_0)^2 +$$
$$+ (\varphi_0 - \varphi_5)^2 + (\varphi_8 - \varphi_7)^2 + (\varphi_7 - \varphi_6)^2)$$
$$\sigma_y^2 = \frac{1}{6}((\varphi_2 - \varphi_1)^2 + (\varphi_1 - \varphi_8)^2 + (\varphi_3 - \varphi_0)^2 +$$
$$+ (\varphi_0 - \varphi_7)^2 + (\varphi_4 - \varphi_5)^2 + (\varphi_5 - \varphi_6)^2)$$
$$\mathrm{cov}(x, y) = \frac{1}{6}((\varphi_2 - \varphi_3)(\varphi_2 - \varphi_1) +$$
$$+ (\varphi_3 - \varphi_4)(\varphi_4 - \varphi_5) + (\varphi_7 - \varphi_6)(\varphi_5 - \varphi_6) +$$
$$(\varphi_8 - \varphi_7)(\varphi_1 - \varphi_8) + (\varphi_0 - \varphi_5)(\varphi_3 - \varphi_0) + (\varphi_1 - \varphi_0)(\varphi_0 - \varphi_7))$$

Шаблоны для вычисления среднего квадратического отклонения и ковариации приведены на рис. 3. Для оценки средних квадратических отклонений производных ошибок аппроксимации использовалась конечно-разностная аппроксимация производных с шагом, равным одному пикселю. Для используемой окрестности 3×3 такая оценка может быть сделана по 6 парам точек. Для оценки ковариации производных может быть использовано несколько разных, но приблизительно равнозначных аппроксимаций. Выбрана одна из них.

Вычисленная корреляционная матрица

$$\begin{pmatrix} \sigma_x^2 & \mathrm{cov}(x, y) \\ \mathrm{cov}(x, y) & \sigma_y^2 \end{pmatrix}$$

характеризует возмущения критериальной функции, определенной на дискретной сетке, относительно аппроксимирующего квадратичного полинома. Можно ожидать, что возмущения самого квадратичного полинома характеризуются матрицей, которую можно аппроксимировать полученной матрицей с некоторым, пока неизвестным, коэффициентом.

Алгоритм установления соответствия ищет положение экстремума критериальной функции соответствия (в точке, где производные по обеим координатам равны нулю), а возмущения вызывают смещение этой точки. Определим разброс ошибок установления соответствия методом эллипсоидов [10]. Так как ошибки установления соответствия определены на плоскости, то эллипсоиды переходят в эллипсы.

Для закона распределения

$$f_*(X) = \frac{1}{2\pi(\det R_*)^{1/2}} \exp\left\{-\frac{1}{2} X^T R_*^{-1} X\right\}$$

вероятность попадания случайной точки в эллипс $X^T R_*^{-1} X \le s^2$ равна:

$$P(s) = 1 - \exp\left(\frac{-s^2}{2}\right).$$

Пусть возмущения квадратичного полинома характеризуются матрицей:

$$\begin{pmatrix} r_1 & r_2 \\ r_2 & r_3 \end{pmatrix} = k \begin{pmatrix} \sigma_x^2 & \mathrm{cov}(x, y) \\ \mathrm{cov}(x, y) & \sigma_y^2 \end{pmatrix}.$$

Здесь k - коэффициент, зависящий от закона распределения, заданной величины вероятности $P(s)$ и других факторов. Тогда обратная матрица R_*^{-1} имеет вид:

$$R_*^{-1} = \frac{1}{r_1 r_3 - r_2^2} \begin{pmatrix} r_3 & -r_2 \\ -r_2 & r_1 \end{pmatrix}$$

В процессе функционирования алгоритмов технического зрения, в частности, при использовании последовательности стереоскопических или монокулярных изображений для решения задач оценки относительных линейных и угловых перемещений имеется возможность уточнения этого коэффициента по величинам невязок направлений на точки визируемой поверхности на последовательности кадров.

Экспериментальная часть

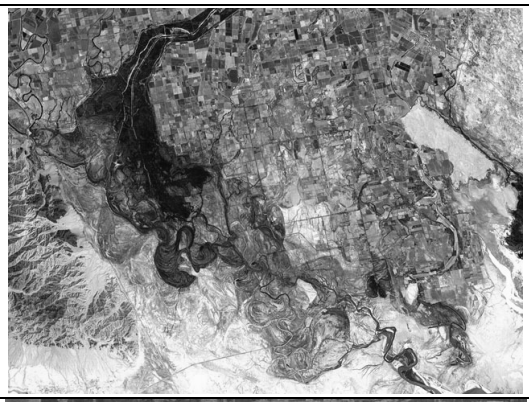
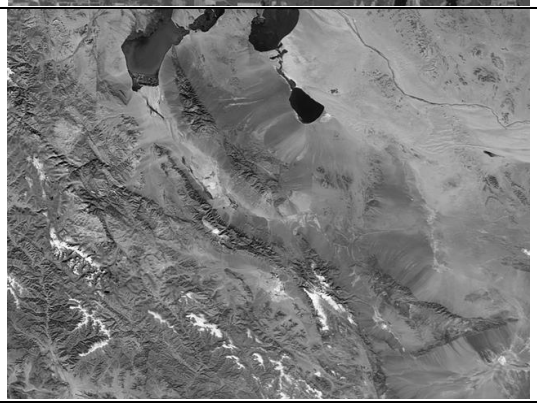
Для проверки полученных соотношений была использована методика, изложенная в [11, 12]. А именно, по заданной траектории движения, профилю визируемой поверхности и условиям съемки были подготовлены наборы матриц линейных преобразований, аппроксимирующих наибольшие локальные проективные искажения для указанных условий. Камера двигалась параллельно поверхности, угол разворота камеры по отношении к вектору скорости был выбран равным 47.5°, 67.5° и 87.5°. Соответственно, величина проективных искажений была наибольшей для угла 47.5° и минимальна для угла 87.5° (оптическая ось камеры почти ортогональна визируемой поверхности). Затем были выбраны три изображения с существенно различающимися корреляционными свойствами. На изображениях размером 640×480 пикселей выбиралось 9945 окрестностей размером 9×9 пикселей. Эти окрестности подвергались искажениям из полученного набора, и запускался процесс установления соответствия по окрестности 17×17 пикселей. После поиска смещения, соответствующего наи-

большему значению функции соответствия, производилась квадратичная аппроксимация функции соответствия и оценка положения экстремума с субпиксельным разрешением. Если ошибка превышала 0.5 пикселя, то такие точки исключались из дальнейшего анализа как имеющие аномальные ошибки установления соответствия.

Точки, для которых матрица вторых производных A не была положительно определенной или параметр s^2 принимал значения, меньше или равные нулю, отбрасывались как потенциально ненадежные. Далее оценивались параметры эллипсоида разброса, и определялся коэффициент k. Так как закон распределения ошибок установления соответствия аппроксимируется гауссовым законом с не очень высокой точностью, то значения этого коэффициента определялись для следующего набора вероятностей попадания ошибки установления соответствия в эллипсоид рассеяния: 0.9, 0.95, 0.975, 0.99.

Как следует из таблицы 1, величина коэффициента k меняется в зависимости от характера изображения в 1.5-2.5 раза. Причем следует отметить, что для изображений с малым радиусом корреляции (самое верхнее изображение) эллипсоид разброса оценивается точнее (т.к. коэффициент k меньше). Для изображений с большим радиусом корреляции возрастает ошибка оценивания возмущений относительно аппроксимирующего полинома (см. рис. 2). Величина коэффициента k в зависимости от ракурса визирования поверхности меняется в 3-6 раз. При увеличении проективных искажений (увеличении угла разворота камеры относительно нормали к поверхности) величина коэффициента k возрастает. Это также обусловлено возрастанием ошибок оценивания возмущений относительно аппроксимирующего полинома. Как уже отмечалось, величину этого коэффициента целесообразно адаптивно подстраивать исходя из величины невязок направлений визирования особенностей на последовательности кадров с помощью специальной петли регулирования.

Таблица 1.

Вероятность	Угол разворота оси камеры относительно вектора скорости			Изображение
	47.5°	67.5°	87.5°	
0.9	5.6	3.1	1.1	
0.95	8.0	4.2	1.6	
0.975	12.1	5.8	2.3	
0.99	19.9	9.5	3.5	
0.9	10.7	6.8	2.8	
0.95	14.9	9.4	3.8	
0.975	19.9	12.6	5.2	
0.99	30.9	19.1	8.8	
0.9	6.9	4.9	2.3	
0.95	9.8	7.0	3.5	
0.975	14.3	9.8	5.0	
0.99	24.7	15.8	8.5	

Заключение

Полученный алгоритм оценки точности установления соответствия имеет незначительную вычислительную сложность и может быть реализован на борту летательного аппарата в реальном времени. Алгоритм позволяет исключать из анализа точки, матрица вторых производных критериальной функции которых не является положительно определенной, а также точки, для которых нарушается условие $s^2 > 0$, как потенциально ненадежные. Кроме того, он позволяет оперативно исключать из анализа точки, для которых величина невязки значительно выйдет из эллипса рассеивания, как аномальные измерения. Это позволяет существенно снизить вычислительные затраты, необходимые при переборном исключении точек, что характерно для семейства алгоритмов RANSAC.

Оценки точности установления соответствия используются для построения модели ошибок системы технического зрения, что необходимо для комплексирования системы технического зрения с другими приборами управления полетом.

Алгоритм пригоден для широкого круга функций соответствия, которые в окрестности экстремума могут быть аппроксимированы полиномом второй степени.

Литература

1. Антипин В. В., Буймов А. Г. Статистический анализ ошибок совмещения изображений по методу наименьших квадратов в условиях окрашенного шума // Автометрия. – 1985. – №3. – С. 27-31.
2. Буймов А. Г. Анализ влияния корреляционных свойств неоднородного яркостного шума на ковариационную матрицу ошибок совмещения изображений // Автометрия. – 1985. – №4. – С. 9-15.
3. Егоров И. В., Юхно П. М. Влияние однородного коррелированного яркостного шума на ошибки совмещения изображений // Автометрия. – 1988. – №1. – С. 105-107.
4. Тарасенко В. П., Тимофеев А. В. Доверительное оценивание точности совмещения изображений в корреляционно-экстремальных системах // Автометрия. – 1990. – №4. – С. 106-111.
5. Степанов О. А. Предельно достижимая точность совмещения гауссовых изображений // Автометрия. – 1990. – №5. – С. 16-23.
6. Sun Z., Ramesh V., Tekalp A. M. Error characterization of the factorization method // Computer Vision and Image Understanding. – 2001. – Vol. 82. Issue 2. – P. 110–137 (гессиан)
7. Chowdhury A. R. Chellappa R. Stochastic approximation and rate-distortion analysis for robust structure and motion estimation // International Journal of Computer Vision. – 2003. – Vol. 55. – №1. – P. 27–53.
8. Chowdhury A.R., Chellappa R. Statistical Error Propagation in 3D Modeling from Monocular Video // Proc. IEEE 2003 Conference on Computer Vision and Pattern Recognition. Workshop on Statistical Analysis in Computer Vision. – 2003. http://www.cse.lehigh.edu/~rjm2/SACV/papers/chowdhury-chellappa.pdf.
9. Sibley G., Matthies L., Sukhatmey G. Bias Reduction and Filter Convergence for Long Range Stereo // Proceedings of the 12th International Symposium of Robotics Research. – 2005.
10. Огарков М. А. Методы статистического оценивания параметров случайных процессов. – М.: Энергоатомиздат, 1990, – 208 с.
11. Гришин В. А. Оценка локальных трансформаций последовательности изображений, используемых для 3D реконструкции // Сб. научн. тр. Нелинейные явления в открытых системах. Вып. 14. – М.: Гос. ИФТП, 2003. – С.80-91.
12. Гришин В. А. Оценка статистических характеристик изображений местности с точки зрения их пригодности для 3D реконструкции // Сб. научн. тр. Искусственный интеллект в технических системах. Вып. 24. – М.: Гос. ИФТП, 2004. – С. 63-74.

Уважаемые коллеги!

Вы можете подписаться на наш журнал в любом отделении связи. Подписной индекс издания по каталогу ОАО «Роспечать» - 82185.

Для тех, кто не успел оформить подписку на первое полугодие 2009 года через ОАО «Роспечать», сохраняется возможность подписки и приобретения журналов непосредственно в редакции по адресу: 107031, г. Москва, Рождественка, 6\9\20, стр. 1, Российское научно-техническое общество радиотехники, электроники и связи им. А.С. Попова, или оформить Заказ в соответствии с требованиями, выставленными на сайте журнала: www.dspa.ru.

Справки по телефонам: (495) 621-71-08, 621-06-10.

Факс: (495) 621-16-39.

E-mail: nto.popov@mtu-net.ru

tor@rgrta.ryazan.ru

УДК 004.932

РЕШЕНИЕ ЗАДАЧИ ОДНОВРЕМЕННОГО СОПРОВОЖДЕНИЯ И ОКОНТУРИВАНИЯ МЕТОДОМ МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ

Слынько Ю.В.

Рассмотрен метод сопровождения выбранного объекта в видеопоследовательности при отсутствии априорной информации о поведении фона, а также о размере и форме объекта. В таких условиях одновременно с сопровождением должно быть произведено оконтуривание объекта. Для решения этих задач применяется метод максимального правдоподобия. Представлен метод оценки функции правдоподобия на основе невязок частей кадра. Итеративный алгоритм применен для поиска максимума оцененной функции правдоподобия. Разработанный алгоритм показал высокое качество работы на широком спектре видеопоследовательностей.

Введение

Сопровождение выделенных объектов в видеопоследовательности – актуальная задача для систем наблюдения, охранных систем, интеллектуальных систем управления автомобилем и робототехники. Кроме того, результаты решения этой задачи могут быть использованы как начальный этап для создания более сложной системы анализа видеопоследовательности.

Предложенный в данной работе алгоритм разрабатывался для управляемых камер с возможностью большого увеличения, что накладывает достаточно серьезные требования на его функциональность. Алгоритм не должен иметь в своей основе представления о статичности и известности фона. Это обусловлено двумя причинами: во-первых, камера может быстро перемещаться (особенно в условиях сильного увеличения), что не дает возможности накопить информацию о фоне; во-вторых, фон может быть сильно нестационарным, например, поверхность моря.

Еще одним условием является отсутствие априорных знаний о размерах и форме объекта. Это требование связанно с тем, что алгоритм разрабатывался для систем, в которых оператор, помимо выделения объектов, вынужден управлять камерой. В этом случае он не имеет возможности точно задать размеры объекта, которые могут варьироваться из-за изменения увеличения.

На данный момент задача сопровождения объектов в общем виде не имеет единственного удовлетворительного решения. Существуют два различных способа ее решения. Первый – на основе отделения движущихся объектов от фона с последующей классификацией (чтобы выбрать на кадре из всех объектов только интересующий). Второй – на основе поиска на следующем кадре выделенного объекта из предыдущего кадра.

В первом случае алгоритм работает с неподвижными камерами и имеет ряд недостатков, например, невозможность сопровождать неподвижные объекты или объекты на сильно изменяющемся фоне. Такие задачи характерны для систем видеонаблюдения и мониторинга (большинство коммерческих продуктов относится именно к этой категории). Здесь сопровождение объекта основано на использовании стационарности хорошо известного «подстилающего» фона. Обзор таких методов дан в [1].

Для сопровождения объектов на более сложном фоне, как правило, используются алгоритмы с накоплением информации (фильтрацией). Базовыми являются алгоритмы фильтрации на основе фильтра Калмана [2, 3]. Среди наиболее изощренных алгоритмов следует выделить фильтрацию методом частиц [4, 5] и с использованием ядра объекта [6, 7].

Основная проблема алгоритмов с накоплением заключается в завязке сопровождения, т.е. в обработке нескольких первых кадров, пока не накоплено достаточное количество информации. Это особенно актуально, когда алгоритм должен управлять камерой, следящей за объектом. Если алгоритм не справится с задачей, объект исчезнет из поля зрения камеры за несколько кадров.

Для решения задачи сопровождения объекта с одновременным нахождением его контура алгоритм классификации (разделения) кадра на объект и фон должен быть сильно взаимозавязан с алгоритмом сопровождения (нахождения сдвига). Один из примеров таких алгоритмов представлен в [8]. Однако здесь заложено жесткое ограничение на форму объекта – она должна быть близка к прямоугольной.

Предложенный в данной статье алгоритм позволяет решить задачу сопровождения и оконтуривания движущихся объектов при отсутствии априорной информации об их форме, размерах и характере изменения фона.

Теоретические основы метода

Будем решать задачу одновременного определения контура объекта и его сдвига методом максимального правдоподобия. В случае отсутствия априорной информации данный подход является оптимальным.

Кадр f_t момента времени t состоит из конечного количества частей – пикселей (для ускорения алгоритма можно использовать части, состоящие из нескольких пикселей; далее будет использоваться термин «часть»). Будем считать, что каждая часть полностью принадлежит либо объекту, либо фону. Первоначальный выбор объекта осуществляется оператором и заключается в выборе части i_0, которая заведомо принадлежит объекту.

Запишем функцию правдоподобия для данной задачи, т.е. принадлежности всех частей кадра к объекту или фону:

$$J(I) = \prod_{i \in I} p(i \in object) \prod_{i \notin I} p(i \notin object), \quad (1)$$

где i – номер части в кадре, I – множество частей, принадлежащих объекту, $p(i \in object)$ - вероятность того, что часть i принадлежит объекту.

Таким образом, для нахождения сдвига необходимо найти максимум правдоподобия, т.е.

$$I = \arg\max_{I} J(I) \quad (2)$$

Основная проблема заключается в нахождении вероятности того, что часть принадлежит объекту. Т.к., по условиям задачи, объект отличается от фона только вектором сдвига, такая вероятность должна зависеть от сдвига объекта и фона.

Вычислим функцию невязки (как функцию сдвига $(\delta x, \delta y)$) для каждой части и всего объекта (подробности вычисления этой функции могут быть найдены в [9]:

$$\widetilde{C}_i(\delta x, \delta y) = \sum_{l,k \in P_i} \left(f_t(l,k) - f_{t-1}(l-\delta x, k-\delta y)\right)^2, \quad (3)$$

$$\widetilde{F}(\delta x, \delta y) = \sum_{i \in I} \widetilde{C}_i(\delta x, \delta y),$$

где P_i - пиксели, принадлежащие части i.

Будем рассматривать вычисленные функции невязки как оценки истинных C_i и F. Минимум этих функций будет оценкой сдвига каждой части или всего объекта. Можно было бы выбрать части объекта, сдвиг которых близок к сдвигу части i_0.Но как показали эксперименты, этот алгоритм чрезвычайно неустойчив из-за больших ошибок определения сдвига отдельной части.

Более точным будет вычисление вероятности $p(i \in object)$ как вероятности того, что минимум C_i и F совпадают. Тогда

$$p(i \in object) = \sum_{\delta x, \delta y} p_i(\delta x, \delta y) p_F(\delta x, \delta y) \text{ или приближенно} \quad (4)$$

$$p(i \in object) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p_i(\delta x, \delta y) p_F(\delta x, \delta y) d\delta x d\delta y \quad (5)$$

где $p_i(\delta x, \delta y)$ и $p_F(\delta x, \delta y)$ - вероятности того, что значение сдвига части или объекта равны $(\delta x, \delta y)$.

Эти вероятности могут быть определены из оценок функций невязок. Для этого заменим распределение оценок сдвига каждой части нормальным распределением:

$$p_i(\delta x, \delta y) = \frac{1}{2\pi\sqrt{\det(\Gamma_i)}} e^{-\frac{1}{2}(\vec{x}-\vec{x}_i)^T \Gamma_i^{-1} (\vec{x}-\vec{x}_i)}, \text{ где} \quad (6)$$

$$\vec{x} = (\delta x, \delta y), \ \Gamma_i = \begin{pmatrix} \sigma_{ix}^2 & \rho_i \sigma_{ix} \sigma_{iy} \\ \rho_i \sigma_{ix} \sigma_{iy} & \sigma_{iy}^2 \end{pmatrix}$$

и аналогично для F.

Параметры этого распределения $\vec{x}_i, \sigma_{ix}, \sigma_{iy}, \rho_i$ выберем исходя из оценки функции невязки (подробности см. в Приложении 1). Тогда, вычисляя интеграл, находим:

$$p(i \in object) = \frac{\exp(\frac{B}{2A})}{2\pi\sqrt{-A}}, \text{ где} \quad (7)$$

$$A = 2\rho_i \rho_F \sigma_{Fx} \sigma_{Fy} \sigma_{ix} \sigma_{iy} +$$

$$+ \sigma_{Fx}^2((\rho_F^2 - 1)\sigma_{Fy}^2 - \sigma_{iy}^2) + \sigma_{Fx}^2((\rho_i^2 - 1)\sigma_{iy}^2 - \sigma_{Fy}^2),$$

$$B = (\sigma_{Fy}^2 + \sigma_{iy}^2)(x_F - x_i)^2 -$$

$$-2(\rho_F \sigma_{Fx} \sigma_{Fy} + \rho_i \sigma_{ix} \sigma_{iy})(x_F - x_i)(y_F - y_i) + (\sigma_{Fx}^2 + \sigma_{ix}^2)(y_F - y_i)$$

Детали реализации

Для решения поставленной задачи – нахождения контура объекта с одновременным определением его сдвига – необходимо максимизировать функцию правдоподобия. Основная проблема заключается в том, что функция невязки F для объекта зависит от контура объекта I. Это затрудняет поиск максимума. Задачу можно решить методом перебора, что чрезвычайно трудоемко и исключает практическое применение. Для сокращения объема вычисления и повышения устойчивости, части выбираются как квадраты размером 8x8 пикселей. Поэтому предлагается следующая схема: контур объекта вычисляется одним из стандартных алгоритмов кластеризации, а затем рекуррентно уточняется, используя функцию правдоподобия.

Задачей алгоритма кластеризации является нахождение первого приближения контура объекта. Кластеризация ведется по сдвигам отдельных частей – т.е. вычисляется множество сдвигов всех частей и делится на две группы, одна из которых соответствует объекту, а вторая – фону. В качестве алгоритма кластеризации были использованы два хорошо известных алгоритма: на основе матрицы сходства и метод К-средних. Описание обоих методов дано в [10].

Задав начальное разбиение кадра на объект и фон, можно найти локальный максимум функции правдоподобия методом последовательных приближений. Для этого вычислим функцию невязки $\widetilde{F}(\delta x, \delta y)$ для объекта по контуру I. Затем для каждой части найдем вероятности $p(i \in object)$ и $p(i \notin object)$. Если $p(i \in object) > p(i \notin object)$, то часть принадлежит объекту, иначе – фону (это следует из критерия максимизации функции правдоподобия). Таким образом, строится новый контур объекта. Итерация повторяется (т.е. для нового контура вычисляется $\widetilde{F}(\delta x, \delta y)$ и т.д.). Когда контур I перестанет изменяться – это будет означать, что достигнут локальный максимум, и итерации следует прекратить.

На практике такой подход позволяет достаточно быстро (за 3-4 итерации) находить локальный максимум, который достаточно близок или вовсе совпадает с глобальным.

Результаты тестирования и выводы

Работа алгоритма была проверена на большом количестве видеопоследовательностей разного характера. Особенно внимательно исследовались случаи существенно нестационарного фона и быстрых перемещений камеры (а значит и фона).

Помимо реальных видеопоследовательностей, качество работы алгоритма проверялось и на модельных, где на нестационарный фон накладывалось изображение движущегося объекта. Такой метод позволяет получить оценки точности алгоритма определения сдвига. На исследованных видеопоследовательностях СКО определения сдвига составила 0.08 пикселей.

На рис. 1 приведены результаты работы алгоритма с различными видеопоследовательностями. В верхнем ряду представлены результаты работы алгоритма на

синтетических кадрах. Два средних ряда – сопровождение человека на нестационарном фоне. В двух нижних рядах – сопровождение лодки на водной поверхности с яркими информативными бликами при перемещающейся камере. Видно, что некоторые части фона ошибочно определятся как принадлежащие объекту. Они в дальнейшем отсеиваются методами морфологической обработки На изображениях нижнего ряда видно, что определились как части объекта даже удочки и антенны.

Таким образом, был создан алгоритм сопровождения и оконтуривания выделенных объектов при нестационарном фоне и отсутствии априорной информации о размерах и контуре объекта.

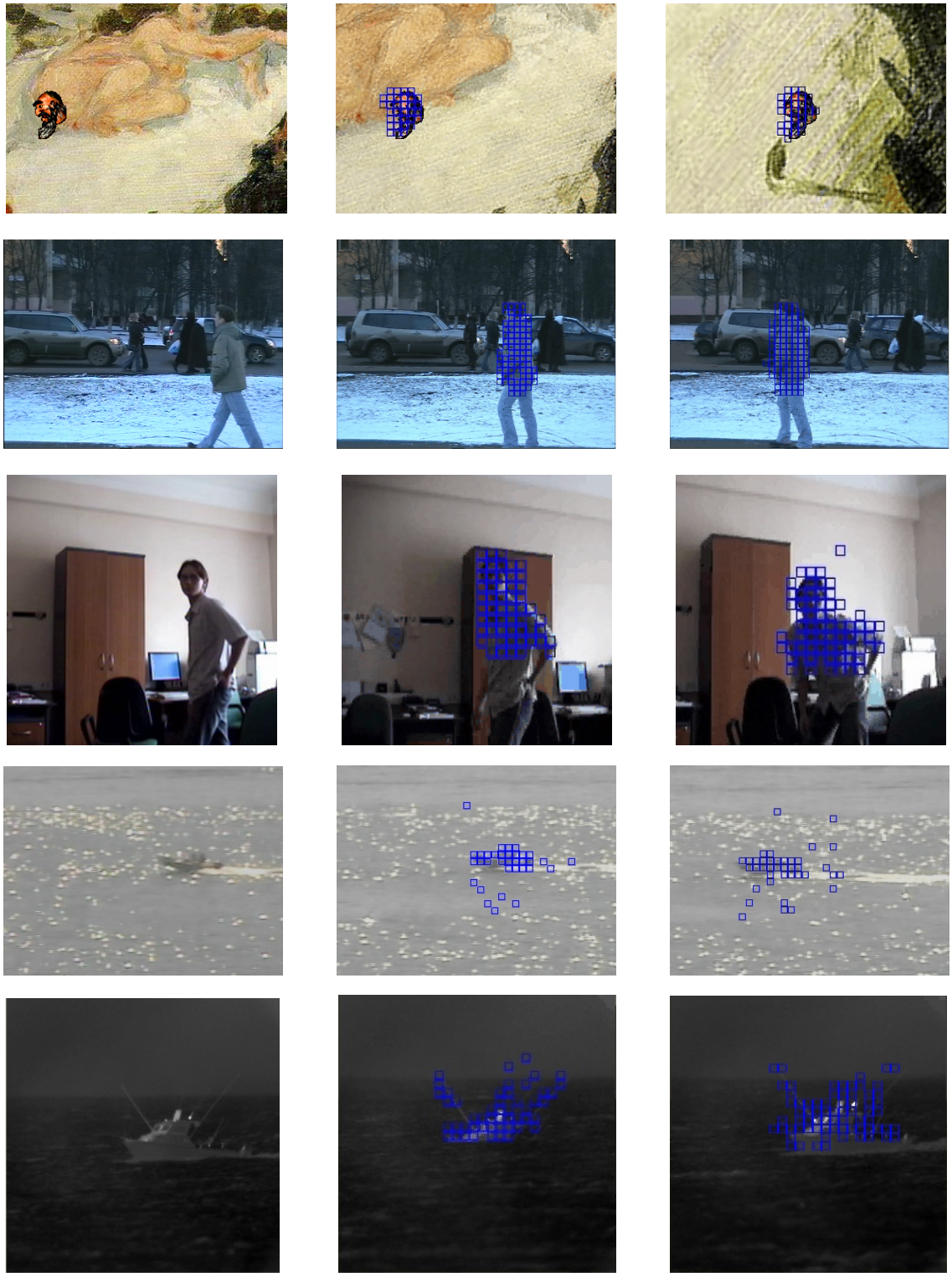


Рис.1. Результаты работы алгоритма. Синими квадратами обозначены части, определенные как принадлежащие объекту

Литература

1. Fatih Porikli. Achieving Real-Time Object Detection and Tracking Under Extreme Conditions.
2. C. Wren, A. Azarbayejani, T. Darell, and A. Pentland, “Pfinder: Real-time tracking of the human body,” IEEE Trans. Pattern Anal. Machine Intell., vol. 19, pp. 780–785, 1997.
3. Y. Boykov and D. Huttenlocher, “Adaptive bayesian recognition in tracking rigid objects,” in Proc. IEEE Conf. on Computer Vision and Pattern Recognition, Hilton Head, SC, vol. II, 2000, pp. 697–704.
4. M. Isard and I. Blake, “Condensation – conditional density propagation for visual tracking,” in Intl. J. of Computer Vision, vol. 29, 1998, pp. 5–28.
5. N. Bouaynaya, W. Qu, and D. Schonfeld, “An online motion-based particle filter for head tracking applications,” in Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Philadelphia, 2005.
6. D. Comaniciu, V. Ramesh, and P. Meer, “Real-time tracking of non-rigid objects using mean shift,” in Proc. IEEE Conf. on Computer Vision and Pattern Recognition, Hilton Head, SC, vol. 1, 2000, pp. 142–149.
7. F. Porikli and O. Tuzel, “Object tracking in low-frame-rate video,” in Proc. of PIE/EI - Image and Video Communication and Processing, San Jose, CA, 2005.
8. S. Avidan, “Ensemble tracking,” in Proc. IEEE Conf. on Computer Vision and Pattern Recognition, San Diego, CA, 2005.
9. Yu. V. Slyn’ko, V. N. Lagutkin, and A. P. Luk’yanov. Development and Investigation of Real-Time Robust Algorithms for Estimating the Parameters of Geometric Transformations of Video-Sequence Frames. Journal of Communications Technology and Electronics, 2007, Vol. 52, No. 3, pp. 332–338.
10. Д. Форсайт, Ж. Понс. Компьютерное зрение. Современный подход. М: Вильямс, 2004.

ПРИЛОЖЕНИЕ 1

Для оценки разброса значений функции невязки C_i предлагается следующая схема. Исходя из предположения о том, что шумы в пикселях распределены нормально и некоррелированы, находится доверительная область нахождения минимума функции невязки (подробности см. в [9]).

$$\widetilde{D} = \left\{ \delta x, \delta y : \widetilde{C}_i(\delta x, \delta y) < m_n + B \cdot \sigma_n \right\}, \quad (8)$$

где m_n - значение функции невязки в минимуме,

$\sigma_n = \frac{m_n \sqrt{2}}{\sqrt{N_n}}$ - дисперсия значений невязки,

N_n – количество точек в области, по которой подсчитана функция C_i (для части размеров 8 на 8 $N_n = 64$), B - порог, определяющий доверительную вероятность ошибки.

По этой области рассчитываются параметры распределения

$$\vec{x}_i = \begin{pmatrix} x_i \\ y_i \end{pmatrix} = \frac{1}{M(\widetilde{D})} \sum_{\delta x, \delta y \in \widetilde{D}} \begin{pmatrix} \delta x \\ \delta y \end{pmatrix},$$

$$\begin{pmatrix} \sigma_{ix} \\ \sigma_{iy} \end{pmatrix} = \frac{1}{M(\widetilde{D})} \sum_{\delta x, \delta y \in \widetilde{D}} \begin{pmatrix} (\delta x - x_i)^2 \\ (\delta y - y_i)^2 \end{pmatrix}, \quad (9)$$

$$\rho_i = \frac{1}{M(\widetilde{D})} \frac{1}{\sigma_{ix}\sigma_{iy}} \sum_{\delta x, \delta y \in \widetilde{D}} (\delta x - x_i)(\delta y - y_i)$$

Вычисление параметров для F производится аналогично.

У в а ж а е м ы е а в т о р ы !

Редакция научно-технического журнала "Цифровая обработка сигналов" просит Вас соблюдать следующие требования к материалам, направляемым на публикацию:

1) Требования к текстовым материалам и сопроводительным документам:

- Текст - текстовый редактор Microsoft Word.
- Таблицы и рисунки должны быть пронумерованы. На все рисунки, таблицы и библиографические данные указываются ссылки в тексте статьи.
- Объем статьи до 12 стр. (шрифт 12). Для заказных обзорных работ объем может быть увеличен до 20 стр.
- Название статьи на русском и английском языках.
- Рукопись статьи сопровождается:
  - краткой аннотацией на русском и английском языках;
  - номером УДК;
  - сведениями об авторах (Ф.И.О., организация, телефоны, электронная почта).

2) Требования к иллюстрациям:

- Векторные (схемы, графики) - желательно использование графических редакторов Adobe Illustrator или Corel DRAW.
- Растровые (фотографии, рисунки) - М 1:1, разрешение не менее 300dpi, формат tiff.

УДК 621.397.2

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МОДИФИКАЦИЙ ИНДЕКСА СТРУКТУРНОГО ПОДОБИЯ

Радченко Ю.С., Радченко Т.А., Булыгин А.В.

Предложены модификации критерия качества восстановленного после сжатия изображения – индекса структурного подобия, основанные на методах непараметрической статистики. Показано преимущество предложенных модификаций при наличии импульсных помех, блочных искажений, размытии контуров по сравнению с исходным критерием.

Введение

При разработке автоматизированных систем оценки качества цифровых изображений, тестировании кодеков для сжатия видео, оценки качества систем передачи видеоинформации весьма актуальной является проблема построения объективного критерия качества изображений. Сложность решения названной проблемы в том, что объективный критерий должен быть достаточно простым и одновременно хорошо соответствовать восприятию зрительной системы человека.

В последнее время среди специалистов по обработке изображений нашли признание критерии структурного подобия MSSIM и MESSIM [1,2]. Как показали исследования, эти критерии вполне адекватно оценивают качество изображений при наличии искажений в виде гауссова шума и размытости контуров. Однако при наличии таких искажений, как импульсный шум, блочность они дают значение качества не вполне соответствующие зрительному восприятию системе человека.

В данной работе предложены модификации критерия MSSIM, основанные на непараметрических методах статистики.

Модификации MSSIM

Критерий структурного подобия MSSIM [1,2] определяется путем усреднения индекса структурного подобия SSIM по всем блокам сравниваемых изображений $\boldsymbol{X}$ и $\boldsymbol{Y}$:

$$\mathrm{MSSIM}(X,Y)=\frac{1}{M}\sum_{j=1}^{M}\mathrm{SSIM}_j(x,y). \qquad (1)$$

Значение $\mathrm{SSIM}_j(x, y)$ для j- го блока вычисляется по формуле:

$$\mathrm{SSIM}_j(x, y) = l(x, y)\cdot c(x, y)\cdot s(x, y). \qquad (2)$$

Здесь l(x,y) - функционал «сравнение яркости», c(x,y) - функционал «сравнение контрастности» и s(x,y) – функционал «сравнение структуры».

В соответствии с [2] функционалы сравнения в блоках вычисляются следующим образом:

$$l(x,y)=\frac{2\mu_x\mu_y+C_1}{\mu_x^2+\mu_y^2+C_1},\quad c(x,y)=\frac{2\sigma_x\sigma_y+C_2}{\sigma_x^2+\sigma_y^2+C_2},$$
$$s(x,y)=\frac{\sigma_{xy}+C_3}{\sigma_x\sigma_y+C_3}, \qquad (3)$$

где μ_x, μ_y , σ^2_x , σ^2_y – выборочные среднее и дисперсия значений яркостной компоненты для блоков изображений $\boldsymbol{X}$ и $\boldsymbol{Y}$ соответственно; σ_{xy} - корреляционный момент между векторами значений в блоках; C_1, C_2, C_3 – малые константы. С точностью до константы C_3 функционал $s(x,y)$ совпадает с выборочным коэффициентом корреляции Пирсона.

Выбор в качестве меры структурного различия функционала s(x,y) в виде (3) наиболее обоснован при сравнении векторов значений, имеющих гауссово распределение. Поэтому критерий MSSIM хорошо отслеживает искажения в виде гауссова шума. Однако, он неудовлетворительно работает при размытии контуров.

Существует модификация критерия структурного подобия MESSIM [4], основанная на информации о границах изображения. В MESSIM вместо (3) используются функционалы:

$$l(x,y)=\frac{2\mu_x\mu_y+C_1}{\mu_x^2+\mu_y^2+C_1},\quad c(x,y)=\frac{2\sigma_x\sigma_y+C_2}{\sigma_x^2+\sigma_y^2+C_2},$$
$$s(x,y)=\frac{\sigma'_{xy}+C_3}{\sigma'_x\sigma'_y+C_3}, \qquad (4)$$

где σ'_x и σ'_y - дисперсия векторов D_x и D_y соответственно; σ'_{xy} – ковариация векторов D_x и D_y; D_x и D_y – векторы направления границ исходного блока изображения и искаженного соответственно.

При неизвестных законах распределения искажений целесообразно воспользоваться оценками соответствующих структурных характеристик на основе непараметрической статистики [5]. В частности, для оценки яркости целесообразно использовать выборочную медиану, обладающую по сравнению с выборочным средним большей устойчивостью, а для оценки структурного подобия вместо коэффициента корреляции Пирсона – один из ранговых коэффициентов корреляции.

Нами предлагаются две модификации критерия структурного подобия на основе непараметрических методов (Mean Nonparametric Structural Similarity): MNSSIM1 и MNSSIM2 [8].

Для MNSSIM1 вместо (3) предлагается использовать функционалы:

$$l(x,y)=\frac{2m_xm_y+C_1}{m_x^2+m_y^2+C_1},\ c(x,y)=\frac{2\sigma m_x\sigma m_y+C_2}{\sigma m_x^2+\sigma m_y^2+C_2},$$
$$s(x,y)=R_S(x,y), \qquad (5)$$

где $m_x=\mathrm{median}(\vec{x})$ и $m_y=\mathrm{median}(\vec{y})$ - медианы векторов значений яркости в блоках изображений $\boldsymbol{X}$ и $\boldsymbol{Y}$ соответственно;

$\sigma m_x^2 = \text{median}\left[(\vec{x} - m_x)^2\right]$, $\sigma m_y^2 = \text{median}\left[(\vec{y} - m_y)^2\right]$ -

медианы квадрата разности векторов значений яркости и медианы, $R_S(x,y)$ – ранговый коэффициент корреляции Спирмена [5].

Для MNSSIM2 предлагается использовать функционалы:

$$l(x,y) = \frac{2m_x m_y + C_1}{m_x^2 + m_y^2 + C_1}, \quad c(x,y) = \frac{2\sigma_x \sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}, \quad s(x,y) = R_S(x,y), \quad (6)$$

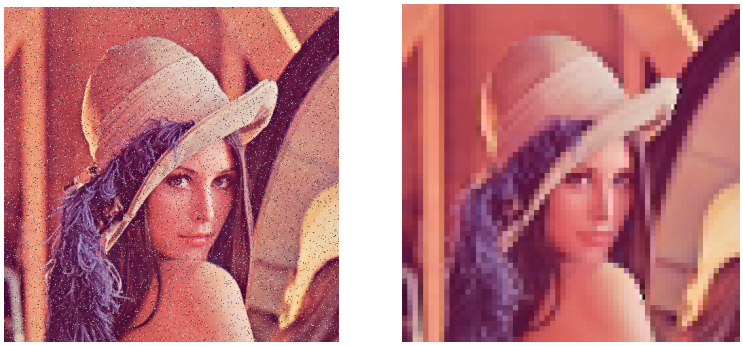
т.е. функционал «сравнение контрастности» c(x,y) оставить тем же, что в MSSIM

Константы C_1 и C_2 в (5), (6) идентичны, используемым при подсчете MSSIM [2].

ния более адекватные зрительному восприятию и близкие между собой.

На рис. 2 представлены зависимости критериев MSSIM, MESSIM, MNSSIM1, MNSSIM2, от параметра интенсивности шума *p* для импульсного шума типа «соль/перец». Аналогичные зависимости имеют место для импульсных шумов «перец» и «соль». Практически во всем диапазоне изменений *p* значения критериев MNSSIM1 и MNSSIM2 близки между собой и более адекватны зрительному восприятию, чем значения критериев MSSIM и MESSIM.

Таким образом, при исследовании влияния искажений в виде импульсного шума следует отдать



а) б)

Рис. 1. Изображение «Лена»: а) шум - "соль/перец", p=0,05. MSSIM = 0,315, MESSIM = 0,331; MNSSIM1 = 0,899, MNSSIM2 = 0,898; б) восстановленное после максимального сжатия при помощи JPEG: MSSIM = 0,634, MESSIM = 0,578, MNSSIM1 = 0,073, MNSSIM2 = 0,079.

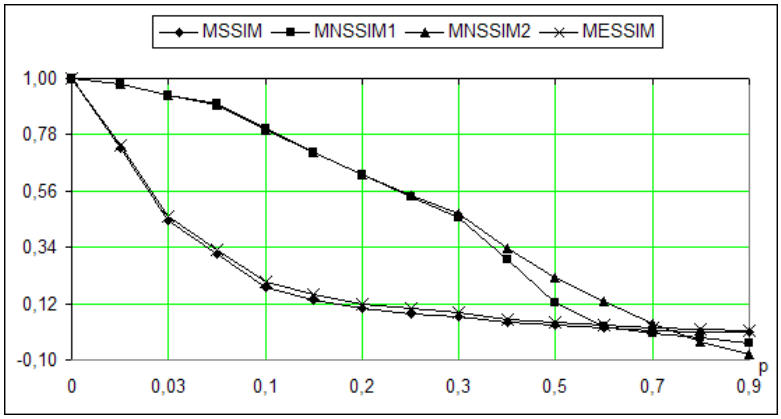


Рис.2 Зависимость значений критериев MSSIM, MNSSIM1, MNSSIM2, MESSIM от интенсивности импульсного шума для типа «соль/перец».

Исследование критериев при наличии импульсных помех

При исследовании влияния импульсных помех на качество восстановленного изображения использовались критерии качества MSSIM, MESSIM, MNSSIM1, и MNSSIM2. Импульсный шум задавался двумя характеристиками: интенсивностью, определяемой вероятностью появления шума в пикселе *p*, и типом шума, имеющим три вариации: "соль", "перец" и "соль/перец". В качестве тестового изображения использовалось изображение «Лена». На рис. 1-а) представлено изображение, искаженное импульсным шумом типа «соль/перец».

Как следует из рис. 1-а), критерии MNSSIM1 и MNSSIM2 при наличии импульсного шума дают значе-

предпочтение модифицированным непараметрическим критериям.

Исследование критериев при наличии блочности

Блочная структура изображения обычно образуется при сжатии и восстановлении изображения и зависит от степени сжатия. Чем больше степень сжатия, тем меньше энтропия сжатого изображения. При исследовании влияния блочности на качество изображения использовались критерии MSSIM, MESSIM. MNSSIM1, MNSSIM2 и алгоритмы сжатия GDCT [6,7] и JPEG. На рис. 3 представлена зависимость критериев MSSIM, MESSIM, MNSSIM1, MNSSIM2 от энтропии (в бит/пиксель) для изображения «Лена», сжатого и восстановленного с помощью алгоритма GDCT.

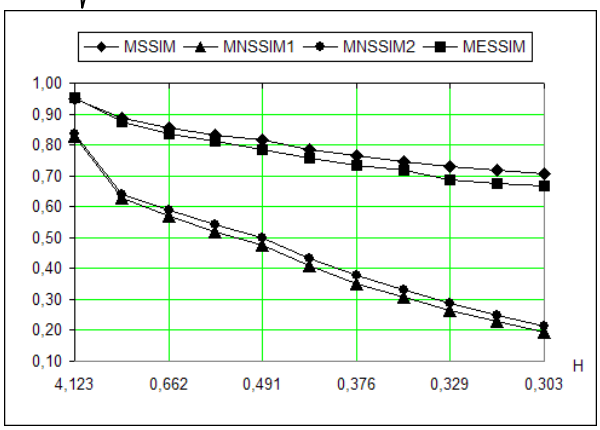


Рис. 3. Зависимость критериев MSSIM, MNSSIM1, MNSSIM2, MESSIM от энтропии.

Можно отметить, что модифицированные непараметрические критерии при малой блочности дают несколько заниженные значения качества, но при большом уровне блочности (высокой степени сжатия, или малой энтропии) более близки к зрительному восприятию человека по сравнению с критериям MSSIM и MESSIM.

Последний вывод иллюстрирует рис. 1-б), где представлено изображение «Лена», восстановленное после максимального сжатия с помощью JPEG.

Следует отметить также большой диапазон изменений критериев MNSSIM1, MNSSIM2 при изменении степени сжатия (энтропии), что говорит о чувствительности этих критериев к блочным искажениям.

Исследование критериев при наличии гауссовых помех

При исследовании влияния гауссовых помех на качество изображения гауссов шум со среднеквадратическим отклонением σ добавлялся в каждую компоненту цвета RGB. В качестве тестовых использовались изображения из базы изображений, предоставленных Laboratory for Image and Video Engineering (LIVE) университета в Остине, Техас. Из указанной базы данных были использованы и значения субъективной оценки качества изображений DMOS [3]. Для зашумленных гауссовым шумом изображений были рассчитаны значения критериев MSSIM, MESSIM, MNSSIM1, MNSSIM2 и ранговые коэффициенты корреляции Спирмена между ними и значениями DMOS. Результаты исследований для изображения «Parrots» представлены в таблице 1.

Как видно из таблицы, при гауссовом шуме значения и диапазон изменения критериев MSSIM и MESSIM выше, чем у критериев MNSSIM1 и MNSSIM2. Все критерии сильно коррелируют с DMOS, причем MSSIM и MESSIM коррелируют несколько больше.

Исследование структурного подобия разных изображений

В ходе проведения исследований был поставлен вопрос, в какой степени рассматриваемые критерии отслеживают структурное подобие изображений. С этой целью критерии MSSIM, MESSIM, MNSSIM1, MNSSIM2 были применены к совершенно разным изображениям. При этом выявилось еще одно преимущество непараметрических модификаций. При сравнении абсолютно разных изображений, имеющих одинаковые размеры, критерии MSSIM и MESSIM дают неоправданно большие значения, в то время как критерии MNSSIM1 и MNSSIM2 имеют значения, практически равные нулю. На рис. 4 в качестве примера представлены сравниваемые изображения и значения критериев MSSIM, MESSIM, MNSSIM1 и MNSSIM2.

Таблица 1.

σ	DMOS	MSSIM	MESSIM	MNSSIM1	MNSSIM2
1,000	68,727	0,027	0,051	0,012	0,017
0,129	47,039	0,252	0,326	0,152	0,208
0,063	38,931	0,518	0,609	0,314	0,359
0,031	28,506	0,788	0,850	0,520	0,541
Коэффициент корреляции Спирмена		**-0,955**	**-0,961**	**-0,929-**	**0,921**



Рис. 4. Сравнение двух изображений: «lighthouse» и «sailing 2». MSSIM = 0,339, MESSIM = 0,454, MNSSIM1 = -0,004, MNSSIM2 = -0,006.

Полученные данные свидетельствуют о том, что непараметрические модификации более соответствуют названию критерия структурного подобия и, может быть, они могли бы найти свое применение для определения смены сюжета в видеопоследовательности

Заключение

Проведенные исследования свидетельствуют о том, что предложенные непараметрические модификации критерия структурного подобия могут быть использованы для оценки качества изображений при импульсном шуме, при гауссовом шуме, при большой блочности изображений и, возможно, при обнаружении полной смены кадра в видеопоследовательности.

Авторы выражают благодарность доктору Х. Шейху за предоставленную базу данных, позволившую провести экспериментальные исследования предложенных критериев.

Литература

1. Z. Wang and A.C. Bovik, "A universal image quality index," *Signal Processing Letters, IEEE* , vol.9, no.3pp.81-84, Mar 2002.
2. Z. Wang, A. Bovik, H. Sheikh, E. Simoncelli, "Image Quality Assessment: Form Error Visibility to Structural Similarity", IEEE Transaction on Image Processing, vol. 13, №4, 2004, 600-612.
3. H.R. Sheikh, M.F. Sabir and A.C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms", *Image Processing, IEEE Transactions on,* vol. 15, no. 11, pp. 3440-3451, Nov. 2006.
4. G-H Chen, C-L Yang, L-M Po, S-L XIE, "Edge-based Structural Similarity for Image Quality Assessment", *Proceedings of the 2006 International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2006)*, Vol 2, Toulouse, France, May 2006, pp 933-936.
5. Кобзарь А.И. Прикладная математическая статистика. Для инженеров и научных работников. - М.: ФИЗМАТЛИТ, 2006.- 816 с.
6. Радченко Ю.С. Сравнительный алгоритм сжатия изображений на основе дискретного косинусного (DCT) и чебышевского (GDCT) преобразований / Ю.С. Радченко, Т.А. Радченко, А.В. Булыгин // Цифровая обработка сигналов, 2006.- №4.-С. 15-19.
7. Радченко Ю.С. Метод сжатия и восстановления изображений на основе быстрых чебышевских преобразований / Ю.С. Радченко // Автометрия. - 2002. - № 4. - С. 32-40.
8. Радченко Ю.С. Модификация индекса структурного подобия (MSSIM) на основе методов непараметрической статистики / Ю.С. Радченко, Т.А. Радченко, А.В. Булыгин // Труды X международной конференции и выставки «Цифровая обработка сигналов и ее применение» DSPA'2008, Москва, 2008. – т. II, с. 505-508.

Уважаемые коллеги!

Предлагаем вам принять участие в формировании тематических выпусков журнала «Цифровая обработка сигналов» и размещению рекламы продукции (услуг) Вашей фирмы на его страницах. В случае положительного решения просим представить в редакцию журнала Ваши предложения по плановому размещению информационных материалов и макет рекламы продукции (услуг) Вашей фирмы с указанием желаемого её месторасположения: обложка (2-я, 3-я или 4-я стр.), цветная внутренняя полоса (объем полосы).

В 2009 году планируется выпуск 4-х номеров журнала (тираж до 1000 экз.). Журнал распространяется по подписке через агентство «Роспечать» в России, СНГ и странах Балтии

(индекс 82185), а также на Выставках: «Цифровая обработка сигналов и ее применение – DSPA'2009», «ExpoElectronica», «ChipEXPO», «СвязьЭкспокомм», «ЭЛЕКТРОНИКА: компоненты, оборудование, технологии» (г. Москва) и др.

Размещение рекламы Вашей фирмы на страницах журнала «Цифровая обработка сигналов» на плановой основе (не менее 2-х полных или 4-х половинчатых рекламных полос в течение года) предоставит Вам следующие возможности и права:

1. Первоочередное право расположения рекламных материалов на всех обложках (кроме 1-й) и страницах журнала.
2. Публикация представленных Вами рабочих (рекламных) материалов (статей) объемом до 6 полос в каждом очередном номере (в счет оплаченной рекламы).
3. Установка баннера Вашего сайта (или логотипа вашей организации) на 1-й странице сайта журнала «Цифровая обработка сигналов» (www.dspa.ru) в течение всего года, что привлечет внимание к продукции (услугам) Вашей фирмы новых участников на рынке DSP-технологий (ежедневно фиксируется до 100 и более посещений сайта www.dspa.ru).
4. Предоставление до 10 экз. очередного выпуска журнала.

Ориентировочная стоимость рекламных услуг:
4-я (внешняя) страница цветной обложки - 20 тысяч рублей.
2-я и 3-я (внутренние) страницы цветной обложки - 13 тысяч рублей.
1\2 цветной внутренней полосы - 7 тысяч рублей.
1\2 черно-белой внутренней полосы – 1 тысяча рублей.

Ждем Ваших предложений.

С наилучшими пожеланиями, зам. главного редактора
д.т.н., профессор Витязев Владимир Викторович

Предложения прошу направлять по адресу: E-mail: tor@rgrta.ryazan.ru или info@dspa.ru

УДК 621.397

АНАЛИЗ ЭФФЕКТИВНОСТИ ПРИМЕНЕНИЯ ОРТОГОНАЛЬНЫХ ПРЕОБРАЗОВАНИЙ ДЛЯ КОДИРОВАНИЯ ДИСКРЕТНЫХ СИГНАЛОВ С КОРРЕЛИРОВАННЫМИ ОТСЧЕТАМИ

Умняшкин С.В.

Статья посвящена проблемам оценки эффективности использования различных ортогональных преобразований статических и динамических изображений. Предложен оригинальный критерий этой оценки.

Введение

Пусть $\mathbf{X}=(X_0,X_1,\ldots,X_{N-1})^T$ - некоторый случайный вектор (дискретный сигнал). Требуется найти такой способ эффективного цифрового кодирования вектора **X**, чтобы при заданной длине *L* (бит) двоичного кода полученный после декодирования вектор $\hat{\mathbf{X}}=(\hat{X}_0,\hat{X}_1,\ldots,\hat{X}_{N-1})^T$ представлял исходный с минимальной ошибкой $\varepsilon^2=M\left(\|\mathbf{X}-\hat{\mathbf{X}}\|^2\right)$. (Здесь и далее в качестве векторной нормы используется евклидова норма.)

Вследствие наличия межкомпонентных связей независимое покомпонентное статистическое кодирование отсчетов дискретного сигнала, следующее за их скалярным квантованием, порождает избыточные, неэффективные коды. Общий подход, который используется для повышения эффективности кодирования дискретных сигналов, состоит в предварительной обработке исходных данных с помощью обратимого дискретного преобразования с матрицей **W**, переводящего вектор **X** в некоторый вектор **Y**=**WX**, в котором зависимость между компонентами ослаблена. Тогда независимое покомпонентное кодирование вектора **Y**, а не вектора **X**, становится обоснованным. Вектор **Y** будем называть вектором трансформант, или обобщенным дискретным спектром.

Ограничим рассмотрение статистических зависимостей между отсчетами вектора **X** рамками корреляционной модели, считая, что для исходного дискретного сигнала **X** заданы вектор математических ожиданий компонент $\mathbf{m_X}=\mathrm{M}(\mathbf{X})=(m_{X_0},\ldots,m_{X_{N-1}})^T$ и ковариационная матрица $\mathbf{K_X}=(\mathrm{cov}(X_k,X_j))_{k,j=0}^{N-1}$. Тогда смысл использования дискретного преобразования для кодирования сигналов состоит в том, чтобы получить вектор трансформант **Y**=**WX** с некоррелированными компонентами. Этого можно добиться, если матрица **W**=**Wopt** составлена из собственных векторов матрицы $\mathbf{K_X}$ – соответствующее дискретное ортогональное преобразование называется преобразованием Карунена-Лоэва (или Хоттелинга). Ковариационная матрица **KY** вектора **Y** имеет в этом случае диагональный вид. Однако отсутствие быстрых алгоритмов вычисления и зависимость параметров оптимального преобразования Карунена-Лоэва от структуры матрицы $\mathbf{K_X}$ существенно ограничивают практические возможности его применения и вынуждают использовать другие преобразования. Возникает вопрос: насколько эффективно применение того или иного преобразования **Y**=**WX** по сравнению с оптимальным преобразованием Карунена-Лоэва? Как численно оценить эту эффективность для того, чтобы выбрать лучший из возможных вариантов обработки?

Энтропия случайной величины

Если источник сообщений ξ имеет N возможных состояний $\{\xi_0,\xi_1,\ldots,\xi_{N-1}\}$ и принимает их вне зависимости от предыдущего состояния с вероятностями $\{p_0, p_1,\ldots, p_{N-1}\}$, то двоичной энтропией источника согласно теоремы Шеннона о кодировании для канала без помех [1] является величина:

$$H(\xi)=-\sum_{k=0}^{N-1}p_k\log_2 p_k \quad \text{(бит)}. \qquad (1)$$

Энтропия (1) представляет собой минимально возможные средние битовые затраты R, которые необходимы для представления одного символа сообщения и является точной нижней оценкой для средней длины кода одного символа: $R\geq H$.

Следуя [1], рассмотрим возможность применения понятия энтропии к случайной величине непрерывного типа **X** (соответствующей непрерывному источнику информации), которая задана функцией плотности вероятности $f(x)$. Для упрощения выкладок положим, что случайная величина **X** может принимать значения из конечного диапазона, соответственно, функция плотности распределения вероятностей $f(x)$ отлична от нуля на конечном интервале. Положим, что $f(x)\neq 0$ непрерывна при $x\in[0,1)$, и

$$\int_0^1 f(x)dx=1.$$

Равномерно разобьем интервал [0,1) на N участков $\Delta_j=[ju,(j+1)u)$ длины $u=1/N$, $j=0,1,\ldots,N-1$. Случайной величине X, попавшей в интервал Δ_j, поставим в соответствие значение $x_j=u\left(j+\frac{1}{2}\right)$. Получаем из непрерывной величины X равномерно проквантованную с шагом u дискретную случайную величину $\widetilde{X}$, энтропия (1) которой равна:

$$H(\widetilde{X})=-\sum_{j=0}^{N-1}\int_{\Delta_j}f(x)dx\cdot\log\int_{\Delta_j}f(x)dx=-\sum_{j=0}^{N-1}f(\theta_j)|\Delta_j|\cdot\log\left(f(\theta_j)|\Delta_j|\right)=$$

$$= -\sum_{j=0}^{N-1} f(\theta_j) u \log_2 f(\theta_j) - \sum_{j=0}^{N-1} \int_{\Delta_j} f(x)dx \log_2 u =$$

$$= -\sum_{j=0}^{N-1} f(\theta_j) \log_2 f(\theta_j) \cdot u - \log_2 u \,, \quad (2)$$

где некоторая точка $\theta_j \in \Delta_j$. При этом математическое ожидание квадрата ошибки квантования $\varepsilon^2 = M((X-\hat{X})^2)$:

$$\varepsilon^2 = \sum_{j=0}^{N-1} \int_{\Delta_j} \left(x - \left(j+\tfrac{1}{2}\right)u\right)^2 f(x)dx =$$

$$= \sum_{j=0}^{N-1} f(\xi_j) \int_{\Delta_j} \left(x - \left(j+\tfrac{1}{2}\right)u\right)^2 dx = \frac{u^2}{12} \sum_{j=0}^{N-1} f(\xi_j) u \,, \quad (3)$$

где некоторая точка $\xi_j \in \Delta_j$.

Увеличивая число разбиений N, в пределе получим для энтропии (2):

$$\lim_{N\to\infty} H(\hat{X}) = \lim_{u\to 0} (\hat{X}) = -\int_a^b f(x) \log_2 f(x) dx - \lim_{u\to 0} \log_2 u \to \infty \quad (4)$$

(в предположении существования стоящего в выражении интеграла). При этом для ошибки (3):

$$\lim_{u\to 0} \frac{\varepsilon^2}{u^2/12} = \lim_{u\to 0} \sum_{j=0}^{N-1} f(\gamma_j) u = \int_0^1 f(x)dx = 1 \cdot$$

Таким образом, при стремлении шага (дискрета) u равномерного квантования к нулю ошибка квантования (3) есть бесконечно малая величина, $\varepsilon^2 \sim u^2/12$ для любой функции $f(x)$, т.е. независимо от закона распределения.

Число бит, необходимое для кодирования дискретной величины $\hat{X}$, с уменьшением шага квантования $u \to 0$ непрерывной величины X бесконечно возрастает: $H(\hat{X}) \sim -\log_2 u$. Однако, если имеются две случайных величины X и Y, определяемые на интервале $t \in [0,1)$ функциями плотности вероятности $f(t)$ и $g(t)$ соответственно, то при бесконечно малом шаге разбиения интервала [0,1) существует предел разности для значений энтропии (4) дискретных величин $\hat{X}$ и $\hat{Y}$ (в случае существования соответствующих интегралов):

$$\lim_{u\to 0} \left(H(\hat{X}) - H(\hat{Y}) \right) = -\int_0^1 f(t) \log_2 f(t) dt + \int_0^1 g(t) \log_2 g(t) dt \,, \quad (5)$$

при этом $\varepsilon^2(\hat{X}) \sim \varepsilon^2(\hat{Y}) \sim u^2/12$.

Для непрерывной случайной величины X с функцией плотности вероятности $f(x)$ (заданной в общем случае на всей числовой оси) величина

$$H(X) = -\int_{-\infty}^{+\infty} f(x) \log_2 f(x) dx \quad (6)$$

является дифференциальной энтропией [1]. Данное название подчеркивает тот факт, что реальный смысл несут именно разности значений (6): если $H(X) < H(Y)$, то при достаточно малом шаге квантования u ошибки квантования эквивалентны, $\varepsilon^2(\hat{X}) \approx \varepsilon^2(\hat{Y}) \approx u^2/12$, но, как следует из (5), для кодирования величины $\hat{Y}$ потребуется примерно на $H(Y)$-$H(X)$ большее число бит, чем необходимо для кодирования $\hat{X}$.

Критерий средней избыточной энтропии [4]

Рассмотрим вектор $\mathbf{Y}=\mathbf{W}\mathbf{X}$, полученный в результате дискретного ортогонального преобразования, для которого матрица обратного преобразования $\mathbf{W}^{-1}=\mathbf{W}^{\mathrm{T}}$. Пусть $f_k(m_k,\sigma_k,x)$ - функция плотности распределения вероятностей для Y_k k-ой компоненты вектора $\mathbf{Y}$, где m_k – математическое ожидание, σ_k – среднеквадратичное отклонение. Обозначив

$f_k^0(x) = f_k(0,1,x)$ и учитывая, что

$$f_k(m_k,\sigma_k,x) = \frac{1}{\sigma_k} f_k^0\left(\frac{x-m_k}{\sigma_k}\right), \quad \int_{-\infty}^{+\infty} f_k(m_k,\sigma_k,x)dx = 1 \,,$$

среднюю дифференциальную энтропию одной компоненты преобразованного вектора можно записать следующим образом:

$$H_{cp} = -\frac{1}{N} \sum_{k=0}^{N-1} \int_{-\infty}^{+\infty} f_k(m_k,\sigma_k,x) \log f_k(m_k,\sigma_k,x) dx =$$

$$= \frac{1}{N} \sum_{k=0}^{N-1} \log \sigma_k \int_{-\infty}^{+\infty} f_k(m_k,\sigma_k,x)\, dx -$$

$$- \frac{1}{N} \sum_{k=0}^{N-1} \int_{-\infty}^{+\infty} f_k^0\left(\frac{x-m_k}{\sigma_k}\right) \log f_k^0\left(\frac{x-m_k}{\sigma_k}\right) d\left(\frac{x}{\sigma_k}\right) =$$

$$= \frac{1}{N} \sum_{k=0}^{N-1} \log \sigma_k - \frac{1}{N} \sum_{k=0}^{N-1} \int_{-\infty}^{+\infty} f_k^0(x) \log f_k^0(x) dx \,. \quad (7)$$

Здесь и далее опущенное в обозначениях основание логарифма означает, что выбор его конкретного значения для построения изложения непринципиален.

Чем меньше средняя энтропия (7), тем эффективнее будет последующее независимое кодирование компонент преобразованного вектора.

Поскольку в рамках корреляционной модели законы распределения компонент вектора $\mathbf{X}$ неизвестны, то и точное определение вида функций плотности распределения $f_k^0(x)$ для компонент вектора $\mathbf{Y}$ также невозможно. Однако компоненты вектора $\mathbf{Y}$ представляют собой определенные взвешенные суммы из компонент вектора $\mathbf{X}$:

$$y_k = \sum_{j=0}^{N-1} w_{k,j} x_j \,,$$

где весовые коэффициенты представляют собой элементы матрицы преобразования

$$\mathbf{W} = \left\{ w_{k,j} \right\}_{k,j=0}^{N-1} \,,$$

поэтому с некоторым приближением можно допустить возможность применения центральной предельной теоремы (ЦПТ) и положить, что распределение каждой из компонент вектора $\mathbf{Y}$ подчиняется нормальному закону (отличия имеют место только в параметрах распределения). Хотя, строго говоря, для корректного использования ЦПТ необходима прежде всего статистическая независимость входящих в сумму случайных величин, именно гауссова плотность часто используется для описания законов распределения коэффициентов унитарных преобразований, в частности, при цифровой обработке изображений [2]. Используется также распределение Лапласа. Более точное описание законов распределения основано на обобщенном нормальном законе (см., например, [3]). Допущение об общем виде законов распределения компонент вектора трансформант $\mathbf{Y}$ для различных преобразований не является строгим, однако

часто применяется на практике, поэтому можно считать, что $\forall k = 0,1,\ldots,N-1$:

$$-\int_{-\infty}^{+\infty} f_k^0(x)\log f_k^0(x)\,dx \approx C.$$

Используем менее жесткое ограничение, а именно:

$$-\frac{1}{N}\sum_{k=0}^{N-1}\int_{-\infty}^{+\infty} f_k^0(x)\log f_k^0(x)\,dx \approx C,$$

где C - некоторая константа, не зависящая от вида используемого преобразования и его размерности. Среднюю энтропию (7) одной компоненты вектора трансформант можем записать тогда в следующем виде:

$$H_{cp} = \frac{1}{2N}\sum_{k=0}^{N-1}\log\sigma_k^2 + C. \qquad (8)$$

Необходимые в выражении (8) значения σ_k^2 дисперсий трансформант Y_k являются диагональными элементами матрицы $\mathbf{K_Y}=\mathbf{W}\mathbf{K_X}\mathbf{W}^T$ и в общем случае могут быть найдены по формуле:

$$\sigma_k^2 = \sum_{m=0}^{N-1} w_{k,m}\sum_{j=0}^{N-1} w_{k,j}\operatorname{cov}(x_m, x_j). \qquad (9)$$

Для оптимального преобразования Карунена-Лоэва расчет по формуле (9) упрощается, т.к. матрица $\mathbf{K_Y}$ имеет диагональный вид и

$$\prod_{k=0}^{N-1}\sigma_k^2 = \det\mathbf{K_Y} = \det\mathbf{W}\cdot\det\mathbf{K_X}\cdot\det\mathbf{W}^{-1} = \det\mathbf{K_X}.$$

Поэтому $\frac{1}{2N}\sum_{k=0}^{N-1}\log\sigma_k^2 = \frac{1}{2N}\log\prod_{k=0}^{N-1}\sigma_k^2 = \frac{1}{2N}\log\det\mathbf{K_X}$. (10)

Рассматривая среднюю безусловную энтропию (8) как характеристику декоррелирующих свойств ортогональных преобразований, естественно в качестве точки отсчёта принять значение (8) для оптимального преобразования Карунена-Лоэва и ввести в рассмотрение следующий параметр:

$$\Delta H(\mathbf{W},\mathbf{K_X}) = H_{cp}(\mathbf{W},\mathbf{K_X}) - H_{cp}(\mathbf{W}_{opt},\mathbf{K_X}),$$

или, с учетом (8), (10):

$$\Delta H(\mathbf{W},\mathbf{K_X}) = \frac{1}{2N}\log\left(\prod_{k=0}^{N-1}\sigma_k^2 \Big/ \det\mathbf{K_X}\right). \qquad (11)$$

Величина (11) характеризует избыточность данных в векторе $\mathbf{Y}$, связанную с неполной декорреляцией компонент (чем больше значение (11), тем меньше эффективность декоррелирующего преобразования с матрицей $\mathbf{W}$). Назовём данную величину *средней избыточной энтропией* [4].

Комментарии к определению (11)

1. Физический смысл величины. Решая задачу оптимального (в смысле минимизации суммарной квадратичной ошибки) распределения заданного количества бит для кодирования компонент случайного вектора, имеющего *N*-мерное нормальное распределение, Хуанг и Шультхайс [5] установили, что при определенных условиях вносимая в результате оптимального скалярного квантования по Максу суммарная квадратичная ошибка пропорциональна произведению дисперсий компонент, а именно:

$$M\left(\left\|\mathbf{Y}-\hat{\mathbf{Y}}\right\|^2\right) \sim N\left(\prod_{k=o}^{N-1}\sigma_k^2\right)^{1/N}, \qquad (12)$$

где σ_k^2 - дисперсия компоненты Y_k, а $\mathbf{Y}$ и $\hat{\mathbf{Y}}$ - оригинальный и скалярно проквантованный векторы, соответственно. В свою очередь, Дэвис и Носратиниа [6] показали, что для ортогональных преобразований при фиксированном значении размерности N величина, стоящая в правой части соотношения (12), принимает минимальное значение для оптимального преобразования Карунена-Лоэва. Поэтому для данного преобразования значение средней энтропии (8) минимально (см. также (10)), и значение величины (11) для любых ортогональных преобразований неотрицательно. Таким образом, чем меньше величина средней избыточной энтропии, тем ближе свойства используемого преобразования к оптимальным. Минимум величины (11), равный нулю, достигается для оптимального преобразования Карунена-Лоэва. Если в формуле (11) взять двоичное основание для логарифма, то физический смысл величины (11) определяется как избыточные (по сравнению с преобразованием Карунена-Лоэва) битовые затраты, приходящиеся в среднем на одну компоненту вектора $\mathbf{Y}$ при независимом эффективном кодировании компонент после равномерного квантования с достаточно малым шагом u.

2. О требовании к ортогональности матрицы $\mathbf{W}$. Отметим, что для анализа с использованием формулы (11) требование ортогональности исследуемого преобразования является принципиальным. Действительно, анализируя среднюю неопределенность компоненты вектора $\mathbf{Y}=\mathbf{W}\mathbf{X}$ по величине средней дифференциальной энтропии (7)-(8), мы фактически предполагали, что компоненты вектора $\mathbf{Y}$ равномерно проквантованы с шагом квантования $u\to 0$, при этом квадратичная ошибка (3), вносимая при квантовании в каждую компоненту вектора $\mathbf{Y}$: $\varepsilon^2 \sim u^2/12$, тогда суммарная ошибка от квантования компонент $\mathbf{Y} \xrightarrow{\text{квантование}} \hat{\mathbf{Y}}$ для любого преобразования:

$$\left\|\mathbf{Y}-\hat{\mathbf{Y}}\right\|^2 \sim Nu^2/12.$$

Эта ошибка всегда будет равна ошибке, внесенной в исходный вектор $\mathbf{X}$, $\left\|\mathbf{Y}-\hat{\mathbf{Y}}\right\| = \left\|\mathbf{X}-\hat{\mathbf{X}}\right\|$, только для ортогональных преобразований. Если анализируемые преобразования не являются ортогональными, то говорить об эквивалентности ошибки кодирования $\left\|\mathbf{X}-\hat{\mathbf{X}}\right\|^2 \sim Nu^2/12$ для разных преобразований нельзя, и сравнение декоррелирующей эффективности преобразований по параметру (11) является некорректным.

3. Эффективность декорреляции как степень концентрации энергии в векторе трансформант. Из формулы (11) следует, что самый «худший» вариант используемого преобразования при фиксированной размерности N дает максимальное значение параметра

$$P = \prod_{k=0}^{N-1}\sigma_k^2. \qquad (13)$$

Известно, что для всех ортогональных преобразований след ковариационной матрицы инвариантен, т.е.

$$\operatorname{trace}\mathbf{K_Y} = \sum_{k=0}^{N-1}\sigma_k^2 = \operatorname{trace}\mathbf{K_X} = \sum_{k=0}^{N-1}\sigma_{X_k}^2.$$

Тогда наложив ограничение

$$\sum_{k=0}^{N-1}\sigma_k^2 = const,$$

методом неопределенных множителей Лагранжа можно убедиться в том, что максимум выражения (13), а также и (11), имеет место, когда $\sigma_0^2 = \sigma_1^2 = \ldots = \sigma_{N-1}^2$. Оптимальность же применения преобразования следует понимать как способность последнего к концентрации основной «энергии» (определяемой следом ковариационной матрицы) исходного дискретного сигнала (вектора $\mathbf{X}$) в малом количестве коэффициентов-трансформант. В этом смысле задача выбора преобразования принимает вид задачи дискретной L_2-аппроксимации: найти такое преобразование, чтобы для заданного набора Ω отбрасываемых коэффициентов преобразования их суммарная дисперсия (фактически, энергия) $\sum_{k\in\Omega}\sigma_k^2$ была минимальной. Использование такого подхода менее универсально, вполне может дать разные результаты сравнения преобразований при различном количестве отбрасываемых коэффициентов. Тем не менее, подобное изучение свойств преобразований описано в литературе [7].

Анализ эффективности применения преобразований для дискретного марковского процесса первого порядка

Для примера приведем расчетные значения параметра (11) при использовании ряда дискретных ортогональных преобразований для обработки вектора $\mathbf{X}$, который представляет собой реализацию стационарного марковского процесса первого порядка и имеет ковариационную матрицу

$$\mathbf{K}_{\mathbf{X}}(\rho) = \begin{pmatrix} 1 & \rho & \ldots & \rho^{N-1} \\ \rho & 1 & \ldots & \rho^{N-2} \\ \ldots & \ldots & \ldots & \ldots \\ \rho^{N-1} & \rho^{N-2} & \ldots & 1 \end{pmatrix}, \qquad (14)$$

$0 \le \rho \le 1$. Данная модель имеет широкое применение и представляет большой интерес для практики, параметр ρ есть коэффициент корреляции соседних отсчетов дискретного случайного процесса [2,7].

Результаты расчета средней избыточной энтропии приведены на рис. 1, где рассмотрены дискретное косинусное преобразование (ДКП), слэнт-преобразование (СП), дискретное преобразование Хартли (ДПХ2), а также дискретные преобразования Адамара (ДПА) и Хаара (ДПХ1) (определения преобразований см. [2,7]). Выводы, которые следуют из анализа графиков рис. 1, подтверждают известные результаты оценки декоррелирующих свойств преобразований, полученные при использовании других методов, например по Пэрл [8]. В частности, критерий (11) подтверждает тот факт, что среди дискретных преобразований, имеющих быстрые алгоритмы вычислений, ДКП дает наиболее близкие к оптимальному преобразованию характеристики декорреляции для модели данных, которые описываются ковариационной матрицей (14).

Подробное сравнение результатов анализа декоррелирующей эффективности преобразований по предложенному критерию (11) и по критерию Пэрл [8] для других видов ковариационных матриц $\mathbf{K}_{\mathbf{X}}$ и других размерностей преобразований, а также более подробное рассмотрение вопросов численного нахождения величины (11) для ряда особых случаев можно найти в работах [4,9,10].

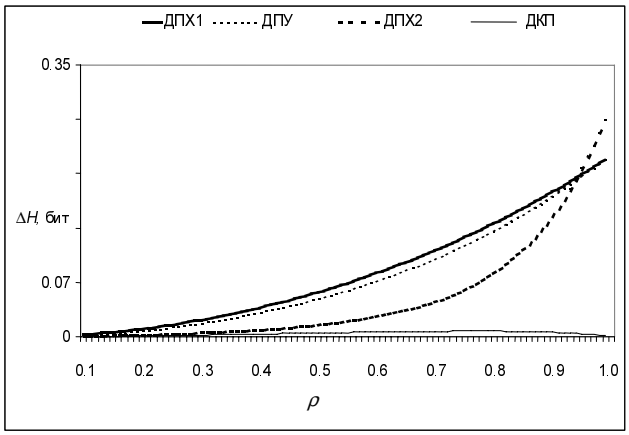


Рис. 1. Значения средней избыточной энтропии (11) при использовании некоторых ортогональных преобразований для обработки 16-ти компонентного вектора данных, имеющего ковариационную матрицу (14)

Применение предложенного критерия (11) во всех рассмотренных случаях дало оценки, которые хорошо согласуются с другими известными теоретическими и экспериментальными результатами.

Литература

1. Дмитриев В.И. Прикладная теория информации. - М.: Высш. шк., 1989. – 320 с.
2. Прэтт У. Цифровая обработка изображений: Пер. с англ. - М.: Мир, 1982. – Кн.1 и 2. – 312 и 480 с.
3. Antonini M., Barlaud M., Mathieu P., and Daubechies I., Image coding using wavelet transform// IEEE Trans. Image Proc. - Vol. 1. - №2, 1992. –P. 205-220.
4. Умняшкин С.В., Кочетков М.Е. Анализ эффективности использования дискретных ортогональных преобразований для цифрового кодирования коррелированных данных // Известия вузов. Электроника. - №6. - 1998. - С. 79-84.
5. Huang J.Y., Schultheiss P. M. Block quantization of correlated Gaussian random variables// IEEE Trans. Commumications. –1963. –V. CS-11 (Sept.) –P. 289-296.
6. Davis G., Nosratinia A. Wavelet-based Image Coding: An Overview // Applied and Computational Control, Signals and Circuits. –1998. –V.1. –№1. –P. 205-269.
7. Eliott D.F., Rao K.R. Fast transforms: algorithms, analyses, applications. - London: Academic Press inc., 1982. - 488 p.
8. Pearl J. On coding and filtering stationary signals by discrete Fourier transforms // IEEE Trans. Inf. Theory. - 1973. - Vol. IT-19. - P. 229-232.
9. Кочетков М.Е., Умняшкин С.В. О сравнении критериев для оценки эффективности декоррелирующих преобразований / М.: МГИЭТ (ТУ), 1998. - 34 с. - Деп. в ВИНИТИ 13.04.98, № 1069-В98.
10. Умняшкин С.В. Эффективность применения ортогональных преобразований для кодирования дискретных сигналов с точки зрения корреляционной теории // Интеллектуальные системы в производстве. - №1 – 2003. – С. 100-123.

УДК 681.775, 628.946.2

ЦИФРОВАЯ ОБРАБОТКА СИГНАЛОВ В РАСПРЕДЕЛЕННОМ ПРОЕКТОРЕ ДЛЯ ФОРМИРОВАНИЯ ЯРКИХ ИЗОБРАЖЕНИЙ НА ШИРОКОФОРМАТНЫХ ЭКРАНАХ

Торчигин А.В.

Анализируется возможность формирования яркого изображения на широкоформатном экране с помощью многих цифровых проекторов, каждый их которых формирует некоторый фрагмент изображения на основе синтеза в реальном масштабе времени воспроизведения информации, предварительно полученной при калибровке проектора. Показано, что использование проекторов, в которых лампа накаливания заменена на матрицу светодиодов, позволяет существенно увеличить яркость изображения и срок службы проектора.

Проблема формирования ярких изображений на экранах большого формата требует разрешения и поныне, поскольку яркость создаваемого кинопроектором изображения обратно пропорциональна его площади, что с неизбежностью вызывает необходимость увеличения мощности светового потока, проектируемого на экран. Эта мощность ограничена способностью кинопленки пропускать интенсивный свет, так как значительная доля светового потока поглощается кинопленкой, что приводит к ее разогреву.

Остается острой проблема получения мощного источника света с заданными спектральными и пространственными характеристиками. Срок службы таких источников составляет около 1000 часов, и при этом их характеристики изменяются во времени. Изменениям также подвержена сама кинопленка, качество которой ухудшается с каждым показом, а также при длительном хранении.

Проблемы ухудшения качества изображения из-за деградации кинопленки решены в современных цифровых LCD и DLP проекторах, в которых используется цифровой метод формирования изображений. Однако и в этих случаях сохраняется проблема ухудшения качества изображения во времени, так как изменяются характеристики формирующих изображение матриц, выходят из строя отдельные пикселы. Кроме того, в этих проекторах остаются проблемы, связанные с ограничениями на мощность светового потока.

В настоящее время источники света и средства для управления их яркостью достигли такого уровня, что целесообразно использовать принцип формирования изображения несколькими проекторами. При этом каждый проектор может быть предназначен для формирования определенного фрагмента общего изображения, и границы этих фрагментов перекрываются. В этом случае должна быть полностью решена проблема граничных областей, в которых изображение формируется двумя или большим количеством проекторов.

Для решения этой задачи в качестве источников света могут быть использованы яркие светодиоды, управляемые современными микропроцессорами.

В настоящее время технология изготовления светодиодов не достигла еще такого уровня, при котором было бы возможно изготовление нескольких миллионов светодиодов, по одному на каждый пиксел высококачественного изображения.

В отличие от ламп накаливания, светодиоды способны к высокочастотной модуляции. Частота модуляции 25 МГц вполне доступна для современных ярких светодиодов. Если обеспечить перемещение изображения каждого светодиода по экрану, то светящиеся линии, формируемые светодиодами, можно рассматривать как некие аналоги строк телевизионного изображения. Наработка на отказ у светодиодов на 2 порядка выше, чем у ламп накаливания. Что касается светового потока, то уже в течение нескольких лет коммерчески доступны светодиоды со световым потоком в 100 люмен [1]. Световой поток, создаваемый типичным проектором, составляет около 2000 люмен, а световой поток самого мощного в настоящее время проектора NEC STARUS NC2500S равен 20 000 люмен. Таким образом, требуется всего 20 и 200 светодиодов, чтобы создать соответственно типичный и рекордный световой поток современных проекторов. Количество используемых светодиодов в проекторе может быть значительно больше. Немаловажным является и то обстоятельство, что светоотдача светодиодов более чем в 10 раз выше, чем у ламп накаливания [2]. Это обстоятельство не только повышает КПД проектора, но и упрощает проблемы его охлаждения. Следует также заметить, что существуют светодиоды, излучающие свет в достаточно узком спектральном диапазоне, соответствующем красному, зеленому и синему спектральным компонентам Эта особенность обеспечивает существенные преимущества светодиодам по сравнению с LCD и DLP проекторами, в которых из белого света, излучаемого лампой накаливания, с помощью светофильтров выделяются требуемые спектральные составляющие. Таким образом, использование светодиодов в качестве управляемых источников света позволяет снять многие ограничения, присущие современным проекторам.

Однако все перечисленные возможности остались бы нереализованными, если бы в настоящее время не было средств, позволяющих оперативно определять и управлять в реальном масштабе времени яркостью светодиодов для получения изображения. Применение современных микропроцессоров позволяет достаточно просто решить возникающие при этом проблемы.

Пусть некоторая матрица светодиодов проектируется на светоотражающий экран с помощью обычного объектива, используемого в проекторах. Если сдвинуть матрицу светодиодов относительно объектива таким образом, что плоскость, в которой расположена матрица, осталась неизменной, то изображение светодиодов на экране также будет также сдвинуто. При этом фокусировка изображения не нарушится. В том случае, когда матрица светодиодов перемешается в своей плоскости по некоторой замкнутой траектории, изображение каждого светодиода на экране будет описывать некоторую замкнутую линию.

При модуляции параметров светодиода яркость различных участков этой линии будет различна. Для современных светодиодов и микроконтроллеров, позволяющих производить модуляцию яркости с частотой 25 Мгц, обеспечивается скорость перемещения элемента изображения светодиода на экране 100 м/с, а яркость может изменяться при смещении этого элемента на расстояние, равное 4мкм. Это существенно меньше, чем требуется. Если допустить, что смена кадров должна происходить с частотой 100 Гц, то за время показа одного кадра имеется возможность 250 000 раз изменить яркость излучения светодиода. Из этого не следует, что один светодиод может участвовать в формировании 250 000 пикселов, так как при модуляции с такой частотой пикселы не полностью разделены в пространстве. Количество пикселов *N*, которое может формироваться одним светодиодом, зависит количества светодиодов в проектируемой матрице и от соотношения длины и ширины замкнутой линии, засвечиваемой одним светодиодом. При этом вполне достижимы значения $N=10^4 - 10^5$.

Пусть в формировании общего изображения на экране 1 участвуют несколько проекторов с регулируемыми светодиодами 2, как показано на Рис.1. Пусть период движения светодиодов во всех проекторах одинаков и состоит из *N* временных интервалов. Допустим, что на короткое время в начале каждого временного интервала поочередно включается первый светодиод на первом устройстве. После этого в следующем периоде включается второй светодиод на первом устройстве и т.д. Затем такая же процедура повторяется со светодиодами на втором устройстве, на третьем устройстве и т.д.

Формируемое светодиодом изображение с помощью цифровой видеокамеры 3 с максимально высоким пространственным и спектральным разрешением передается в персональный компьютер 4, который в рассматриваемом случае выступает в качестве HOST машины многопроцессорной вычислительной системы, включающей в себя микропроцессоры, управляющие яркостью светодиодов в каждом проекторе. В HOST машине полученное с видеокамеры изображение анализируется и составляется таблица, содержащая координаты создаваемого светодиодом пятна на экране в каждый *q*-ый момент времени ($q=0, 1, \ldots N-1$). Кроме того, таблица содержит такие параметры пятна, как его интенсивность и цвет.

Пусть теперь требуется с помощью указанных устройств создать изображение прямоугольной картины, имеющей *w* пикселов по ширине и *h* пикселов по высоте. Пусть картина представлена в BMP формате, где каждый пиксел характеризуется 3 байтами, определяющими интенсивность соответственно красного, зеленого и синего цветов.

Центр картины разместим в центре области, засвечиваемой всеми включенными светодиодами во всех проекторах. Выберем масштаб картины таким, чтобы она полностью попадала в область, засвечиваемую всеми проекторами, когда в каждом из них включены все имеющиеся светодиоды. Пусть суммарное количество светодиодов во всех проекторах равно K. Пронумеруем последовательно эти светодиоды от 0 до $K-1$.

Далее, на основе собранной в HOST машине информации поочередно для каждого пиксела картины на экране составляется таблица, содержащая следующую информацию:

- координаты пиксела (номер пиксела в строке $0<=i<=w-1$, номер строки в формируемом изображении $0<=j<=h-1$);
- номер светодиода $0<=p<=K-1$ и номер временного интервала $0<=q<=N-1$, когда пятно, создаваемое этим светодиодом, попадает на рассматриваемый пиксел;
- яркость и цвет этого пятна.

В том случае, если используются светодиоды, цвет которых не в точности совпадает с красным, зеленым и синим цветом, эта информация содержится в 3 байтах, в которых указывается интенсивность каждого из 3 цветов точно так же, как в 3 байтах BMP файла.

Далее необходимо провести проверку того, что у каждого пиксела может быть обеспечена необходимая яркость каждого их трех цветов. С этой целью по полученной для каждого пиксела информации составляются три матрицы *Rij, Gij, Bij,* элементы которых равны суммарной интенсивности, создаваемой всеми проекторами соответственно красными, зелеными и синими светодиодами в пикселе с координатами i,j. После этого определяются минимальные *rmin*, *gmin*, *bmin* и максимальные элементы *rmax*, *gmax*, *bmax* в каждой из указанных матриц. Ситуация, при которой хотя одно из отношений *rmin*/*rmax*, *gmin*/*gmax*, *bmin*/*bmax* близко к нулю, свидетельствует о том, что при заданном расположении проекторов для некоторых пикселов не может быть обеспечена необходимая яркость.

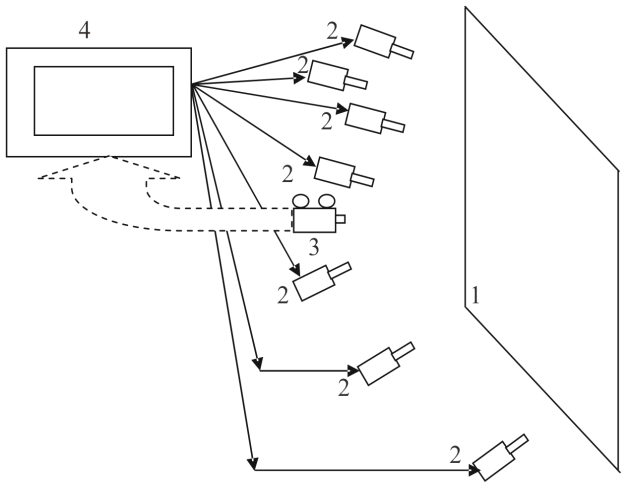


Рис.1. Схема формирования на экране одного изображения многими проекторами

В этом случае требуется скорректировать расположение проекторов и добиться ситуации, при которой минимальное из указанных соотношений больше некоторого предела L. Этот предел характеризует степень равномерности засветки экрана всеми проекторами и определяет результирующую яркость создаваемого изображения.

Действительно, при формировании произвольного изображения может оказаться, что пиксел, для которого проекторы обеспечивают минимальную интенсивность, должен иметь максимальную яркость. В этом случае в формируемом изображении наиболее яркий пиксел будет иметь интенсивность, ослабленную в L раз по сравнению с максимально возможной. Таким образом, как и следовало ожидать, равномерность засветки экрана всеми проекторами весьма важна, так как непосредственно влияет на яркость изображения. Заметим, что в отличие от обычных проекторов, где неравномерность засветки экрана приводит к нарушению пропорциональности в интенсивности отдельных фрагментов картины, в рассматриваемом случае такая пропорциональность сохраняется. Уменьшается лишь яркость картины. Для сохранения пропорций между яркостью различных пикселов достаточно допустить, что яркость найденного пиксела с минимальной освещенностью характеризуется 3 байтами с значениями 255, 255, 255. В этом случае может быть реализована соответствующая яркость любого пиксела изображения, характеризуемого байтами со значениями r, g, b $(0<=r<=255,\ 0<=g<=255,\ 0<=b<=255)$. Действительно, для этого достаточно лишь выбрать соответствующим образом доли временных интервалов, в течение которой должны быть включены соответствующие светодиоды.

Рассмотрим более подробно алгоритм формирования управляющих последовательностей для каждого светодиода в проекторе. Пусть требуется сформировать изображение, задаваемое матрицей A_{ij} $(i=0,1..w-1;\ j=0,1..h-1)$, где w и h количество пикселов в изображении соответственно по горизонтали и вертикали.

$$\sum_{p=0}^{K-1}\sum_{q=0}^{N-1}\alpha_{pq}I_{ij}^{pq}=A_{ij}(i=0,1..w-1;\ j=0,1..h-1)$$

Элемент матрицы A_{ij} представляет собой 24-разрядное целое, определяющее интенсивность красного, зеленого и синего цветов. Эта информация может быть получена из файла типа BMP, задающего формируемое изображение. На основе этой информации должно быть получено K последовательностей T_q, определяющих время, на которое включается светодиод с номером p $(p=0,\ 1..K-1)$ во временном интервале с номером q $(q=0,1..\ N-1)$. Эти K последовательностей могут быть представлены в виде матрицы T_{pq}. Значение элемента матрицы T_{pq} может изменяться в пределах от 0 до 1 и определяет долю временного интервала, в течение которой должен быть включен светодиод с номером p во временном интервале с номером q.

В том случае, когда изображение формируется с помощью красных, зеленых и голубых светодиодов, задача упрощается и сводится к трем одинаковым более простым задачам по формированию красной, зеленой и голубой составляющих общего изображения. В этом случае 3 байта элемента матрицы A_{ij} содержат лишь один отличный от нуля байт. Заметим, что в рассматриваемом случае формирование заданного цветного изображения возможно и при использовании не совсем чистых цветов, то есть при использовании светодиодов, спектр излучения которых характеризуется 3-мя отличными от нуля байтами.

Рассмотрим частный случай, при котором $K*N>>w*h$. Это условие означает, что количество потенциально возможных пикселов, которые могут быть сформированы распределенным проектором, существенно больше количества пикселов, заданных в BMP файле.

При калибровке проектора для каждого p-го светодиода и q –го временного интервала на основе измерений интенсивности пятна в различных участках экрана в этом временном интервале составляется матрица I_{ij}^{pq}, которая задает интенсивность света в квадрате на экране с координатами i,j (квадрат с номером i в j-ой строке), создаваемого светодиодом с номером p во временном интервале с номером q. Сторона квадрата e определяется соотношением $e=W/w$, где W – ширина изображения на экране. При калибровке проектора создается $K*N$ таких матриц. В рассматриваемом частном случае размеры пятна, создаваемого одним светодиодом на экране, существенно меньше размеров квадрата. Поэтому не вызовет большой погрешности предположение, что пятно полностью находится в некотором квадрате. Из этого предположения следует, что в матрице I_{ij}^{pq} лишь один элемент отличен от нуля и равен 1. Таким образом, задача сводится к нахождению линейной комбинации матриц I_{ij}^{pq}, которая дает матрицу A_{ij}. Иными словами, задача сводится к решению следующей системы из $w*h$ линейных уравнений, содержащей $K*N$ неизвестных.

Вообще говоря, эта система имеет множество решений, так как количество неизвестных α_{pq} больше количества уравнений. Однако на решения накладываются условия $0<=\alpha_{pq}<=1$. Для рассматриваемого частного случая эти условия выполнимы и решение этой системы находится простым перебором. Для каждого элемента матрицы A_{ij} просматриваем аналогичные элементы во всех $K*N$ матрицах I_{ij}^{pq} и выбираем матрицы, в которых эти элементы отличны от нуля. В предположении, что изображение хотя бы одного светодиода на экране в некоторый временной интервал находится в квадрате с координатами i,j, получим, что такие матрицы существуют. В том случае, если найдена лишь одна матрица, коэффициент α_{pq} определяется соотношением

$\alpha_{pq}=A_{ij}/255.$

Если найдено несколько матриц, то существует множество решений. Простейшим из них является решение, при котором коэффициенты α_{pq} у всех матриц, кроме первой, равны нулю. Коэффициент у первой матрицы определяется указанным соотношением. Найденные коэффициенты α_{pq} являются соответствующими элементами искомой матрицы T_{pq}. Значения элементов матрицы находятся в интервале от 0 до 1 и определяют долю q-го временного интервала, в течение которого должен быть включен p-ый светодиод. Заметим, что степень соответствия полученной информации реально определяется параметрами ви-

деокамеры, используемой для сбора исходной информации. По этой причине необходимо использовать высококачественную видеокамеру.

Строки матрицы T_{pq} рассылаются в микропроцессоры проекторов. При этом в каждый микропроцессор рассылаются лишь строки с номерами, соответствующими номерам светодиодов, управляемых этим микропроцессором. Алгоритм работы микропроцессора реализует включение на полную яркость светодиод с номером p во временной интервал с номером q на время, определяемое значением элемента T_{pq}. Современные 32-разрядные весьма дешевые микропроцессоры [4] в состоянии одновременно управлять яркостью около сотни светодиодов. При необходимости в одном проекторе может быть использовано несколько микропроцессоров, каждый из которых управляет отведенной для него группой светодиодов. Проблема управления яркостью светодиодов полностью решена в уличных крупноформатных экранах. В настоящее время несколькими фирмами выпускаются микросхемы, предназначенные значительно упростить управление яркостью светодиодов [5].

Заметим, что добавление новых проекторов сводится только к изменению параметра K. Это обстоятельство позволяет рассматривать множество проекторов как один распределенный проектор, автоматически настраиваемый на показ общего изображения в зависимости от характеристик и взаимного расположения компонентов такого проектора. Так как в описанном алгоритме отсутствует понятие отдельного проектора и все проекторы рассматриваются в рамках одного распределенного проектора, имеющего K светодиодов, то исчезает понятие границы между отдельными проекторами. Изображение в областях, засвечиваемых несколькими проекторами, формируется этими проекторами совместно.

Разумеется, описанный алгоритм может быть оптимизирован по многим направлениям. Однако и без оптимизации он вполне работоспособен. Как и в традиционных проекторах, перед показом изображений проекторы должны быть настроены. Настройка традиционных проекторов сводится к ориентации проектора в пространстве и фокусировке изображения. Дальнейшая настройка всей системы производится в автоматическом режиме при выполнении соответствующей программы один раз перед началом показа.

Заметим, что для точного отображения цветов, указанных в BMP-файле, не обязательно использовать светодиоды чисто красного, зеленого и синего цветов. Возможны небольшие отклонения цвета светодиодов от стандарта, которые могут быть учтены в процессе формирования изображений. Действительно, пусть требуется сформировать пиксел, в котором интенсивность красного, зеленого и синего цветов определяется соответственно параметрами $\{r, g, b\}$. Пусть для формирования заданной палитры используются красные, зеленые и синие светодиоды, спектр излучения которых определяется соответственно следующими выражениями $\{R, \Delta R_G, \Delta R_B\}$, $\{G, \Delta G_R, \Delta G_B\}$, $\{B, \Delta B_R, \Delta B_G\}$, где $R+\Delta R_G+\Delta R_B=1$, $G+\Delta G_R+\Delta G_B=1$, $B+\Delta B_R+\Delta B_G=1$, $\Delta R_G<<1$, $\Delta R_B<<1$, $\Delta G_R<<1$, $\Delta G_B<<1$, $\Delta B_R<<1$, $\Delta B_G<<1$. Интенсивности указанных светодиодов $\{r',g',b'\}$, необходимые для получения заданной палитры, могут быть определены из следующей системы уравнений

$$r'R+g'\Delta G_R+b'\Delta B_R=r$$
$$r'\Delta R_G+g'G+b'\Delta B_G=g$$
$$r'\Delta R_B+g'\Delta G_B+b'B=b$$

Решая эту систему уравнений с учетом указанных ограничений, получим следующие выражения для интенсивности светодиодов

$$r'=r(1+\Delta R_G+\Delta R_B)-g\Delta G_R-b\Delta B_R)$$
$$g'=g(1+\Delta G_R+\Delta G_B)-r\Delta R_G-b\Delta B_G)$$
$$b'=b(1+\Delta B_R+\Delta B_G)-r\Delta R_B-g\Delta G_B)$$

Таким образом, при рассматриваемом подходе цветопередача может быть настроена программно.

Дополнительно к привлекательным свойствам проекторов на основе светодиодов можно добавить следующие моменты:

- отсутствует необходимость в точной ориентации проекторов;
- для получения необходимой яркости изображения количество проекторов может, как увеличиваться, так и уменьшаться;
- исключаются проблемы традиционных проекторов, связанные с искажениями, вносимыми оптикой, такие как астигматизм, кома, сферическая и хроматическая аберрации, бочкообразные и подушкообразные искажения и т. п. [6];
- устраняются эффекты, связанные со старением оборудования, так как при настройке проекторов эти обстоятельства автоматически принимаются во внимание;
- возможно точное воспроизведение цветов в изображении;
- настройка проекторов с помощью программного обеспечения гораздо быстрее, дешевле, точнее, гибче и оперативнее, чем настройка аппаратуры с помощью механических регуляторов.

Таким образом, появление ярких светодиодов и дешевых микропроцессоров, обеспечивающих в реальном времени модуляцию яркости многих светодиодов, позволяет по-новому подойти к решению вековой проблемы – увеличению яркости изображений на широкоформатных экранах.

Литература

1. Кононов В. Рекордно яркие LED от Nichia http://www.3dnews.ru/news/rekordno_yarkie_led_ot_nichia/
2. Бурняшев А.. Современные мощные светодиоды и их оптика // Современная электроника. – 2006.-№1.- С. 24-27.
3. Мощные светодиоды различных производителей:
   www.leds.ru
   www.lumileds.com/products
   www.osram-os.com/goldendragon
   www.osram.ru/index.php
   www.upec.com
   www.laminaceramics.com
4. Редькин П. Микроконтроллеры ARM7 семейства LPC2000. –М.: Додэка-XXI, 2007 450 с.
5. Драйверы для светодиодов www.mblock.com
6. Ландсберг Г.С.. Оптика. М.:, Наука, 1976.

УДК 539.25: 620.187

СПОСОБ ФОРМИРОВАНИЯ ИЗ 2D 3D ИЗОБРАЖЕНИЯ В СКАНИРУЮЩЕЙ МИКРОСКОПИИ

Григоров И.Г., Строшков В.П., Зайнулин Ю.Г.

Описана методика формирования трехмерного изображения в цифровой форме реального физического объекта, путем получения микрорельефа его поверхности методом растровой электронной микроскопии и коррекцией данных трехмерного преобразования методами сканирующей зондовой микроскопии.

Возможности современных вычислительных средств и методов машинной графики стимулируют развитие новых направлений в разработке методов формирования “реалистичных изображений”. Речь идет об изображениях, которые можно наблюдать визуально с помощью оптических или оптоэлектронных устройств. Необходимость в создании такого типа изображений возникает в тех случаях, когда это может быть использовано для оценки качества сканирования трехмерного объекта, либо для получения дополнительной информации о пространственных характеристиках объекта. В этом случае формирование трехмерного (3-D) изображения должно обеспечивать передачу всей совокупности особенностей объекта: объемность, расположение деталей, полутона, текстура поверхности и т. д. [1].

Применительно к микроскопии методика построения 3-D изображения микрообъекта имеет свои особенности относительно методов машинной графики, обусловленные спецификой процесса формирования изображения в микроскопе.

Основным отличием является то обстоятельство, что для систем машинной графики источником входной информации являются не сами физические объекты, а их математические модели. Общим же для всех систем, формирующих 3-D изображения, является отображение объекта в виде синтезированного 3-D изображения на экране дисплея. При этом обрабатывающие программы должны передавать ощущение глубины пространства, пространственной формы и структуры объектов. Определяющим при этом является уровень соответствия (подобия) синтезируемого изображения оригиналу.

К наиболее наглядным и прямым методам визуализации 3-D параметров микро- и нанообъектов, относятся методы растровой электронной микроскопии (РЭМ) и сканирующей зондовой микроскопии (СЗМ): туннельной (СТМ) и атомно-силовой (АСМ) [2, 3]. Методы СЗМ, в отличие от РЭМ, позволяют непосредственно в процессе сканирования поверхности исследуемого объекта создавать ее трехмерный образ, однако диапазон измеряемого перепада высот (± 1 мкм), а также размер поля сканирования (менее 50x50 мкм^2), существенно меньше чем у РЭМ, для которого они составляют сотни микрон.

Для того чтобы получить наиболее полную информацию об исследуемом объекте с помощью РЭМ, необходимо иметь четкое представление о процессе формирования изображения, включая:

1 - принцип сканирования, используемый для построения изображения;

2 - природу механизмов формирования контраста, возникающих из-за взаимодействия электронного пучка с образцом;

3 - характеристики регистрирующих детекторов и их выходных сигналов;

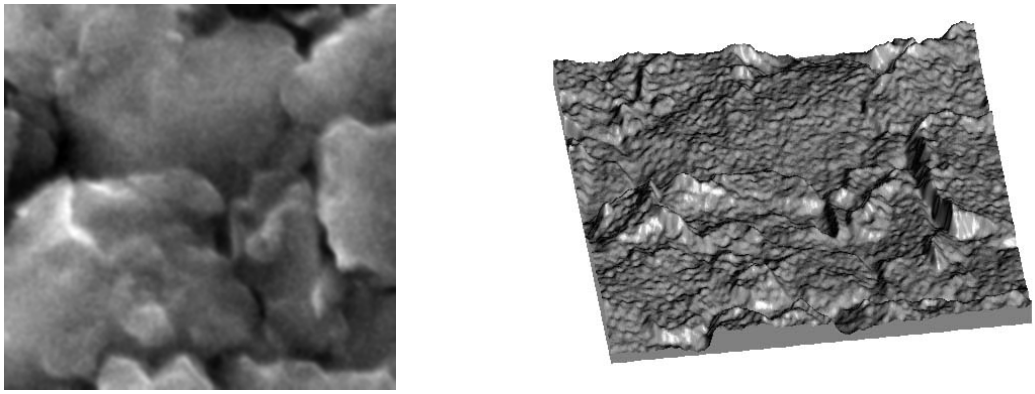
4 - влияние на качество окончательного изображения предварительной обработки сигналов.

Кажущаяся простота интерпретации изображений объектов с шероховатой поверхностью, понятная даже для людей, незнакомых с работой РЭМ, может зачастую содержать гораздо больше информации, чем видит глаз. Поэтому, чтобы быть уверенным, что изображение правильно сформировано и синтезировано с помощью компьютера, необходимо иметь четкое представление о процессе его формирования [4].

Электронная оптика РЭМ позволяет сформировать электронный пучок, который характеризуется тремя параметрами: ток пучка ($10^{-12} \div 10^{-6}$ А), диаметр пучка (1 нм ÷1 мкм) и угол схождения ($10^{-4} \div 10^{-2}$ ср.). Соотношение между ними дает уравнение для яркости видеосигнала. Контрастность видеосигнала можно определить, как разность между максимальным и минимальным сигналами яркости, деленную на максимальный сигнал яркости.

При сканировании электронного пучка по поверхности образца внутри области взаимодействия происходит как упругое, так и неупругое рассеяние, в результате чего в детекторах возникают сигналы за счет упругих (отраженных) вторичных и поглощенных электронов, характеристического и непрерывного рентгеновского излучения, катодолюминесцентного излучения [4-7]. Детекторы и видеосистема РЭМ преобразует эти сигналы в пропорциональный сигнал яркости на экране монитора. Поэтому на экране можно наблюдать только монохромное (одноцветное) изображение в градациях “серого”. Следовательно, синтезированное изображение, полученное с помощью РЭМ, должно соответствовать оригиналу только, как модель одноцветного зрения [8].

Видеосистема РЭМ синтезирует двухмерное изображение, в котором теряется информация, определяемая бинокулярным зрением. Она способна воспроизводить лишь условия монокулярного наблюдения. Объемность изображения оригинала воспринимаются благодаря характеру теней, изменению тона, загораживанию одних объектов другими и линейной перспективе. Существенное значение для восприятия объема и пространства изображения, полученного с помощью РЭМ, имеет опыт наблюдателя, благодаря которому объемная структура может быть “достроена”.



а) б)

Рис. 1. Пример формирования трехмерного изображения микрорельефа поверхности:
а) 2-D изображение 12х12 мкм поверхности излома сплава, РЭМ, увеличение 5000;
б) 3-D изображение поверхности излома сплава с тем же увеличением, сформированное по его 2-D изображению.

Максимальное приближение к оригиналу можно обеспечить, если математически создать модель съемки, т. е. обрабатывающая программа должна учесть расположение объекта относительно первичного электронного пучка и регистрирующих детекторов, геометрическую форму объектов, их взаимное положение, условия энергетического подобия, размер и положение теней и т. д. [1].

В данной работе предлагается методика комплексного использования РЭМ и СЗМ (СТМ и АСМ) для формирования 3-D изображения поверхности, исследуемых с большой протяженностью (>100 мкм) и шероховатостью (>1мкм) объектов.

Первоначально рассмотрим способ моделирования 3-D изображения, полученного методом РЭМ. Для формирования 2-D изображения использовался парносимметричный детектор РЭМ BS-301 Tesla в режиме A+B. Контрастность и яркость такого изображения пропорциональны изменению высот деталей рельефа относительно базового уровня, что соответствует фронтально освещенному объекту оптического изображения (рис. 1, а) [4-7]. Организация цифрового образа двумерного (2-D) изображения осуществлялась при помощи контроллера ввода изображения с РЭМ в персональный компьютер (ПК) [9]. Данный контроллер разработан в Институте химии твердого тела УрО РАН и позволяет вводить в ПК изображение с РЭМ за один кадр сканирования (оптимальный размер кадра 1900х1500 пиксель). Для создания 3-D изображения (рис. 1, б), введенного в ПК с РЭМ 2-D, дополнительно использовались сканирующий зондовый микроскоп СММ-2000Т и пакет программ Scan Master (www.zproton.ru).

Так как поле сканирования СЗМ было значительно меньше поля сканирования РЭМ, то, первоначально, для получения точных количественных характеристик шероховатости поверхности определялись среднеквадратичная шероховатость (R_q) и максимальная высота профиля (R_{max}) 3-D изображений, полученных на СЗМ в наиболее характерных участках исследуемой поверхности.

В дальнейшем, критерием для выбора увеличения при съемке на РЭМ являлось получение «представительного» вида изображения поверхности образца. Иначе говоря, увеличение выбиралось таким образом, чтобы хорошо были видны особенности микрорельефа поверхности, а характер их расположения на поверхности был бы равномерен и соответствовал среднему удельному значению их плотности относительно площади поверхности.

Программа Scan Mast позволяет преобразовывать 2-D изображение в цифровом формате BMP в градациях серого (монохромное изображение) в 3-D изображение, если предварительно заданы максимальные параметры (x, y, z) трехмерного образа изображения.

Параметры x и y задавались увеличением при съемке на РЭМ, а для определения параметра z использовались данные R_{max}, полученные на СЗМ. Кроме того, данные R_q и R_{max} использовались для корректировки при преобразовании 2-D изображения с РЭМ в 3-D программой Scan Master путем сравнения и подгонки параметров, получаемых при данном преобразовании с данными СЗМ анализа шероховатости образца [10].

Литература

1. Иванов В. П. Трехмерная компьютерная графика / В. П. Иванов, А. С. Батраков. М.: Радио и связь, 1995. 224 с.
2. Барндо Д. Микроструктура материалов. Методы исследования и контроля / Д. Барндо, У. Каплан. М.: Техносфера, 2004. 384 с.
3. Миронов В. Основы сканирующей зондовой микроскопии. М.: Техносфера, 2004. 144 с.
4. Гоулдстейн Дж. Растровая электронная микроскопия и рентгеновский микроанализ / Дж. Гоулдстейн, Д. Ньбери, П. Эчлин и др.: В 2 т. М.: Мир, 1984. Т. 1. 303 с.
5. Кальнер В. Д. Практика микрозондовых методов исследования металлов и сплавов / В. Д. Кальнер, А. Г. Зильберман. М.: Металлургия, 1981. 216 с.
6. Физическое металловедение / Под ред. Р. У. Кана и П. Хаазена: В 3 т. М.: Металлургия, 1987. Т. 1. 640 с.
7. Гаранин В. К. Применение электронно-зондовых приборов для изучения минерального вещества / В. К. Гаранин, Г. П. Кудрявцева. М.: Недра, 1983. 216 с.
8. Прэтт У. Цифровая обработка изображений: В 2 т. М.: Мир, 1982. Т.1. 312 с.
9. Григоров И.Г. Изучение особенностей микроструктуры и физико-механических характеристик сплавов на основе карбонитрида титана методом цифровой растровой микроскопии. Автореф. дис. … канд. хим. наук. Екатеринбург, 2002. 31 с.
10. Строшков В.П. Высокоточное электрохимическое формообразование сложнопрофильного инструмента для деталей машин / В.П. Строшков, В.А. Пшеничников, В.Л. Кожевников. Екатеринбург: Изд. УрО РАН, 2005. 212 с.

УДК 621.397.2

ОБРАБОТКА ИЗОБРАЖЕНИЙ ДВУМЕРНЫМИ НЕРЕКУРСИВНЫМИ ЦИФРОВЫМИ ФИЛЬТРАМИ ВТОРОГО ПОРЯДКА

Приоров А.Л.

Определены области существования двумерных нерекурсивных цифровых фильтров нижних и верхних частот второго порядка с симметричными коэффициентами и монотонными амплитудно-частотными характеристиками. Установлено максимальное подавление сигнала, возможное в фильтрах такого типа, получено уравнение линии среза. Приведены примеры применения таких фильтров для обработки искаженных изображений.

Введение

Существует множество областей, в которых используется цифровая обработка двумерных массивов данных. Это, например, обработка изображений, геофизические исследования земной коры, компьютерная томография в медицине и множество других важных научно-практических задач [1].

Хотя двумерные нерекурсивные цифровые системы для достижения тех же частотных свойств, что и рекурсивные, требуют большего объема вычислений, но они имеют и свои преимущества, к которым следует отнести их устойчивость при любых значениях своих параметров, а также возможность реализации линейной фазовой характеристики, что весьма важно при обработке изображений.

Частотные свойства одномерных нерекурсивных цифровых систем второго порядка хорошо изучены. В зависимости от значений двух коэффициентов этой системы, она может реализовать фильтры верхних или нижних частот, а также полосовой или режекторный фильтры. В тоже время частотные свойства двумерных фильтров изучены недостаточно и требуют дальнейших исследований.

Исследование частотных свойств двумерных нерекурсивных фильтров

Для исследования частотных свойств двумерных цифровых фильтров верхних и нижних частот второго порядка с симметричными коэффициентами использована методика, впервые предложенная в [2] для двумерной рекурсивной цифровой системы первого порядка. Структурная схема двумерной нерекурсивной цифровой системы второго порядка с симметричными коэффициентами приведена на рис. 1. Здесь $a = a_{10} = a_{01}$, $b = a_{20} = a_{02}$, $c = a_{11}$.

Частотная характеристика такой системы имеет вид:

$$H(e^{j\omega_1}, e^{j\omega_2}) = 1 + a\left(e^{-j\omega_1} + e^{-j\omega_2}\right) + b\left(e^{-j2\omega_1} + e^{-j2\omega_2}\right) + ce^{-j(\omega_1+\omega_2)}. \quad (1)$$

Квадрат модуля частотной характеристики (1) в этом случае запишется следующим образом:

$$\left|H(e^{j\omega_1}, e^{j\omega_2})\right|^2 = H(\omega_1, \omega_2) = (1 + a(\cos(\omega_1) + \cos(\omega_2)) + b(\cos(2\omega_1) + \cos(2\omega_2)) + c\cos(\omega_1 + \omega_2))^2 + (a(\sin(\omega_1) + \sin(\omega_2)) + c\sin(\omega_1 + \omega_2) + b(\sin(2\omega_1) + \sin(2\omega_2)))^2. \quad (2)$$

Из условий симметрии коэффициентов достаточно рассмотреть поведение функции $H(\omega_1, \omega_2)$ на осях ω_1 и ω_2, а также на диагоналях $\omega_1 = \omega_2$ и $\omega_1 = -\omega_2$ в верхней полуплоскости частотной плоскости (ω_1, ω_2).

Установлено, что рассматриваемая система имеет 12 экстремумов по направлениям осей и диагоналей. Система обладает монотонной амплитудно-частотной характеристикой (АЧХ) при отсутствии экстремумов у функций, получаемых из (2). Для исследования частотных свойств двумерной нерекурсивной цифровой системы второго порядка с симметричными коэффициентами, также как и в рекурсивном случае, применим аппарат дифференциальной геометрии в пространстве [2].

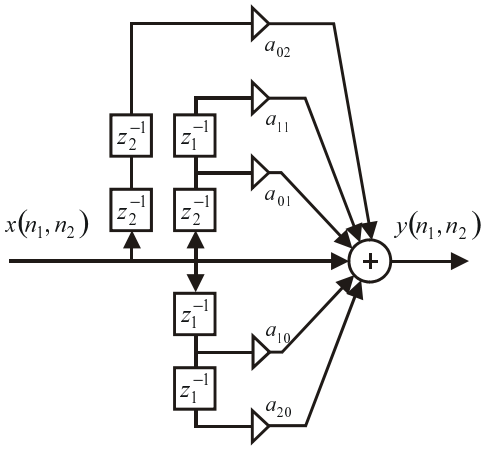


Рис. 1. Структурная схема двумерной нерекурсивной цифровой системы второго порядка с симметричными коэффициентами.

В точке $(0,0)$ условие положительности имеет вид:

$$\begin{cases}(a+c)(1+a+2b)+4b(1+a+b)<0,\\ [(a+c)(1+a+2b)+4b(1+a+b)]^2-[(a+2b)+c(2b-1)]^2>0,\end{cases} \quad (3)$$

а условие отрицательности:

$$\begin{cases}(a+c)(1+a+2b)+4b(1+a+b)>0,\\ [(a+c)(1+a+2b)+4b(1+a+b)]^2-[(a+2b)+c(2b-1)]^2>0,\end{cases} \quad (4)$$

В точке (π, π) условие положительности имеет вид:

$$\begin{cases}(a-c)(1-a+2b)+4b(a-1-b)>0,\\ [(a-c)(1-a+2b)+4b(a-1-b)]^2-[(a+2b)+c(2b-1)]^2>0,\end{cases} \quad (5)$$

а условие отрицательности:

$$\begin{cases}(a-c)(1-a+2b)+4b(a-1-b)<0,\\ [(a-c)(1-a+2b)+4b(a-1-b)]^2-[(a+2b)+c(2b-1)]^2>0.\end{cases} \quad (6)$$

Для того чтобы система была фильтром нижних частот, необходимо выполнение условий (4) и (6). Для того чтобы система была фильтром верхних частот, необходимо, чтобы были выполнены условия (3) и (5).

Для фильтра нижних частот максимум АЧХ приходится на точку $(0,0)$, а максимальное подавление входного сигнала на точку (π,π), поэтому подавление ФНЧ:

$$G = 20\lg\left(\frac{1+2a+c+2b}{1-2a+c+2b}\right). \quad (7)$$

Для фильтра верхних частот максимум АЧХ приходится на точку (π,π), а максимальное подавление входного сигнала на точку $(0,0)$, поэтому подавление ФВЧ:

$$G = 20\lg\left(\frac{1-2a+c+2b}{1+2a+c+2b}\right). \quad (8)$$

В качестве примера рассмотрим типичные ФНЧ и ФВЧ и найдем для них по формулам (7-8) подавление входного сигнала. Для ФНЧ с параметрами $a=0.62,\ b=0.09,\ c=0.1$ затухание $G=-35.99$ дБ (рис. 2). Для ФВЧ с параметрами $a=-0.63,\ b=0.1,\ c=0.1$ затухание $G=-36.12$ дБ (рис. 3).

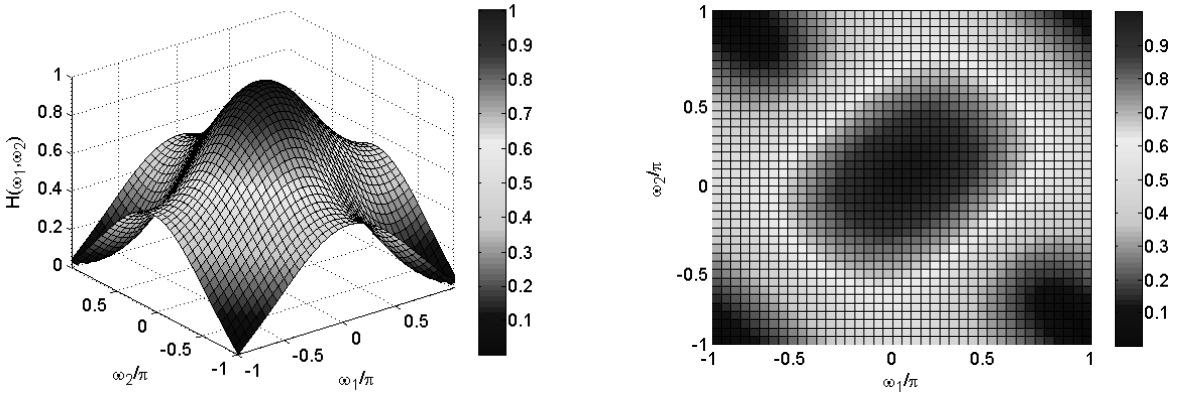
При проектировании цифровых систем одним из важнейших параметров является линия среза. Выполним нормировку значения квадрата модуля частотной характеристики фильтра нижних частот. Для этого вычислим значение $|H(0,0)|^2=(1+2a+c+2b)^2$. Далее получим:

$$\frac{|H|^2}{|H_{0,0}|^2}=\frac{1}{2};\ \ 2|H|^2=|H_{0,0}|^2. \quad (9)$$

В уравнении (9) используем разложение функций $\sin(\omega)$ и $\cos(\omega)$ в ряд Тейлора второго порядка в окрестности нуля. Подставляя результаты этого разложения в (3), получим:

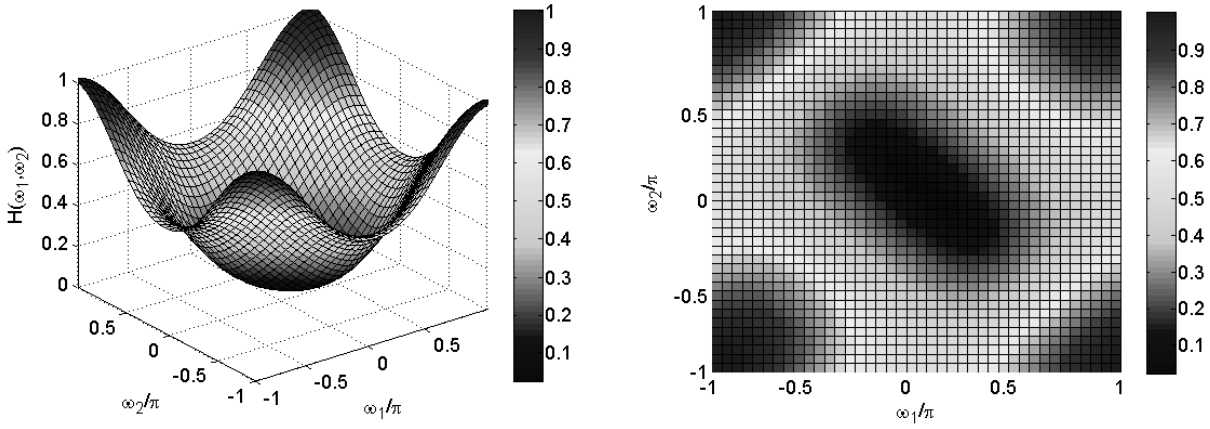
$$\begin{aligned}&[(a+c+2b)(2b-1)+2b(2a-1)+ac]\,\omega_1^2+\\ &+[(a+c+2b)(2b-1)+2b(2a-1)+ac]\,\omega_2^2+\\ &+2[c(1-2b)]-(a+2b)^2\,\omega_1\omega_2=(1+2a+c+2b)^2.\end{aligned} \quad (10)$$

Соотношение (10) является уравнением линии среза (окружности) двумерного нерекурсивного цифрового фильтра второго порядка с симметричными коэффициентами во втором приближении.



а) б)

Рис. 2. Типичный вид АЧХ двумерного цифрового фильтра нижних частот



а) б)

Рис. 3. Типичный вид АЧХ двумерного цифрового фильтра верхних частот

Результаты моделирования

Описанные выше двумерные нерекурсивные цифровые фильтры второго порядка с симметричными коэффициентами могут быть применены для решения задачи удаления периодических помех из изображений.

В качестве модели периодической помехи использована двумерная синусоида:

$$r(x,y)=A\sin\left[2\pi\omega_{10}(x+\Delta x)/M+2\pi\omega_{20}(y+\Delta y)/N\right], \quad (11)$$

где A – амплитуда, ω_{10} и ω_{20} определяют пространственные частоты соответственно по осям x и y, Δx и Δy – сдвиги фаз относительно начала отсчета, M и N – число строк и столбцов матрицы $r(x,y)$.

Дискретным преобразованием Фурье размером $M\times N$ от $r(x,y)$ является функция

$$R(\omega_1,\omega_2)=j\frac{A}{2}[\left(e^{2\pi\omega_1\Delta x/M}\right)\delta(\omega_1+\omega_{10},\omega_2+\omega_{20})-$$
$$-\left(e^{2\pi\omega_2\Delta y/N}\right)\delta(\omega_1-\omega_{10},\omega_2-\omega_{20})],$$

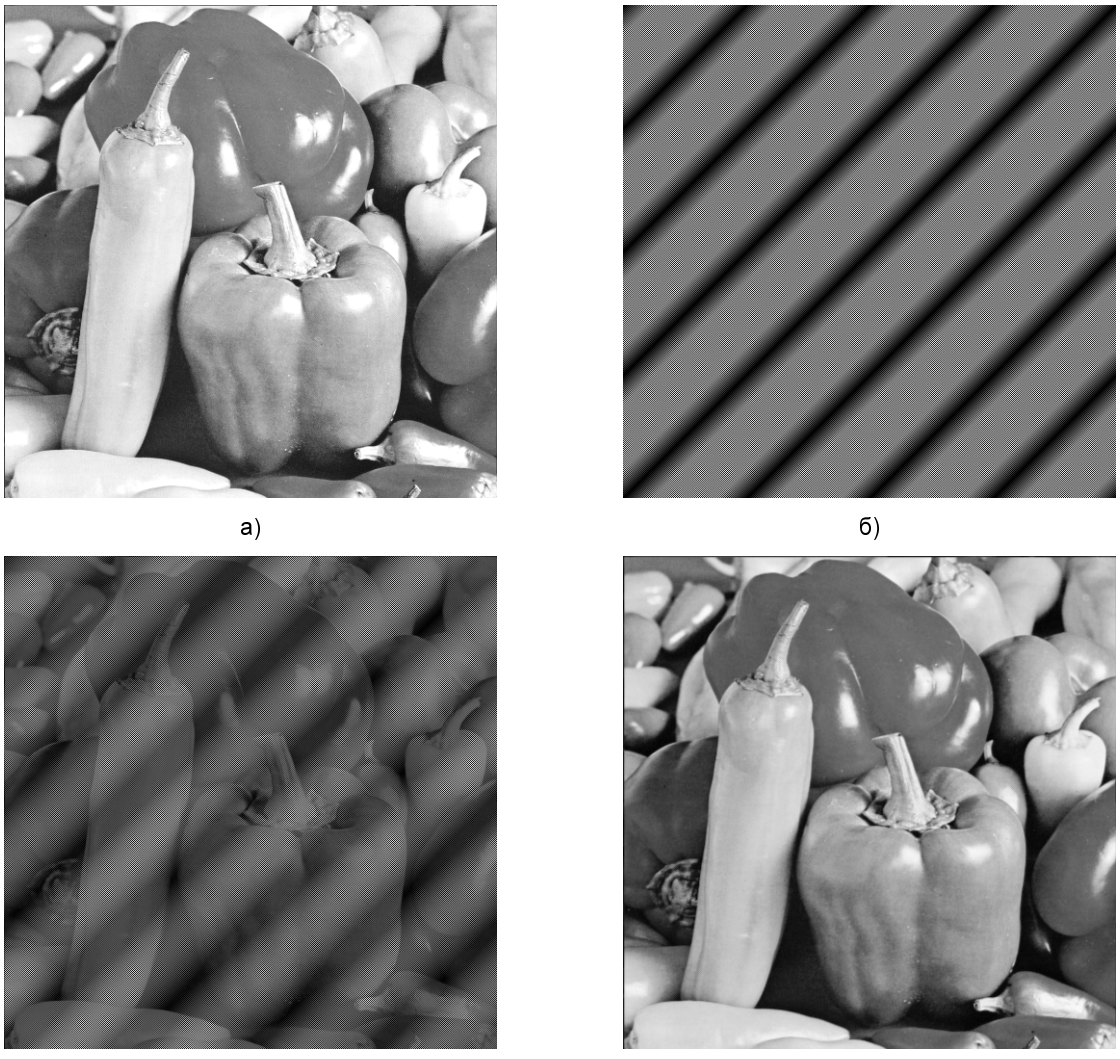
которая имеет вид пары комплексно сопряженных единичных импульсов, находящихся в точках $(\omega_{10},\omega_{20})$ и $(-\omega_{10},-\omega_{20})$ соответственно.

За основу метрики оценки качества цифрового изображения в работе принято пиковое отношение сигнал/шум (ПОСШ), определяемое по следующей формуле:

$$ПОСШ=20\log_{10}\frac{255}{\sqrt{СКО}},\quad СКО=\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(x_{i,j}-y_{i,j}\right)^2,$$

где M и N – число строк и столбцов матриц исходного $x_{i,j}$ и восстановленного $y_{i,j}$ изображений.

Примеры исходных изображений, периодических помех (построенных на основе модели, введенной выше), искаженных и восстановленных изображений приведены на рис. 4 и рис. 5. Помеха на рис. 4 представляет собой единичную высокочастотную синусоиду, описанную с использованием формулы (11). Помеха на рис. 5 представляет собой суперпозицию двух низкочастотных синусоид.



а) б) в) г)

Рис. 4. Обработка тестового изображения ФНЧ: а) исходное тестовое изображение «Перцы»; б) синусоидальная помеха; в) искаженное изображение, ПОСШ = 7,67 дБ; г) восстановленное изображение с использованием рассматриваемого ФНЧ, ПОСШ = 31,96 дБ

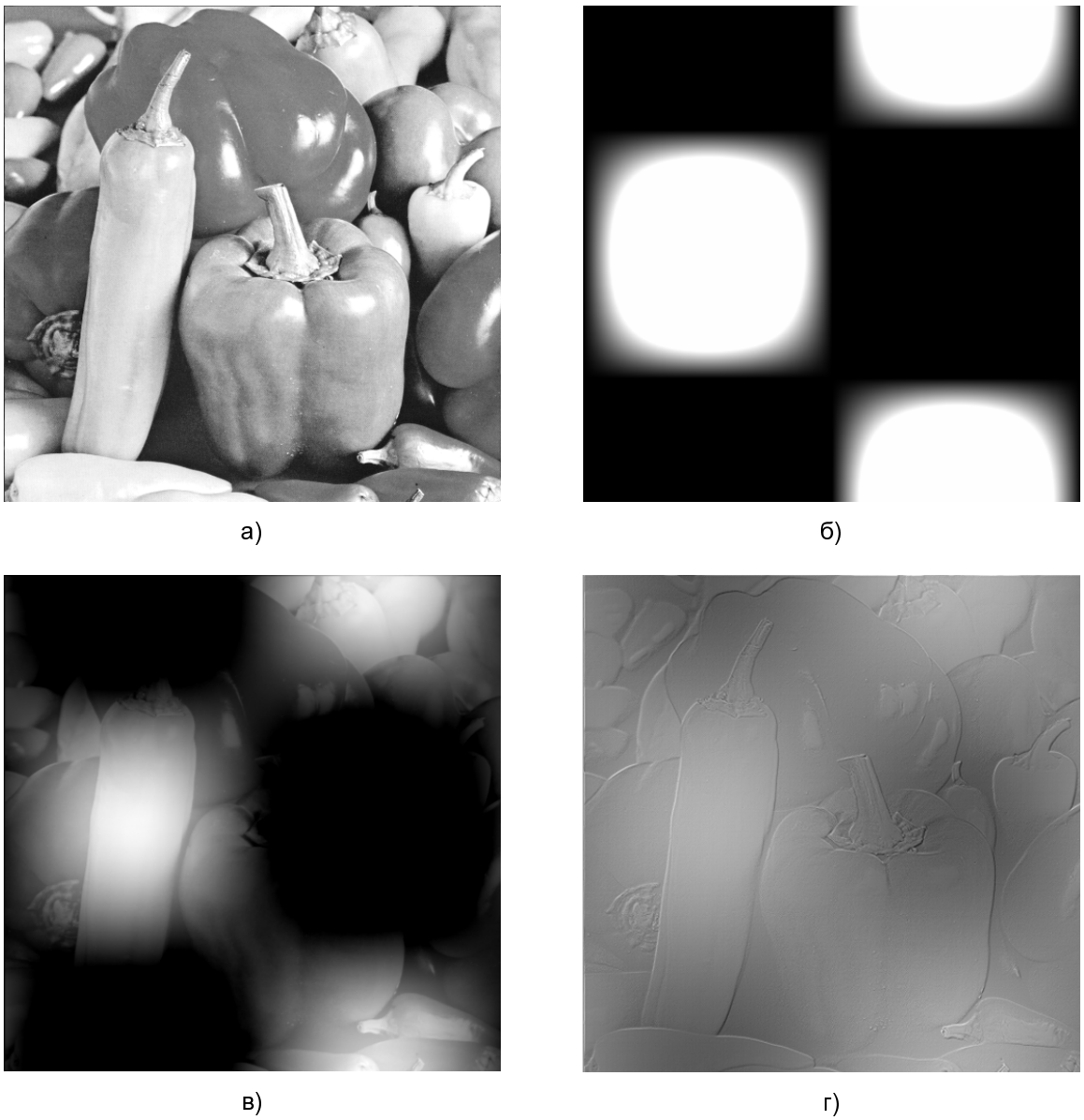


Рис. 5. Обработка тестового изображения ФВЧ: а) исходное тестовое изображение «Перцы»; б) синусоидальная помеха; в) искаженное изображение; г) восстановленное изображение с использованием рассматриваемого ФВЧ

Изображение, восстановленное с использованием ФНЧ с параметрами $a = 0.62,\ b = 0.09,\ c = 0.1$, приведено на рис. 4г. Видно, что высокочастотная синусоидальная помеха полностью удаляется указанным фильтром.

Результаты использования ФВЧ с параметрами $a = -0.63,\ b = 0.1,\ c = 0.1$ представлены на рис. 5. Как видно из рис. 5г, восстановленное изображение, полученное с использованием ФВЧ, содержит информацию о деталях (границах, контурах и т.д.) исходного тестового изображения. Помимо удаления низкочастотных компонент собственно изображения частично удалена и низкочастотная помеха, представленная на искаженном изображении рис. 5в. Неполное удаление помехи связано с видом АЧХ рассматриваемого ФВЧ (рис. 3).

Литература

1. Даджион Д., Мерсеро Р. Цифровая обработка многомерных сигналов. – М.: Мир, 1988.
2. Брюханов Ю.А., Приоров А.Л., Мясников Е.А., Калинин С.А. Частотные свойства двумерных рекурсивных цифровых систем первого порядка // Изв. высш. учеб. заведений. Радиоэлектроника. 1995. № 4. С. 26-30.

УДК 621.397

ФРАКТАЛЬНАЯ ОБРАБОТКА ИЗОБРАЖЕНИЙ

Перегуда Е.С.

Решается задача сокращения вычислительной сложности фрактального анализа. Разработаны три алгоритма сокращения вычислительной сложности и приведено их обоснование. Представлены результаты исследований, показывающие эффективность предложенных алгоритмов.

Введение.
Постановка задачи

Задача фрактальной обработки изображений представляет большой интерес как с научной, так и с практической точек зрения. Теория фрактального преобразования и построения аттрактора хорошо исследована. Данная теория обширна. Она затрагивает множество областей математики, физики, химии, биологии.

Важной перспективой приложения фрактального преобразования к задаче сжатия с потерями является возможность расширения теории фрактального множества на другие типы медиа данных, допускающих сжатие с потерями. К таким типам данных можно отнести аудио и видео информацию. Расширение фрактального множества на данные типы информации позволяет говорить о возможности создания универсального алгоритма сжатия данных с потерями. Создание такого универсального алгоритма представляет большой практический интерес.

Алгоритмическое решение обратной задачи фрактального преобразования можно представить в следующем виде: кодируемое множество разбивается на два класса – рангов R и доменов D, и каждому рангу ищется соответствующий домен, который после применения нужного преобразования в точности соответствует рангу. Сложность расчета заключается в переборе всех комбинаций рангов с доменами и применения всех выбранных преобразований. В качестве преобразований часто выбирают самые простые изометрические аффинные преобразования. Большинство исследователей выбирают в качестве базового алгоритм Фишера [1]. Выражение расчета метрики отклонения домена и ранга $\|E(D,R)\|$ имеет следующий вид:

$$\|E(D,R)\| = \sqrt{\sum_{i=0}^{n-1}\sum_{j=0}^{m-1}\left(\varphi\circ\left(s\cdot d_{ij}+o\right)-r_{ij}\right)^2}\ , \qquad (1)$$

где d_{ij} - значение точки домена D, r_{ij} - значение точки ранга R, φ - оператор преобразования, связанный с перестановкой точек в пространстве (чаще всего это аффинные изометрическое преобразования), s и o - масштабирование и смещение яркости точки. Параметры *s* и *o* рассчитываются, исходя из условия экстремума метрики отклонения. Выражение фрактальной аппроксимации $s\cdot d_{ij}+o$ является линейной функцией, поэтому выражение метрики отклонения (1) обладает только одной точкой экстремума-минимума. Исходя из этого, выражение (1) дифференцируют относительно переменных *s* и *o*. В итоге, исходя из условия экстремума-минимума, получают следующие выражения оптимума для переменных *s* и *o*:

$$s = \frac{n\cdot m\left(\sum_{i=0}^{n-1}\sum_{j=0}^{m-1}\left(\varphi\circ d_{ij}\right)\cdot r_{ij}\right)-\left(\sum_{i=0}^{n-1}d_{ij}\right)\left(\sum_{i=0}^{m-1}r_{ij}\right)}{n\cdot m\sum_{i=0}^{n-1}\sum_{j=0}^{m-1}d_{ij}^2-\left(\sum_{i=0}^{n-1}\sum_{j=0}^{m-1}d_{ij}\right)^2}; \qquad (2)$$

$$o = \frac{1}{n\cdot m}\left(\sum_{i=0}^{n-1}\sum_{j=0}^{m-1}r_{ij}-s\cdot\sum_{i=0}^{n-1}\sum_{j=0}^{m-1}d_{ij}\right). \qquad (3)$$

Данные выражения позволяют выполнить фрактальный анализ с наименьшим отклонением аттрактора от исходного изображения. Данный факт определяет высокое качество аппроксимации фракталом любого произвольного изображения, что обуславливает выбор алгоритма Фишера в качестве базового при разработке ускоренного фрактального анализа.

Сложность алгоритма заключается в необходимости выполнения множества операций умножения, деления, сложения, вычитания и извлечения корня. Наиболее затратными по времени выполнения являются операции умножения и деления, в то время как операцию извлечения квадратного корня можно исключить. Операции деления в выражениях (2) и (3) производятся только на переменные

$$n\cdot m\sum_{i=0}^{n-1}\sum_{j=0}^{m-1}d_{ij}^2-\left(\sum_{i=0}^{n-1}\sum_{j=0}^{m-1}d_{ij}\right)^2 \text{ и } n\cdot m,$$

которые носят постоянный характер и зависят только от текущего анализируемого домена. Такая постановка вопроса позволяет произвести предварительные расчеты данных значений и преобразовать значение данной переменной из знаменателя в числитель. Таким образом, появляется возможность исключить из расчета операции деления. Следующими по сложности являются операции умножения. Среди всех операций умножения в выражениях (1), (2), (3) следует выделить три наиболее часто повторяемые:

1. Операция умножения при расчете свертки текущего домена и ранга $\left(\sum_{i=0}^{n-1}\sum_{j=0}^{m-1}\left(\varphi\circ d_{ij}\right)r_{ij}\right)$

на каждый пиксель в выражении (2). Важно отметить, что для каждого нового преобразования $\varphi\circ$ необходимо заново пересчитывать свертку домена и ранга.

2. Две операции умножения на каждый пиксель из выражения (1).

Данные операции невозможно заменить на более простые операции сложения и вычитания. Исключение операций умножения позволило бы значительно сократить вычислительную сложность фрактальных алгоритмов.

Полный перебор всех комбинаций можно представить в виде следующего графа состояний (рис. 1). В этом графе узлы представляют из себя множество рангов $\{R_j\}$, посещение которых возможно, если найдены наиболее подходящие для них домены из множества $\{D_i\}$, и метрика отклонения $\|E(D_i,R_j)\|$ минимальна, при этом количество доменов и рангов равны соответственно N и M. Таким образом, ребро графа характеризует домен, а его вес характеризует метрику отклонения домена от ранга, на который указывает ребро. Очевидно, что качество решения обратной задачи фрактального преобразования будет тем выше, чем короче путь прохождения через все узлы графа. Важно отметить, что каждому домену можно применить несколько преобразований φ. На рис. 1 это выражается в восьми дополнительных ребрах $\varphi=\{1,2,3,4,5,6,7,8\}$ для каждого домена. Число восемь определяется восемью аффинными изометрическими преобразованиями, применяемыми к прямоугольному сегменту изображения.

Важно отметить сходство решения данной задачи с известной задачей о коммивояжере. Но в отличие от нее граф задачи на рис. 1 располагается в линейном одномерном пространстве, а это значит, что отсутствуют промежуточные узлы. Так же в задаче о коммивояжере решалась задача нахождения наикротчайшего пути по заданному графу узлов и ребер, в то время как в рассмотренном графе длина ребер априори неизвестна, и необходимо для каждого ранга рассчитать вес ребра отклонения домена для всех выбранных преобразований. Именно необходимость многократного пересчета комбинаций доменов и рангов и выбора ребра наименьшего веса и определяет вычислительную сложность решения обратной задачи фрактального преобразования. С увеличением изображения увеличивается количество рангов, доменов, и количество требуемых математических операций увеличивается в прогрессии степени натурального числа.

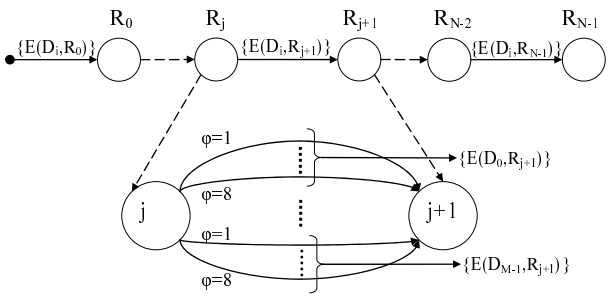


Рис. 1. Граф состояний алгоритма решения обратной задачи фрактального преобразования.

Автором было предложено три алгоритма сокращения вычислительной сложности алгоритма решения обратной задачи фрактального преобразования [5-13]. В совокупности сокращение времени работы алгоритма составляет 30 – 40 раз по сравнению с алгоритмом без включения данных алгоритмов.

В качестве отправной точки в поиске методов сокращения сложности рассматривают каждый ранговый и доменный блок как вектор $\vec{R}$ и $\vec{D}$ соответственно, в линейном векторном пространстве $\Re^n$ где $n=N\times N$ - количество пикселей в квадратном ранговом блоке со сторонами N. Преобразование из квадратного блока со стороной N в вектор длинной $n=N^2$ может быть произведено путем построчного сканирования блока.

Алгоритмы сокращения вычислительной сложности

Первый алгоритм. Первоначально при подстановке выражения (3) в выражение (1) получим следующее выражение:

$$\|E(D,R)\|=\sqrt{\sum_{i=0}^{n-1}\sum_{j=0}^{m-1}\left(\varphi\circ\left(s\cdot d_{ij}+\frac{1}{n\cdot m}\left(\sum_{i=0}^{n-1}\sum_{j=0}^{m-1}r_{ij}-s\cdot\sum_{i=0}^{n-1}\sum_{j=0}^{m-1}d_{ij}\right)\right)-r_{ij}\right)^2} \quad (4)$$

Обозначим среднею яркость домена и ранга следующим образом:

$$\overline{R}=\frac{1}{n\cdot m}\left(\sum_{i=0}^{n-1}\sum_{j=0}^{m-1}r_{ij}\right), \quad (5)$$

$$\overline{D}=\frac{1}{n\cdot m}\left(\sum_{i=0}^{n-1}\sum_{j=0}^{m-1}d_{ij}\right). \quad (6)$$

Средняя яркость независима от пространственного преобразования. Результатом подстановки соотношений (5) и (6) в (4) является:

$$\|E(D,R)\|=\sqrt{\sum_{i=0}^{n-1}\sum_{j=0}^{m-1}\left(\varphi\circ\left(s\cdot\left(d_{ij}-\overline{D}\right)\right)-\left(r_{ij}-\overline{R}\right)\right)^2}. \quad (7)$$

Представление блоков рангов и доменов в векторной форме позволяет ввести обозначения соответствующих векторных операций в выражение расчета метрики отклонения между доменом и рангом. Если для упрощения расчетов операцию пространственного преобразования φ исключить, то

$$\|E(D,R)\|=\left\|s\cdot\vec{\overline{D}}-\vec{\overline{R}}\right\|,$$

Где $\vec{\overline{D}}=\left\{d_{ij}-\overline{D}\right\}$ - вектор доменного блока с нулевой средней яркостью, $\vec{\overline{R}}=\left\{r_{ij}-\overline{R}\right\}$ - вектор рангового блока с нулевой средней яркостью.

Как видно при векторном представлении преобразований происходит переход к векторным операциям. В таком линейном пространстве операции сложения и вычитания векторов, операции умножения и сложения скаляров с векторами линейны и расчет метрики имеет следующий вид:

$$\|E(D,R)\|=\left\|s\cdot\vec{\overline{D}}-\vec{\overline{R}}\right\|=\left\|s\cdot\vec{D}^{n-1}-\vec{R}^{n-1}\right\|, \quad (8)$$

где $\vec{D}^{n-1}$ и $\vec{R}^{n-1}$ - векторы доменного и рангового блоков в ортогональном пространстве S^n с Евклидовой метрикой без постоянной составляющей.

Ортогональное дополнение пространств домена $\vec{D}^{n-1}$ и ранга $\vec{R}^{n-1}$ в пространстве S^n также является векторным пространством, состоящим из одного вектора постоянной составляющей.

Выражение (8) описывает простое условие отклонения с одним неизвестным. Частным случаем является условие полного подобия:

$$\left\| E\left(\vec{D}^{n-1}, \vec{R}^{n-1}\right)\right\| = 0, \quad (9)$$

классическое решение которого имеет вид:

$$s = \vec{R}^{n-1} \cdot \left(\vec{D}^{n-1}\right)^{+}. \quad (10)$$

Решение с псевдообратной матрицей с вычислительной точки зрения представляет большую сложность. Но учитывая тот факт, что эта матрица имеет смысл наименьшего квадратичного приближения, возможно, следует выбрать более простую операцию приближения, с меньшей точностью, но более быструю.

Решение можно получить при рассмотрении среднеквадратичной метрики:

$$\sqrt{r_1^2 + r_2^2 + \ldots + r_{n-1}^2} = \sqrt{\left(s \cdot d_1\right)^2 + \left(s \cdot d_2\right)^2 + \ldots + \left(s \cdot d_{n-1}\right)^2};$$

$$s = \frac{\sqrt{r_1^2 + r_2^2 + \ldots + r_{n-1}^2}}{\sqrt{\left(d_1\right)^2 + \left(d_2\right)^2 + \ldots + \left(d_{n-1}\right)^2}} = \frac{\left\|\vec{R}^{n-1}\right\|}{\left\|\vec{D}^{n-1}\right\|} \quad (11)$$

Подставив выражение (11) в (8), получим:

$$\left\| E\left(\vec{D}^{n-1}, \vec{R}^{n-1}\right)\right\| = \left\| \frac{\left\|\vec{R}^{n-1}\right\|}{\left\|\vec{D}^{n-1}\right\|} \cdot \vec{D}^{n-1} - \vec{R}^{n-1}\right\| = \left\| \frac{\vec{D}^{n-1}}{\left\|\vec{D}^{n-1}\right\|} - \frac{\vec{R}^{n-1}}{\left\|\vec{R}^{n-1}\right\|}\right\| = 0, \quad (12)$$

если принять во внимание что множитель метрики рангового блока $\left\|\vec{R}^{n-1}\right\|$ постоянен при переборе доменов $\vec{D}^{n-1}$ из множества $\{D_i\}$. Дроби вида

$$\frac{\vec{D}^{n-1}}{\left\|\vec{D}^{n-1}\right\|} = \phi\left(\vec{D}^{n-1}\right) \text{ и } \frac{\vec{R}^{n-1}}{\left\|\vec{R}^{n-1}\right\|} = \phi\left(\vec{R}^{n-1}\right)$$

являются выражениями нормализации векторов домена и ранга в нормированном метрическом пространстве.

Решение (12) соответствует условию полного подобия. Во всех других случаях решения носят характер интерполяции и стремятся к минимизации метрики отклонения домена и ранга

$$\left\| E\left(\vec{D}^{n-1}, \vec{R}^{n-1}\right)\right\|.$$

Операция нормализации применяется к векторам доменных и ранговых блоков только один раз и в дальнейшем может использоваться только их нормированные значения. Таким образом, исключается одна операция умножения. Вторая операция умножения исключается путем замены операции свертки на операцию деления одного скаляра на другой. Третью операцию умножения исключают путем перехода от метрики в Евклидовом пространстве к метрике по абсолютному значению:

$$\left\| E\left(\vec{D}^{n-1}, \vec{R}^{n-1}\right)\right\| = \left|\left(\frac{d_1}{\left\|\vec{D}^{n-1}\right\|} - \frac{r_1}{\left\|\vec{R}^{n-1}\right\|}\right)\right| + \ldots + \left|\left(\frac{d_{n-1}}{\left\|\vec{D}^{n-1}\right\|} - \frac{r_{n-1}}{\left\|\vec{R}^{n-1}\right\|}\right)\right| \quad (13)$$

Результатом реализации является сокращение времени анализа в 4 раза. При этом исключаются наиболее часто повторяющиеся операции умножения, и алгоритм требует только одну операцию сложения и вычитания.

Второй алгоритм. Далее приводится обоснование исключения цикла перебора восьми изометрических аффинных преобразований. Данное исключение основывается на том факте, что при выборе в качестве ортогонального дискретного косинусного преобразования ДКП (DCT) появляется возможность использовать свойство четности ядра данного преобразования. Данное свойство выражается в тождестве:

$$DCT(Y(x,y)) = DCT(Y(x,(N-1)-y)),$$

где $Y(x,y)$ - прямое изображение, а $Y(x,(N-1)-y)$ - его зеркальное отражение. Данное тождество имеет смысл только для четных коэффициентов ДКП. Таким образом, для изометрических аффинных преобразований одного изображения четные коэффициенты ДКП идентичны. Это значит, что в произведении над изображением операции пространственного преобразования φ, четные коэффициенты ДКП остаются неизменными. Это позволяет представить метрику отклонения домена и ранга $\left\| E\left(\vec{D}^{n-1}, \vec{R}^{n-1}\right)\right\|$ в следующем виде:

$$\left\| E\left(\vec{D}^{n-1}, \vec{R}^{n-1}\right)\right\| = \left\| E\left(\vec{D}^{n-1}, \vec{R}^{n-1}\right)_{ЧЕТ}\right\| + \left\| E\left(\vec{D}^{n-1}, \vec{R}^{n-1}\right)_{НЕЧЕТ}\right\| \quad (14)$$

Задавшись параметром качества метрики отклонения домена и ранга ε, получим:

$$\left\| E\left(\vec{D}^{n-1}, \vec{R}^{n-1}\right)\right\| = \left\| E\left(\vec{D}^{n-1}, \vec{R}^{n-1}\right)_{ЧЕТ}\right\| + \left\| E\left(\vec{D}^{n-1}, \vec{R}^{n-1}\right)_{НЕЧЕТ}\right\| \le \varepsilon \quad (15)$$

Если неравенство (15) тождественно, то домен аппроксимирует ранг с искажением не больше чем заданное ε. С учетом оператора преобразования φ получим:

$$\left\| E\left(\varphi \circ \vec{D}^{n-1}, \vec{R}^{n-1}\right)\right\| =$$

$$\left\| E\left(\vec{D}^{n-1}, \vec{R}^{n-1}\right)_{ЧЕТ}\right\| + \left\| E\left(\varphi \circ \vec{D}^{n-1}, \vec{R}^{n-1}\right)_{НЕЧЕТ}\right\| \le \varepsilon$$

и условие расчета метрики отклонения $\left\| E\left(\varphi \circ \vec{D}^{n-1}, \vec{R}^{n-1}\right)\right\|$ можно представить в виде:

$$\left\| E\left(\varphi \circ \vec{D}^{n-1}, \vec{R}^{n-1}\right)\right\| = \begin{cases} \left\| E\left(\vec{D}^{n-1}, \vec{R}^{n-1}\right)_{ЧЕТ}\right\| + \left\| E\left(\varphi \circ \vec{D}^{n-1}, \vec{R}^{n-1}\right)_{НЕЧЕТ}\right\|; \text{если } \left\| E\left(\vec{D}^{n-1}, \vec{R}^{n-1}\right)_{ЧЕТ}\right\| \le \varepsilon \\ \text{Следующий элемент множества } \{D_i\}; \text{если } \left\| E\left(\vec{D}^{n-1}, \vec{R}^{n-1}\right)_{ЧЕТ}\right\| > \varepsilon \end{cases} \quad (16)$$

Это условие является условием ветвления алгоритма динамического программирования.

Результатом реализации является сокращение времени анализа от 1 до 32 раза, в зависимости от состава изображения. Данный элемент динамического программирования позволяет сравнивать домен с рангом на основе только четверти всех коэффициентов.

Третий алгоритм. Последним алгоритмом сокращения сложности является алгоритм выбора оптимального изометрического аффинного преобразования с использованием неравномерного распределение энергии по коэффициентам ДКП. Данное свойство применяется при зигзагообразном представлении коэффициентов ДКП, хорошо известном из спецификации JPEG. Исходя из этого, достаточно, чтобы знаковое распределение младших индексов рангового и доменного блоков

совпадало. Таким образом зачастую появляется возможность быстрого установления наиболее подходящего аффинного преобразования на основе сравнения знаков только младших индексов (например, трех) коэффициентов. Исходя из этого, действие разработанного алгоритма заключается в использовании прямого метода адресации – поиска номера аффинного преобразования в ячейке памяти, адресуемой трехразрядным двоичным кодом.

Результатом реализации является сокращение времени анализа в 8 раз.

Работу разработанных алгоритмов лучше всего представить в виде графа состояний. На рис. 2 дан граф состояний улучшенного алгоритма решения обратной задачи фрактального преобразования.

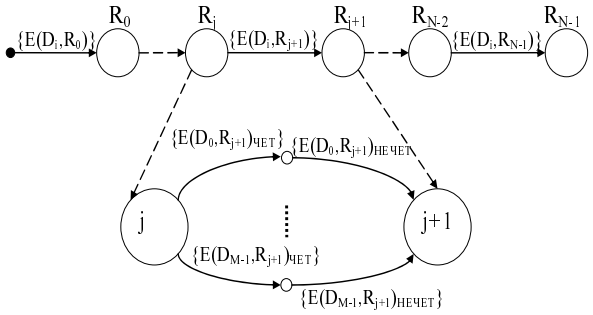


Рис. 2. Граф состояний улучшенного алгоритма решения обратной задачи фрактального преобразования.

При сравнении рисунков 1 и 2 видна значительная разница, между базовым алгоритмом и новым разработанным алгоритмом. Наличие в новом алгоритме промежуточной вершины между двумя соседними вершинами – рангами позволяет ввести элемент динамического программирования.

Динамическое программирование реализуется в виде разработанного второго алгоритма сокращения вычислительной сложности. Третий алгоритм позволяет выбрать оптимальные нечетные коэффициенты домена. Таким образом, между соседними вершинами-рангами каждому домену соответствует только одно ребро. Время перебора всех доменов сокращается примерно в восемь раз. Первый алгоритм исключает три операции умножения и две операции сложения, что соответствует сокращению времени работы в три – четыре раза. Суммарное сокращение времени работы составляет 30 – 40 раз.

Статистические исследования

Ниже приводятся статистические исследования улучшенного алгоритма решения обратной задачи фрактального преобразования. В качестве тестового было выбрано изображение «Lena».

Для оценки качества преобразования использована величина пикового отношения сигнал/шум (PSNR).

При реализации разработанных алгоритмов сокращения вычислительной сложности на базе алгоритма Фишера возникает проблема с применением ДКП. На практике данное преобразование не может быть применимо к изображениям произвольного размера, поэтому первоначально следует отказаться от разбиения изображения по схеме квадродерева и установить фиксированный размер рангов. Для выбора оптимального размера рангов необходимо сравнить результаты анализа при нескольких фиксированных размерах.

Определим следующую серию размеров рангов: 16×16, 8×8, 4×4. Зависимость качества от рангов представлена на рис. 3. Количество доменов при этом является фиксированным и равно 62000.

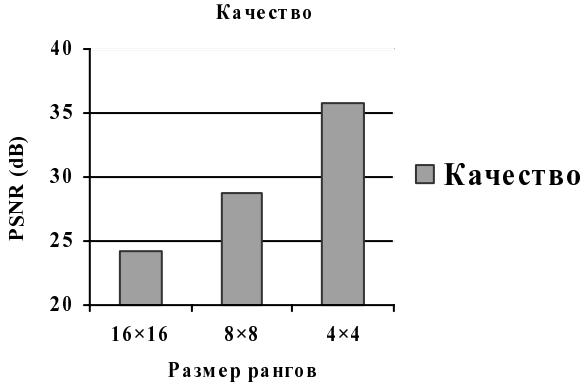


Рис. 3. График PSNR для размеров рангов 16×16, 8×8, 4×4

Как видно из рис. 3 оптимальный размер ранга составляет 4×4.

Далее приводятся результаты статистических исследований. Данные исследования представлены в виде графика на рис. 4. На данном графике представлено две кривые: время выполнения программы на базе алгоритма Фишера (1) и время выполнения программы на базе разработанных алгоритмов (2).

Как видно из графика на рис. 4 эффективность разработанных алгоритмов значительно превосходит алгоритм Фишера (30 – 40 раз). Причем эта эффективность растет с увеличением числа доменов.



Рис. 4. График времени выполнения программы на базе алгоритма Фишера (1) и времени выполнения программы на базе разработанных алгоритмов (2)

Кроме исследований сокращения времени работы приведены исследования качества. Исследования представлены в виде графика качества (рис. 5). На данном графике представлено две кривые: качество работы программы на базе алгоритма Фишера (1) и качество работы программы на базе разработанных алгоритмов (2). Как видно из графика на рис. 6 каче-

ство работы разработанных алгоритмов хуже, чем качество работы алгоритма Фишера. Это связано с тем, что в разработанных алгоритмах обработка данных производится в нормированном пространстве. При этом динамический диапазон ограничивается, что и приводит к небольшой неточности. Как видно из графика, данная неточность невелика (не более 1 – 1,5 дБ), и визуально разница незаметна.

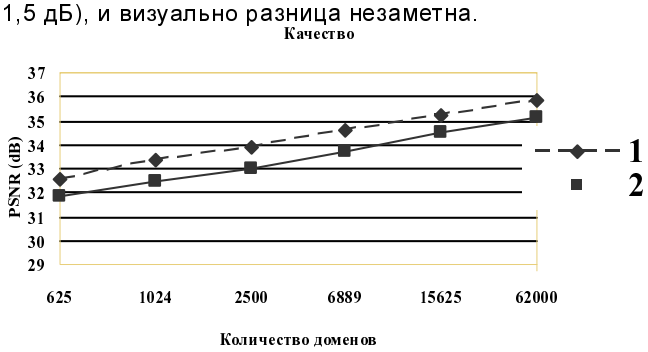


Рис. 5. График качества работы программы на базе алгоритма Фишера (1) и качества работы программы на базе разработанных алгоритмов (2).

Кроме исследований времени работы алгоритмов и качества анализа приведены исследования степени сжатия нового алгоритма фрактального анализа. Результаты сравнения качества сжатия от степени сжатия фрактального алгоритма и алгоритмов JPEG и JPEG2000 представлены в виде графиков на рис. 6.

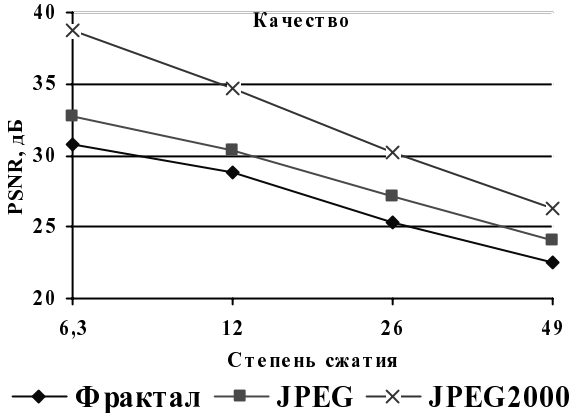


Рис. 6. Сжатие изображения «Lena»

Как видно из графиков, качество фрактального сжатия изображений уступает аналогам стандарта JPEG.

Заключение

В заключении важно отметить, что разработанные алгоритмы сокращения вычислительной сложности показали значительную эффективность, позволив уменьшить количество выполняемых математических операций и выбрать оптимальные операции преобразования, применяемые к изображениям. При этом, как очевидно, возможно дополнительно применить уже существующие алгоритмы классификации, такие как нейронные сети, для достижения еще большего сокращения времени решения обратной задачи фрактального преобразования.

Литература

1. Уэлстид С. Фракталы в вейвлеты для сжатия изображений в действии. Учебное пособие./ С. Уэлстид – М.: Издательство Триумф, 2003 – 320 с.: ил.
2. Ватолин Д. С. Алгоритмы сжатия изображений / Д. С. Ватолин – М.: Диалог-МГУ, – 1999. – 273 с.
3. Шабаршин А. А. Метод фрактального сжатия изображений / А. А. Шабаршин // Научные школы УПИ-УГТУ. – 1997. – №1. – С. 70-82.
4. Zhang N., Yan B. Hybrid image compression method based on fractal geometry / N. Zhang, B. Yan // Electronics Letters 30 (1994), P. 406-408.
5. Перегуда Е.С. О применимости фрактальной обработки в технике передачи изображения// Телевидение: передача и обработка изображений. Материалы 4-й международной конференции, Санкт-Петербург, 24-26 мая 2005. с. 42 – 43.
6. Pereguda E.S. Acceleration of Algorithm of Fractal Image Compression// IEEE International Siberian Conference on Control and Communication (SIBCON-2005). Proceedings, Tomsk, Russia, October 21-22, 2005. p. 159 – 162.
7. Pereguda E.S. Acceleration of Algorithm of Fract al Compression of Images// Proceeding of The KOREA – RUSSIA Joint – Workshop 2006 on Signal Transmission, Processing, Sensor and Monitoring Systems, Khabarovsk, Russia, October 26 – 28, 2006, p. 46 – 48.
8. Сай С.В. и Перегуда Е.С. Методы сокращения объема вычислений в алгоритмах фрактального сжатия изображений// Вестник Тихоокеанского Государственного Университета. №1, июнь, 2006. с. 9 – 14.
9. Перегуда Е.С. Методы ускорения фрактальной обработки изображений// Теоретические и прикладные вопросы современных информационных технологий. Материалы Седьмой Всероссийской научно-технической конференции, Улан-Удэ, 24-30 июля 2006. Часть 1 с. 98 – 103.
10. Перегуда Е.С. Методы ускорения фрактального сжатия изображения// НАУКА. ТЕХНОЛОГИИ. ИННОВАЦИИ, Материалы всероссийской научной конференции молодых ученых, Новосибирск, 07-10 декабря 2006 г., Часть 1. с. 186 – 187.
11. Перегуда Е.С. Ускорение фрактального алгоритма в системах сжатия и передачи изображений// Телекоммуникации №6 2007, с. 2 – 7.
11. Перегуда Е.С. Архиватор цветного изображения на основе фрактального преобразования// Свидетельство об официальной регистрации программы для ЭВМ № 50200600375, ВНТИЦ, 2006.
12. Сай С.В., Перегуда Е.С. Алгоритм поиска мелких структур фото и видео изображений// Телевидение: передача и обработка изображений. Материалы 6-й международной конференции, Санкт-Петербург, 19-21 июня 2007. с. 81 – 82.

УДК 621.501

ОБОБЩЕННАЯ ИДЕНТИФИКАЦИЯ ЛИНЕЙНОГО ОБЪЕКТА ПО СИГНАЛАМ ВХОД-ВЫХОД

Минко А.А., Витязев В.В.

Предложен алгоритм обобщенной идентификации на основе интегральных двухпараметрических преобразований Гаусса линейного стационарного объекта по его реакции на произвольный из заданного класса сигнал. Проведена экспериментальная проверка работоспособности разработанного алгоритма, подтвердившая его весьма высокую эффективность.

Введение

Рассмотрим задачу идентификации многомерного объекта (рис. 1), имеющего вектор входа $V(t)=(v_1(t)\dots v_q(t))^T$ и вектор выхода $Y(t)=(y_1(t)\dots y_k(t))^T$.

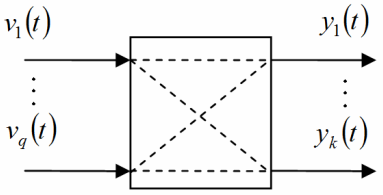


Рис. 1 – Структурная схема многомерного объекта

Будем считать, что дрейф параметров на требуемом интервале наблюдения T объекта пренебрежимо мал, а входной вектор не выводит объект за границы линейных диапазонов, так что допустимо использование линейных стационарных моделей по всем каналам вход-выход. Целью работы является разработка алгоритма идентификации отдельных каналов объекта, испытываемого по технологии «один вход - все выходы». Это позволяет в некоторых случаях сократить время испытаний объекта в k раз и, соответственно, увеличить время управления объектом, согласно идее дуального управления Фельдбаума.

Постановка задачи

Рассмотрим один из каналов стационарного многомерного объекта, динамические свойства которого при известном входном сигнале $v(t), t\in[0,T]$ допускают описание ЛДУ вида

$$\sum_{i=0}^{n} a_i y^{(n-i)}(t) = \sum_{j=0}^{m} b_j v^{(m-j)}(t),\ n=\overline{1,10},\ m\le n-1. \qquad (1)$$

Задача обобщенной идентификации заключается в нахождении не только вектора $P=(a_0\dots a_{n-1}\ b_0\dots b_m)^T$, составленного из коэффициентов левой и правой частей ЛДУ канала объекта (1), но и вектора $Y_{T-}=\left(y(T-)\dots y^{(n-1)}(T-)\right)^T$, составленного из фазовых координат канала объекта в предоконечный момент времени $t=T-0=T-$.

Одним из вариантов решения поставленной задачи является использование проекций ЛДУ объекта на координатные функции Гаусса [1]:

$$Q_l(t-r,s)=\frac{(-1)^l}{\sqrt{2\pi s}}\frac{d^l}{dt^l}e^{-\frac{(t-r)^2}{2s}},\ l=0,1,2,\dots,$$

которые допускают представление

$$Q_l(t-r,s)=\frac{1}{s^l\sqrt{2\pi s}}H_l(t-r,s)e^{-\frac{(t-r)^2}{2s}}, \qquad (2)$$

где $H_l(t-r,s),\ l=0,1,2,\dots$ - полиномы Эрмита, R_+^2 - верхняя полуплоскость над прямой $t\in R$. Параметр r функций Гаусса является параметром сдвига, а s – параметром масштаба.

Графики функций $Q_l(t-r,s),\ l=\overline{0,4}$ при r=0 и s=1 представлены на рис. 2.

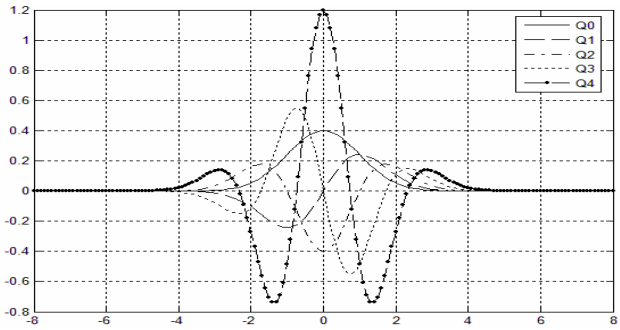


Рис. 2 – Графики функций $Q_l(t-r,s),\ l=\overline{0,4}$

Графики функций $Q_l(t-r,s),\ l=\overline{5,10}$ при r=0 и s=1 представлены на рис. 3.

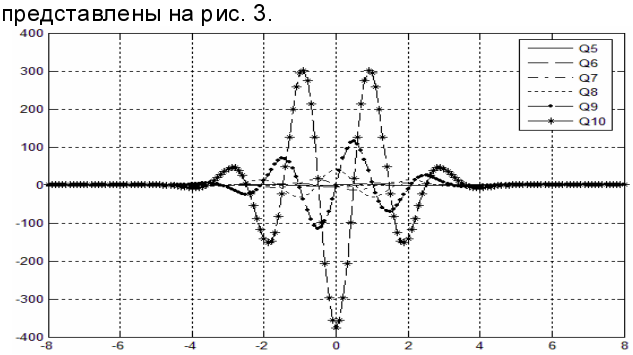


Рис. 3 – Графики функций $Q_l(t-r,s),\ l=\overline{5,10}$

Проекции сигнала $x(t),\ t\in[0,T]$ на координатные функции Гаусса определяются выражением

$$G_x^{(l)}(r,s)=\int_0^T x(t)Q_l(t-r,s)dt,\ l=0,1,2,\dots \qquad (3)$$

и называются интегральными преобразованиями Гаусса (ПГ) порядка l. Они обладают следующими свойствами.

1. ПГ линейны.

2. $G^{(l)}_{x^{(k)}}(r,s)=G^{(l+k)}_{x}(r,s)$.

3. $\lim\limits_{s\to 0+} G^{(l)}_{x}(r,s)=x^{(l)}(t),\ t=r$ - формула обращения.

ПГ позволяют алгебраизовать ЛДУ (1) и допускают численное определение на основе подходящих квадратурных формул, поскольку сигналы $v(t)$ и $y(t),\ t\in[0,T]$, вследствие типового подхода к сбору информации, заданы в виде сеток

$$[t_i,v_i]_0^N,\ [t_j,y_j]_0^N,$$

где $v_i=v(t_i),\ y_j=y(t_j),\ t_i\neq t_j$ при $i=j$. Порядок «l» ПГ целесообразно ограничить числом одиннадцать. Это ограничение обусловлено тем, что вычисление ПГ более высоких порядков требует вычисления полиномов Эрмита порядков $l>11$, а это, в свою очередь, будет вносить всё более возрастающую ошибку в вычисление самих ПГ.

Как правило, определены не отсчеты, а их оценки

$$\hat{v}_i=v_i+n_{vi};\ \hat{y}_j=y_j+n_{yj}, \tag{4}$$

где n_v, n_y - помехи, обусловленные рядом факторов. Поскольку сами ПГ обладают сглаживающими свойствами [2] при приемлемом уровне помех, аспекты получения наиболее достоверных оценок

$$\hat{G}^{(l)}_{v}(r,s),\ \hat{G}^{(l)}_{y}(r,s), l=0,1,2,\ldots$$

в работе не рассматриваются.

Учет финитных свойств входного и выходного сигналов объекта позволяет рассматривать их в виде

$$x(t)=x_p(t)1_T(t), \tag{5}$$

где $x_p(t)-$регулярная функция, заданная на открытом интервале $(0,T)$, а

$$1_T(t)=\begin{cases}1(t), t\in(0,T),\\ 0, t\notin[0,T]\end{cases}- \tag{6}$$

финитный единичный скачок. Характер изменения сигналов в точках 0 и T в данной работе не рассматривается. Производные финитных функций порядка «k», входящие в выражение (1), вычисляются по формуле [3]:

$$x^{(k)}(t)=x_p^{(k)}(t)1_T(t)+\sum_{i=1}^{k}x_p^{(k-i)}(0+)\delta^{(i-1)}(t)- \\ -\sum_{i=1}^{k}x_p^{(k-i)}(T-)\delta^{(i-1)}(t-T), \tag{7}$$

где $x_p^{(k-i)}(0+)\delta^{(i-1)}(t)$ - составляющая, содержащая постначальные фазовые координаты объекта, а $x_p^{(k-i)}(T-)\delta^{(i-1)}(t-T)$ - составляющая, содержащая предоконечные фазовые координаты.

Далее приведены варианты алгоритма идентификации на основе ПГ рассматриваемого стационарного объекта по его реакции на произвольный допустимый входной сигнал вначале при нулевых предначальных условиях, а затем при ненулевых.

Теоретическая часть

Не останавливаясь на вариантах алгоритма обобщенной идентификации объектов простой структуры $(n\le 4)$, перейдем к обсуждению алгоритма идентификации объекта, описываемого ЛДУ

$$\sum_{i=0}^{10}a_i y^{(10-i)}(t)=\sum_{j=0}^{9}b_j v^{(9-j)}(t),\ a_{10}=1 \tag{8}$$

10-го порядка, при нулевых $Y_{0-}=0$ предначальных условиях с полным оператором $B(p)$ в правой части ДУ. Для этих объектов условия баланса сингулярных и регулярных составляющих имеют следующий вид:

$$\begin{aligned}&a)\left(\sum_{i=0}^{k-1}y_p^{(k-1-i)}(0+)-\sum_{j=0}^{k-2}b_j v_p^{(k-2-j)}(0+)\right)\delta^{(10-k)}(t)=0;\\ &б)\left(\sum_{i=0}^{k-1}y_p^{(k-1-i)}(T-)-\sum_{j=0}^{k-2}b_j v_p^{(k-2-j)}(T-)\right)\delta^{(10-k)}(t)=0,\\ &k=\overline{1,10};\\ &в)\sum_{i=0}^{10}a_i y_p^{(10-i)}(t)=\sum_{j=0}^{9}b_j v_p^{(9-j)}(t),\ a_{10}=1.\end{aligned} \tag{9}$$

Будем считать, что заданный интервал $(0,T)$ позволяет выбрать такие значения $(r_{k_q},s_{k_q})\in R_+^2$ параметров ПГ, что выполняется условие:

$$\left|Q_l(T-r_{k_q},s_{k_q})\right|\le\varepsilon,$$

где ε - некоторое малое число, то есть не учитывать вектор Y_{T-} предоконечных значений фазовых координат выходного сигнала объекта. Данный прием уменьшает размерность формируемой матрицы СЛАУ с $4n$ до $3n$. При выборе таких значений $(r_{k_q},s_{k_q})\in R_+^2$ параметров ПГ, что появляется возможность неучета вектора Y_{0+} постначальных значений фазовых координат выходного сигнала объекта, приводит к тому, что алгоритм идентификации становится нечувствительным к быстро затухающим составляющим переходного процесса, а это, в свою очередь, уменьшает порядок определяемого ДУ модели объекта. Проекции ЛДУ (9,в) на координатные функции Гаусса $Q_l(t-r_{k_q},s_{k_q}),\ l=\overline{0,11}$, для всех точек $(r_{k_q},s_{k_q})\in R_+^2$ определяются алгебраическим соотношением

$$\sum_{i=0}^{10}a_i G_y^{(h+10-i)}(r_{k_q},s_{k_q})=\sum_{j=0}^{9}b_j G_v^{(h+9-j)}(r_{k_q},s_{k_q}), \\ k_q=\overline{1,q},\ q\in[10,20], \tag{10}$$

где $h=0$ при $k_q=\overline{1,20}$ и $h=\overline{0,1}$ при $k_q=\overline{1,10}$.

Начальное значение r_{k_1} параметра r ПГ целесообразно выбирать в начале сеток

$$[t_i,v_i]_0^N,\ [t_j,y_j]_0^N$$

реализации вход-выход объекта. Шаг h_r изменения параметра r должен выбираться так, чтобы все значения r_{k_q}, $k_q=1,q$ удовлетворяли условию

$$r_{k_q}\le\frac{t_0+t_N}{2}.$$

При малом шаге h_r изменения параметра r СЛАУ, формируемые из ПГ сеток сигналов вход-выход объекта, имеют матрицы, минимальные значения сингулярных чисел которых близки к нулю, что свидетельствует о вырожденности СЛАУ. Начальное значение s_{k_1} параметра s ПГ следует выбирать в диапазоне от 0,1 до 1÷3. Малые значения параметра s используются тогда,

когда реализация вход-выход имеет пренебрежимо малые шумовые компоненты (отношение сигнал-шум ≥50÷100). Сверху значение параметра s ограничивают условия, позволяющие не учитывать правые краевые эффекты при вычислении ПГ реализации вход-выход объекта.

При минимальном значении числа точек

$$(r_{k_q}, s_{k_q}) \in R_+^2,\ k_q = \overline{1,10}$$

потребуется 230 ПГ для формирования системы линейных алгебраических уравнений (СЛАУ):

$$\begin{pmatrix} G_y^1 & G_v^1 \\ \vdots & \vdots \\ G_y^{10} & G_v^{10} \end{pmatrix} P = (-1) \begin{pmatrix} G_{y0}^1 \\ \vdots \\ G_{y0}^{10} \end{pmatrix}, \quad (11)$$

где блоки определены как

$$G_y^{k_q} = \begin{pmatrix} G_y^{(10)}(r_{k_q}, s_{k_q}) \dots G_y^{(1)}(r_{k_q}, s_{k_q}) \\ G_y^{(11)}(r_{k_q}, s_{k_q}) \dots G_y^{(2)}(r_{k_q}, s_{k_q}) \end{pmatrix};$$

$$G_v^{k_q} = (-1) \begin{pmatrix} G_v^{(9)}(r_{k_q}, s_{k_q}) \dots G_v^{(0)}(r_{k_q}, s_{k_q}) \\ G_v^{(10)}(r_{k_q}, s_{k_q}) \dots G_v^{(1)}(r_{k_q}, s_{k_q}) \end{pmatrix}; \quad (12)$$

$$G_{y0}^{k_q} = \left(G_y^{(0)}(r_{k_q}, s_{k_q})\ \ G_y^{(1)}(r_{k_q}, s_{k_q})\right)^T;$$

$$P = \left(a_0 \dots a_9\ \ b_0 \dots b_9\right)^T,\ k_q = \overline{1,10}.$$

При другом подходе, полярном первому, для формирования СЛАУ

$$\left(G_{yk_q}^{(10)} \dots G_{yk_q}^{(1)} - G_{vk_q}^{(9)} \dots - G_{vk_q}^{(0)}\right) \times \left(p_{k_q}\right) = -\left(G_{yk_q}^{(0)}\right), \quad (13)$$

где

$$k_q = \overline{1,20},$$

$$G_{yk_q}^{(l)} = G_y^{(l)}(r_{k_q}, s_{k_q}),\ l = \overline{0,10};$$

$$G_{vk_q}^{(p)} = G_v^{(p)}(r_{k_q}, s_{k_q}),\ p = \overline{0,9},$$

требуется 420 ПГ сигналов вход-выход объекта.

Затем формируется банк моделей-кандидатов, исходя из гипотез о возможном описании модели объекта ДУ с 1-го по 10-й порядок:

1) одна модель 1-го порядка $(n=1,\ m=0)$;

2) две модели 2-го порядка $(n=2,\ m=0,1)$;

3) три модели третьего порядка $(n=3,\ m=\overline{0,2})$ и т.д. Всего банк содержит 55 моделей-кандидатов. Для каждой из них формируется и решается СЛАУ соответствующего порядка, позволяющая вычислить вектор $P = \left(a_0 \dots a_{n-1}\ b_0 \dots b_m\right)^T$ параметров и вектор $Y_{0+} = \left(y(0+) \dots y^{(n-1)}(0+)\right)^T$ постначальных значений фазовых координат выходного сигнала. Для моделей-кандидатов определяются выходные сигналы $y_{Mi}(t),\ i = \overline{1,55}$, $t \in (0,T)$. В результате, объект идентифицируется той моделью, для которой критерий соответствия модель-объект:

$$J = \left\| y(t) - y_M(t) \right\|, \quad (14)$$

где $\|\cdot\|$ - подходящая норма, принимает значение $J \to \min\limits_{j \in [1,55]}$.

Если значение критерия определенной модели-кандидата удовлетворяет условию $J \le \varepsilon_\partial$, где ε_∂ - допустимая погрешность аппроксимации выходного сигнала объекта выходным сигналом модели, то модель объекта пригодна для оценки вектора Y_{T-} предоконечных значений фазовых координат выходного сигнала объекта и прогнозирования реакций объекта на входные сигналы $v(t),\ t > T$. В противном случае необходимо взять другой набор точек $(r_{k_q}, s_{k_q}) \in R_+^2$ и повторить процедуру идентификации.

После выбора модели объекта определяется вектор $Y_{T-} = \left(y(T-) \dots y^{(n-1)}(T-)\right)^T$ для прогнозирования реакций объекта на входные сигналы $v(t)$ при $t > T$. Для реализации этого алгоритма требуется информация о векторе $V_{0+} = \left(v(0+) \dots v^{(m)}(0+)\right)^T$ постначальных значений координат фазового состояния входного сигнала, причем, в общем случае, он должен иметь $v^{(q_m)}(t),\ t = 0+,\ q_m = \overline{0,m}$. Таким образом, «произвольность» входного сигнала следует рассматривать как «произвольность в заданном классе».

Наглядную оценку необходимого числа ПГ сигналов вход-выход для различных порядков n ЛДУ моделей объекта приведены в таблице 1.

Таблица 1

n/q	**1**	**2**	**3**	**4**	**5**	**6**
0	2	-	-	-	-	-
1	5	6	-	-	-	-
2	11	14	17	20	-	-
3	17	22	27	32	37	42
4	23	30	37	44	51	58
5	-	38	47	56	65	74
6	-	46	57	68	79	90
7	-	-	67	80	93	106
8	-	-	-	92	107	122
9	-	-	-	-	-	138
n/q	**7**	**8**	**9**	**10**	**11**	**12**
4	65	72	-	-	-	-
5	83	92	101	110	-	-
6	101	112	123	134	145	156
7	119	132	145	158	171	184
8	137	152	167	182	197	212
9	155	172	189	206	223	240
10	-	-	-	230	249	268
n/q	**13**	**14**	**15**	**16**	**17**	**18**
7	197	210	-	-	-	-
8	227	242	257	272	-	-
9	257	274	291	308	325	342
10	287	306	325	344	363	382

В таблице 1 обозначено:

q - число значений $(r_{k_q}, s_{k_q}) \in R_+^2$ параметров ПГ, используемых для формирования СЛАУ полного ранга;

n - порядок ЛДУ, представляющего модель динамических свойств объекта;

M_{nq} - число ПГ сигналов вход-выход, покрывающих соответствующую часть расширенной матрицы СЛАУ с учетом того, что максимальный порядок ПГ равен одиннадцати.

Прочерки в ячейках правого верхнего угла таблицы указывают на отсутствие необходимости использования больших значений q, а прочерки в ячейках

в левом нижнем углу таблицы – на то, что для формирования СЛАУ с полным рангом, равным $2n$, с данным значением числа q, потребуются ПГ порядка, большего одиннадцати. По-видимому, числа M_{nq} могут представлять самостоятельный интерес, так как матрица состоит из разных чисел, отстоящих друг от друга на шаги, одинаковые в каждой строке и в каждом столбце. Ненулевые числа M_{nq} можно определить как $M_{nq} = 2n(2n+1),\ n = 1,2,\ldots$ Соседние столбцы состоят либо только из четных, либо только из нечетных чисел.

Нулевые предначальные условия предполагают идентификацию объекта по его «стартовой реакции» или «реакции разгона», заданной входным сигналом. Отклонения от нуля предначальных условий $Y_{0-} \neq 0$ приводит к модификации алгоритма обобщенной идентификации, которая рассматривается ниже.

Возможны две ситуации с ненулевыми предначальными значениями фазовых координат объекта: в первой из них вектор $Y_{0-} = \left(y(0-)\ldots y^{(n-1)}(0-)\right)^T$ известен, во второй – нет. Если вектор $Y_{0-} \neq 0$ известен, это приводит к уточнению условий баланса регулярных и сингулярных составляющих в ЛДУ объекта, задающего финитные сигналы. Более наглядно подобные уточнения проявляются во второй ситуации, когда $Y_{0-} \neq 0$ неизвестен и требует определения в процессе идентификации.

Требуемое число ПГ сигналов вход-выход в зависимости от порядка n и числа q значений параметров $(r_{k_q}, s_{k_q}) \in R_+^2$ ПГ приведены в таблице 2. Обозначения в таблице 2 совпадают с обозначениями таблицы 1. Числа M_{nq} имеют постоянные шаги в каждой из строк и в каждом из столбцов таблицы 2. Конечные значения M_{nq} в каждой из строк равны $(2n+1)3n = 6n^2 + 3n$.

Таблица 2

n/q	**1**	**2**	**3**	**4**	**5**	**6**
1	7	8	9	-	-	-
2	15	18	21	24	27	30
3	23	28	33	38	43	48
4	-	38	45	52	59	66
5	-	-	57	66	75	84
6	-	-	-	80	91	102
7	-	-	-	-	107	120
8	-	-	-	-	-	138
9	-	-	-	-	-	-
n/q	**7**	**8**	**9**	**10**	**11**	**12**
1	-	-	-	-	-	-
2	-	-	-	-	-	-
3	53	58	63	-	-	-
4	73	80	87	94	101	108
5	93	102	111	120	129	138
6	113	124	135	146	157	168
7	133	146	159	172	185	198
8	153	168	183	198	213	228
9	173	190	207	224	241	258

Анализируя данные таблицы 2, нетрудно заметить, что учет ненулевых предначальных условий для каждого n по объему вычислительных затрат на формирование СЛАУ приблизительно эквивалентен случаю нулевых предначальных условий для ДУ порядка $(n+1)$. Порядок СЛАУ, используемого для оценки всех параметров обобщенной модели, равен $3n$. В этой связи целесообразно ограничить порядок n ЛДУ, представляющего модель объекта, величиной, равной семи. В этом случае ЛДУ, задающее связь между финитными сигналами $y(t) = y_p(t)1(t),\ \ v(t) = v_p(t)1(t),\ t \in R$, имеет вид

$$\sum_{i=0}^{7} a_i \sum_{k=0}^{6-i} \left(y_p^{(6-i-k)}(0+) - y^{(6-i-k)}(0-)\right)\delta^{(k)}(t) + \\ + \sum_{i=0}^{7} a_i y_p^{(7-i)}(t)1(t) = \sum_{j=0}^{6} b_j v_p^{(m-j)}(t)1(t) + \\ + \sum_{j=0}^{6} b_j \sum_{k=0}^{m-1-j} v_p^{(m-1-k)}(0+)\,\delta^{(k)}(t), \qquad (15)$$

$$m \in [0,6],\ a_7 = 1.$$

Функция $y_p(t),\ t \in R$ будет решением ЛДУ (15), если он обращает его в тождество на всей числовой оси $t \in R$. Это означает, что должны выполняться условия баланса сингулярных и регулярных составляющих ДУ (15):

$$а)\ \sum_{i=0}^{k} a_i \left(y_p^{(k-i)}(0+) - y^{(k-i)}(0-)\right)\delta^{(k)}(t) = \\ = \sum_{j=0}^{k+q} b_j v_p^{(k+w-j)}(0+), \qquad (16)$$

где $w = m - n$ - разность порядков операторов в правой и левой частях ДУ модели объекта, $w \in \overline{-7,-1},\ k = \overline{0,6}$,

$$б)\ \sum_{i=0}^{7} a_i y_p^{(7-i)}(t) = \sum_{j=0}^{m} b_j v_p^{(m-j)}(t),\ a_7 = 1,\ t > 0. \qquad (17)$$

Проекции ДУ (17) на координатные функции Гаусса $Q_l(t - r_{k_q}, s_{k_q}),\ l = \overline{0,11},\ (r_{k_q}, s_{k_q}) \in R_+^2$ описываются зависимостью между ПГ сигналов вход-выход:

$$-\sum_{i=0}^{7} a_i \sum_{w=0}^{6-i} y_p^{(6-i-w)}(0+) Q_{k+w}(-r,s) + \\ + \sum_{i=0}^{7} a_i G_y^{(k+7-i)}(r,s) = \sum_{j=0}^{m} b_j G_v^{(k+m-j)}(r,s) - \qquad (18) \\ - \sum_{j=0}^{m} b_j \sum_{p=0}^{m-1-p} v_p^{(m-1-p-j)}(0+) Q_{k+p}(-r,s),$$

$$m \in \overline{0,6}.$$

Используя некоторое число точек

$$(r_{k_q}, s_{k_q}) \in R_+^2,\quad k_q = \overline{1,q},\ q \in [5,21]$$

таких, что $\left|Q_l(T - r_{k_q}, s_{k_q})\right| \le \varepsilon$, где ε - некоторое малое число, на основе соотношения (18) формируем СЛАУ

$$GP = G^0, \qquad (19)$$

где неизвестные параметры модели отмечены вектором $P = \left(a_0 \ldots a_6\ b_0 \ldots b_m\ z_0 \ldots z_6\right)^T$. Размерность вектора P зависит от порядка m оператора правой части ДУ (17). Кроме этой зависимости матрицы G и G^0 зависят от числа $q \in [5,21]$ точек $(r_{k_q}, s_{k_q}) \in R_+^2$. При $m = 6$ и $q = 5$ матрицы имеют вид

$$G = \left(G^1 \ldots G^5\right)^T; \\ G^0 = (-1)\left(G_y^{(0)}(r_1,s_1) \ldots G_y^{(0)}(r_5,s_5)\right)^T, \qquad (20)$$

где блоки определяются выражениями

$$G^{k_q}=\begin{pmatrix} G^{(7)}_{yk_q}\dots G^{(1)}_{yk_q} & -G^{(6)}_{vk_q} & -G^{(0)}_{vk_q}Q_{6k_q}\dots Q_{0k_q}\\ \vdots\ \ \vdots\ \ \vdots & \vdots & \vdots\ \ \vdots\ \ \vdots\\ G^{(11)}_{yk_q}\dots G^{(6)}_{yk_q} & -G^{(10)}_{vk_q} & -G^{(4)}_{vk_q}Q_{10k_q}\dots Q_{4k_q}\end{pmatrix},$$

$$k_q=\overline{1,4}; \qquad (21)$$

$$G^5=\left(G^{(7)}_{y5}\dots G^{(1)}_{y5}-G^{(6)}_{v5}\dots-G^{(0)}_{v5}Q_{65}\dots Q_{05}\right),$$

где

$$G^{(k)}_{yk_q}=G^{(k)}_y(r_{k_q},s_{k_q}),\ k=\overline{1,11},\ k_q=\overline{1,4};$$
$$G^{(l)}_{vk_q}=G^{(l)}_v(r_{k_q},s_{k_q}),\ l=\overline{0,10},\ k_q=\overline{1,4};$$
$$Q_{mk_q}=Q_m(-r_{k_q},s_{k_q}),\ m=\overline{0,10},\ k_q=\overline{1,4}.$$

При другом подходе, полярном первому, когда $q=21,\ m=6$, структуру матриц G и G^0 иллюстрируют следующие выражения:

$$G=\left(G^1\dots G^{21}\right)^T;$$
$$G^0=(-1)\left(G^{(0)}_y(r_1,s_1)\dots G^{(0)}_y(r_{21},s_{21})\right)^T, \qquad (22)$$

где

$$G^{k_q}=\left(G^{(7)}_{yk_q}\dots G^{(1)}_{yk_q}-G^{(6)}_{vk_q}\dots-G^{(0)}_{vk_q}Q_{6k_q}\dots Q_{0k_q}\right),$$
$$k_q=\overline{1,21};$$
$$G^{(k)}_{yk_q}=G^{(k)}_y(r_{k_q},s_{k_q}),\ k=\overline{1,7},\ k_q=\overline{1,21}; \qquad (23)$$
$$G^{(l)}_{vk_q}=G^{(l)}_v(r_{k_q},s_{k_q}),\ l=\overline{0,6},\ k_q=\overline{1,21};$$
$$Q_{mk_q}=Q_m(-r_{k_q},s_{k_q}),\ m=\overline{0,6},\ k_q=\overline{1,21}.$$

Вектор $Z=\left(z_0\dots z_6\right)^T$ определяется постначальными значениями фазовых координат сигналов вход-выход:

$$z_k=\sum_{w=0}^{w+k}b_q v_p^{(w+k-w)}(0+)-\sum_{h=0}^{k}a_w y_p^{(k-h)}(0+),$$
$$k=\overline{0,6}, \qquad (24)$$

где $w=m-n$ - разность между порядками операторов правой и левой частей ДУ модели объекта.

Анализируя выражение (24), нетрудно заметить, что произвольный сигнал $v(t),\ t\in(0,T)$ должен быть идентифицирован так, чтобы все его производные $v^{(q_m)}(t),\ t=0+,\ q_m=\overline{0,m}$ были априори известны. В противном случае, вычислить вектор $Y_{0+}=\left(y(0+)\dots y^{(n-1)}(0+)\right)^T$ постначальных значений фазовых координат выходного сигнала на основе только результатов решения СЛАУ (17) становится невозможным. Аналогичное условие должно иметь место и для вектора Y_{0-} предначальных значений фазовых координат выходного сигнала.

Последовательно выдвигая гипотезы о возможном описании модели объекта ДУ с 1-го по 7-й порядки, формируется банк моделей-кандидатов (всего 28 моделей). Для каждой из них формируется и решается СЛАУ соответствующего порядка, позволяющая вычислить вектор $P=\left(a_0\dots a_{n-1}\ b_0\dots b_m\ z_0\dots z_{n-1}\right)^T$ параметров и вектор $Y_{0+}=\left(y(0+)\dots y^{(n-1)}(0+)\right)^T$ постначальных значений фазовых координат выходного сигнала. Для моделей-кандидатов определяются выходные сигналы $y_{Mi}(t),\ i=\overline{1,28},\ t\in(0,T)$. Объект идентифицируется той моделью, для которой критерий соответствия модель-объект (14) принимает значение $J\to\min\limits_{j\in[1,28]}$. Если значение критерия определенной модели-кандидата удовлетворяет условию $J\le\varepsilon_\partial$, где ε_∂ - допустимая погрешность аппроксимации выходного сигнала объекта выходным сигналом модели, то модель объекта пригодна для оценки вектора Y_{T-} предконечных значений фазовых координат выходного сигнала объекта и прогнозирования реакций объекта на входные сигналы $v(t),\ t>T$. В противном случае необходимо взять другой набор точек $(r_{k_q},s_{k_q})\in R^2_+$ и повторить процедуру идентификации

Экспериментальные исследования

Вначале рассмотрим пример обобщенной идентификации объекта простой структуры, имеющий достаточно высокую точность по сравнению с идентификацией объектов усложненной и сложной структуры [4]. Сгенерируем случайным образом объект, описываемый ДУ четвертого порядка:

$$\sum_{i=0}^{4}a_i p^{4-i}y(t)=\sum_{j=0}^{3}b_j p^{3-j}v(t),$$
$$a_4=1, t\in(0,T),$$

а операторы передаточной функции явно имеют вид:

$$B(p)=0{,}243p^3+1{,}593p^2+3{,}953p+2{,}119;$$
$$A(p)=0{,}112p^4+0{,}739p^3+2{,}93p^2+3{,}168p+1. \qquad (25)$$

Подадим на вход объекта сигнал $v(t)=\sin\omega t,\ \omega=1$ и с помощью какого-либо решателя ДУ получим выходной сигнал $y(t), t\in(0{,}10)$. Положим, что объект стартует из нулевых $Y_{0-}=0$ предначальных условий. Воспользовавшись рассмотренным алгоритмом идентификации и выбрав параметры ПГ *r*=0,5 и *s*=0,5 , а так же шаг изменения параметра *r*: *h*=0,5, получаем банк моделей-кандидатов (до десятого порядка), из которых выбирается одна на основе критерия:

$$J=\frac{\sum\limits_{i=0}^{N}\left(y(t_i)-y_M(t_i)\right)^2}{\sum\limits_{i=0}^{N}y_i^2}100\%. \qquad (26)$$

Операторы передаточной функции полученной в результате идентификации «наилучшей» модели объекта, имеют вид:

$$B(p)=0{,}243p^3+1{,}59p^2+3{,}947p+2{,}118; \qquad (27)$$
$$A(p)=0{,}112p^4+0{,}738p^3+2{,}925p^2+3{,}165p+1,$$

значение критерия (26) – $J=5{,}1\cdot10^{-9}\%$. Минимизация критерия *J* возможна за счет выбора параметров ПГ *r* и *s*, а также выбора шага изменения *h* параметра *r*. Рекомендации по выбору этих параметров были подробно рассмотрены в работе [4].

Вычислим вектор $Y_{0+}=\left(y(0+)\dots y'''(0+)\right)^T$ постначальных фазовых координат модели объекта по расчетной формуле:

$$y(0+)=y'(0+)=0;$$
$$y''(0+)=\frac{b_0}{a_0}\omega; \qquad (28)$$
$$y'''(0+)=\frac{1}{a_0}\left(b_1\omega-a_1y''(0+)\right).$$

В результате имеем:

$$y(0+)=y'(0+)=0;$$
$$y''(0+)=2{,}168;$$
$$y'''(0+)=-0{,}086.$$

Наложим на выходной сигнал объекта (25) аддитивный гауссовский шум с математическим ожиданием, равным нулю, и дисперсией, равной 0,05. Теперь необходимо идентифицировать объект по входному сигналу $v(t)=\sin t$ и оценке

$$\hat{y}(t)=y(t)+n_y$$

выходного сигнала. Воспользовавшись вышеизложенным алгоритмом и выбрав параметры ПГ *r*=0,1 и *s*=1,5 , а так же шаг изменения параметра *r*: *h*=0,6, получаем банк моделей-кандидатов, из которых выбирается одна на основе критерия (26). Операторы передаточной функции полученной модели имеют вид:

$$\begin{aligned}B(p)&=0{,}236p^3+0{,}663p^2+1{,}198p+2{,}215;\\ A(p)&=0{,}057p^4+0{,}419p^3+0{,}715p^2+1{,}864p+1,\end{aligned}\qquad(29)$$

а значение критерия (26) - $J=4{,}11\%$. Компоненты вектора фазовых координат определенной модели объекта в постначальный момент времени в данном случае имеют вид:

$$y(0+)=y'(0+)=0;$$
$$y''(0+)=4{,}102;$$
$$y'''(0+)=-18{,}358.$$

На рис. 4 представлены входной сигнал объекта и выходные сигналы идентифицированной модели $y_M(t)$ и объекта $y(t)$.

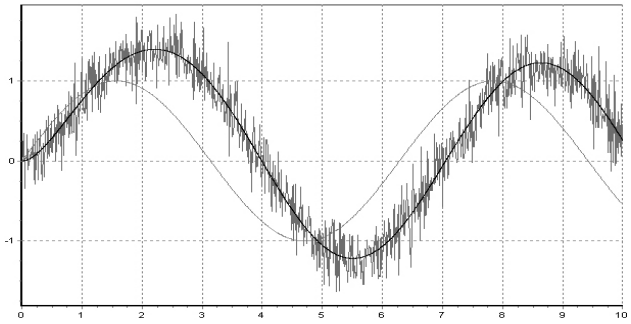


Рис. 4 – Графики $v(t), y_M(t), \hat{y}(t)$

На рис. 4 видно, что выходной сигнал $y_M(t)$ полученной модели повторяет характер изменения зашумленного выходного сигнала $\hat{y}(t)$ объекта, очищенного от шума.

Теперь рассмотрим сгенерированный случайным образом объект, описываемый ДУ десятого порядка:

$$\sum_{i=0}^{10}a_i p^{10-i}y(t)=\sum_{j=0}^{9}b_j p^{9-j}v(t),$$
$$a_{10}=1, t\in(0,T),$$

операторы передаточной функции которого, имеют вид:

$$\begin{aligned}B(p)&=6{,}443\cdot10^{-4}p^9+1{,}87\cdot10^{-2}p^8+\\ &+6{,}04\cdot10^{-2}p^7+0{,}344p^6+0{,}648p^5+1{,}137p^4+\\ &+1{,}12p^3+2{,}484p^2+4{,}875p+2{,}358;\end{aligned}\qquad(30)$$

$$\begin{aligned}A(p)&=8{,}664\cdot10^{-4}p^{10}+7{,}49\cdot10^{-3}p^9+\\ &+7{,}695\cdot10^{-2}p^8+0{,}382p^7+1{,}722p^6+\\ &+5{,}247p^5+11{,}972p^4+19{,}652p^3+\\ &+18{,}041p^2+7{,}423p+1.\end{aligned}$$

Подадим на вход полученного объекта сигнал $v(t)=e^{-t}$, и с помощью какого-либо решателя ДУ получим выходной сигнал $y(t), t\in(0,10)$. Положим, что объект стартует из нулевых $Y_{0-}=0$ предначальных условий. Воспользовавшись предложенным в данной работе алгоритмом и выбрав параметры ПГ *r*=0,1 и *s*=0,5 , а так же шаг изменения параметра *r*: *h*=0,2, получаем банк моделей-кандидатов, из которых выбирается одна на основе критерия (26). Операторы передаточной функции полученной в результате идентификации «наилучшей» модели объекта имеют вид:

$$\begin{aligned}B(p)&=6{,}96\cdot10^{-4}p^8+1{,}916\cdot10^{-2}p^7+\\ &+4{,}212\cdot10^{-2}p^6+0{,}32p^5+0{,}319p^4+\\ &+0{,}867p^3+0{,}282p^2+2{,}393p+2{,}359;\\ A(p)&=3{,}6\cdot10^{-7}p^{10}+9{,}288\cdot10^{-4}p^9+\\ &+6{,}934\cdot10^{-3}p^8+7{,}426\cdot10^{-2}p^7+\\ &+0{,}322p^6+1{,}469p^5+3{,}90p^4+8{,}322p^3+\\ &+11{,}451p^2+6{,}382p+1,\end{aligned}\qquad(31)$$

значение критерия (26) - $J=2{,}665\cdot10^{-5}\%$.

Вычислим вектор $Y_{0+}=\left(y(0+)\dots y^{(9)}(0+)\right)^T$ постначальных фазовых координат модели объекта по расчетной формуле:

$$y^{(q)}(0+)=0,\ q=n-m-1;$$
$$y^{(n-m)}(0+)=\frac{b_0}{a_0};\qquad(32)$$
$$y^{(n-m+k)}(0+)=\frac{1}{a_0}\left(\sum_{j=0}^{k}(-1)^{k-j}b_j-\sum_{p=1}^{k-1}a_p y^{(k-p)}(0+)\right),$$
$$a_0\neq0,\ k=\overline{0,n-1},\ m=\overline{0,n-1},\ n=\overline{1,10}.$$

В результате имеем:

$$y(0+)=0;\ y'(0+)=1931{,}79;\ y''(0+)=-4{,}93\cdot10^6;$$
$$y'''(0+)=1{,}27\cdot10^{10};\ y^{(4)}(0+)=-3{,}25\cdot10^{13};$$
$$y^{(5)}(0+)=8{,}35\cdot10^{16};\ y^{(6)}(0+)=-2{,}14\cdot10^{20};$$
$$y^{(7)}(0+)=5{,}51\cdot10^{23};\ y^{(8)}(0+)=-1{,}42\cdot10^{27};$$
$$y^{(9)}(0+)=3{,}63\cdot10^{30}.$$

Наложим на выходной сигнал объекта (30) аддитивный гауссовский шум с математическим ожиданием, равным нулю, и дисперсией, равной 0,005. Теперь необходимо идентифицировать объект по входному сигналу $v(t)=e^{-t}$ и оценке

$$\hat{y}(t)=y(t)+n_y$$

выходного сигнала. Воспользовавшись вышеизложенным алгоритмом и выбрав параметры ПГ *r*=0,01 и *s*=0,8, а так же шаг изменения параметра *r*: *h*=0,25, получаем банк моделей-кандидатов, из которых выбирается одна на основе критерия (26). Операторы передаточной

функции полученной модели имеют вид:

$$B(p) = 1{,}674 \cdot 10^{-4} p^8 + 7{,}389 \cdot 10^{-4} p^7 + \\ + 2{,}58 \cdot 10^{-2} p^6 + 2{,}895 \cdot 10^{-2} p^5 + 0{,}36 p^4 + \\ + 1{,}827 \cdot 10^{-2} p^3 + 1.038 p^2 - 0{,}224 p + 2{,}473; \quad (33)$$

$$A(p) = 6{,}366 \cdot 10^{-5} p^9 + 9{,}669 \cdot 10^{-4} p^8 + \\ + 1{,}032 \cdot 10^{-2} p^7 + 7{,}104 \cdot 10^{-2} p^6 + 0{,}339 p^5 + \\ + 1{,}235 p^4 + 3{,}026 p^3 + 5{,}349 p^2 + 5{,}533 p + 1,$$

значение критерия (26) - $J = 5{,}01\%$. Из выражения (33) следует, что передаточная функция «наилучшей» модели имеет девятый, а не десятый, как у объекта, порядок.

Вычислим по формуле (32) вектор постначальных значений фазовых координат модели объекта:

$$y(0+) = 0;\ y'(0+) = 2{,}63;\ y''(0+) = -30{,}98;$$
$$y'''(0+) = 440{,}44;\ y^{(4)}(0+) = -4544{,}69;$$
$$y^{(5)}(0+) = 23798{,}3;\ y^{(6)}(0+) = -7719{,}86;$$
$$y^{(7)}(0+) = -515208;\ y^{(8)}(0+) = -613843.$$

На рис. 5 представлены входной сигнал объекта, выходной сигнал определенной модели объекта $y_M(t)$ и выходной $\hat{y}(t)$ зашумленный сигнал объекта.

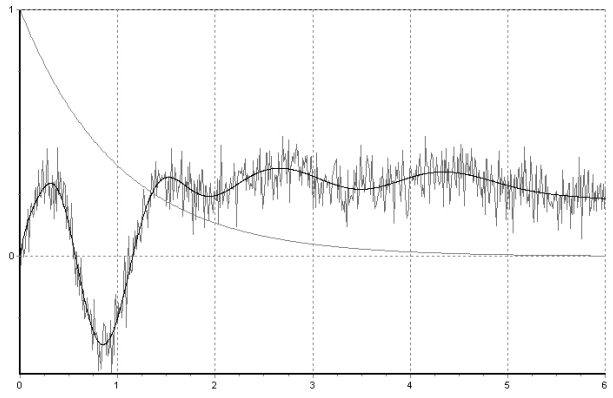


Рис. 5 – Графики $\mathbf{v(t), y_M(t), \hat{y}(t)}$

На рис. 5 видно, что выходной сигнал $y_M(t)$ полученной модели является математическим ожиданием зашумленного выходного сигнала $\hat{y}(t)$ объекта.

Заключение

В статье предлагается метод обобщенной идентификации канала многомерного линейного стационарного объекта, основанный на двупараметрических интегральных преобразованиях Гаусса, позволяющих осуществлять удобную алгебраизацию ЛДУ, описывающего динамику канала объекта. Приведен конкретный алгоритм идентификации при произвольном входном сигнале из заданного класса, позволяющий определить коэффициенты ЛДУ канала и векторы значений его фазовых координат в предоконечный момент времени. Данный подход позволяет далее решать задачу управления объектом согласно идеи дуального управления Фельдбаума. Его удалось реализовать за счет предложенной в работе модификаций ЛДУ, задающих финитный выходной сигнал при финитном входе. Для каналов объекта простой структуры предложенный метод довольно точен. Для каналов высокого порядка, коэффициенты ЛДУ модели мало совпадают, но его решение достаточно точно аппроксимирует реакцию канала объекта.

Следует особо подчеркнуть неизбежный субъективизм процедур подтверждения модели объекта. Как показывает практика, решение данного вопроса никогда не следует полностью передоверять каким бы то ни было статистическим процедурам [5].

Литература

1. Уваев А.И. Идентификация линейных стационарных объектов на основе преобразований Гаусса. //Сб. трудов Всесоюзной научно-технической конференции «Автоматизация исследования, проектирования и испытаний сложных технических систем».- Калуга, 1989. - с.167-168.
2. Уваев А.И., Минко А.А. Исследование преобразований Гаусса гармонических сигналов. // Проблемы математического моделирования и обработки информации в научных исследованиях: сборник научных трудов. Рязань, 1999. - с. 83-88.
3. Уваев А.И., Минко А.А. Особенности обобщенной параметрической идентификации объектов управления на основе модели вход-выход. // Новые информационные технологии в научных исследованиях и образовании. Материалы XII Всероссийской научно-технической конференции студентов. Рязанский государственный радиотехнический университет, 2007. - с. 210-212.
4. Минко А.А. Идентификация линейного объекта по реакции на гармонический сигнал. Вестник РГРТУ №1(выпуск 23), 2008. - с. 63-68.
5. Льюнг Л. Идентификация систем. Теория для пользователя: Пер. с англ./ Под ред. Я.З. Цыпкина. – М.: Наука. Гл. ред. Физ.-мат. Лит., 1991. – 432 с.

УДК 621.396.2

ИССЛЕДОВАНИЕ СПЕКТРА ПОЛИНОМИАЛЬНЫХ СПЛАЙНОВ

Просочкин А.С.

Рассматривается изменение спектра сигнала, аппроксимируемого полиномиальным сплайном произвольной степени m, представленным через базисные B-сплайны, коэффициенты которых определяются в виде взвешенной суммы произведений отсчётов исходного сигнала на соответствующие значения B-сплайнов.

Для приближения, восстановления и фильтрации сигналов, заданных в виде функций времени $f(t)$ на равномерной сетке t_i $(i=-\infty, \ldots, -2, -1, 0, 1, 2, \ldots, +\infty,$ $\Delta t=t_i - t_{i+1} = const$ - период дискретизации сигнала) значениями $f(t_i)$ с некоторой погрешностью, широко используются методы полиномиальной сплайн-аппроксимации [1]. Распространённым способом построения полиномиальных сплайнов произвольной степени m является их представление через базисные *B*-сплайны в виде [1]

$$S^m(t) = \sum_i C_i \cdot B_i^m(t), \qquad (1)$$

где C_i - коэффициенты аппроксимации;

$B_i^m(t)$ - базисные *B*-сплайны степени m, которые имеют отличные от нуля значения на $(m+1)$ участках сплайна, называемых интервалом носителем *B*-сплайна.

Значения *B*-сплайна на границах участков называются его узлами.

В общем случае каждый участок базисного *B*-сплайна может включать d интервалов $\Delta t=t_i - t_{i+1}$ $(d=1, 2, 3, \ldots)$. На рис. 1 приведено множество *B*-сплайнов третьей степени $(m=3)$ при $d=2$.

Спектр дискретного сигнала $f(t)$, заданного в виде значений $f_i = f(t_i)$ на сетке t_i, определяется как [2]

$$\dot{F}(\omega) = \sum_{i=-\infty}^{\infty} f_i \cdot e^{-j\omega i}. \qquad (2)$$

Из формулы (2) следует, что спектр дискретного сигнала является периодическим с периодом равным 2π.

Спектр сплайна (1), заданного в виде значений $S^m(t_i)$ на сетке t_i, принимает вид

$$\dot{F}_{Sm}(\omega) = \sum_{i=-\infty}^{\infty} S^m(t_i) \cdot e^{-j\omega i}. \qquad (3)$$

Так как при аппроксимации сигнала полиномиальным сплайном происходит изменение спектра, рассмотрим вопрос, как изменится спектр $\dot{F}_{Sm}(\omega)$ сплайна $S^m(t)$ произвольной степени m, заданного в виде значений $S^m(t_i)$ на равномерной сетке t_i по отношению к спектру $\dot{F}(\omega)$ исходного сигнала $f(t_i)$ $(i=-\infty, \ldots, -1, 0, 1, \ldots, +\infty)$.

Значение сплайна степени m в момент времени t_i при $d=1$ определяется в виде [1]

$$S^m(t_i) = \sum_{k=-(l-1)}^{(l-1)} C_{i+k} \cdot B_{-k}^m,$$

где l – число узлов базисного *B*-сплайна на половине интервала носителя, в которых его значения отличны от нуля;

B_k^m – величина *B*-сплайна степени m в k-м узле его интервала носителя.

Величина l зависит от степени сплайна и определяется выражением

$$l = \begin{cases} \dfrac{m+1}{2} & \text{- для сплайнов нечётной степени } m, \\ \dfrac{m}{2}+1 & \text{- для сплайнов чётной степени } m. \end{cases} \qquad (4)$$

Если каждый участок включает d интервалов Δt, то величина интервала носителя *B*-сплайна равна $d \cdot \Delta t \cdot (m+1)$, а значение сплайна степени m в момент времени t_i определяется по формуле

$$S^m(t_i) = \sum_{k=-(l-1)}^{(l-1)} C_{i+d\cdot k} \cdot B_{-d\cdot k}^m. \qquad (5)$$

Например, для кубического сплайна $(m=3,\ l=2)$ при $d=2$ (см. рис. 1) из формулы (5) получим

$$S^3(t_i) = C_{i-2} \cdot B_2^3 + C_i \cdot B_0^3 + C_{i+2} B_{-2}^3.$$

Значения B_2^3, B_0^3, B_{-2}^3 отмечены на рис.1 точками.

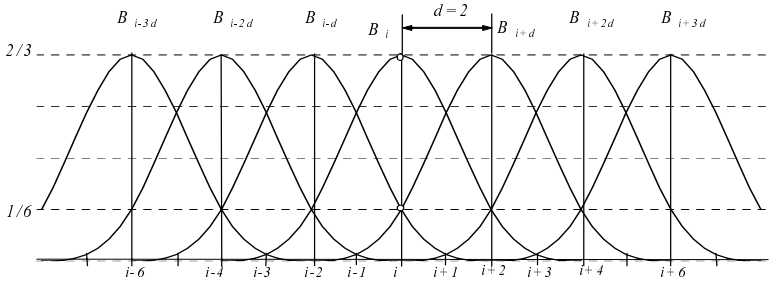


Рис. 1. Множество базисных кубических B-сплайнов при d=2.

Подставив значения $S^m(t_i)$ из формулы (5) в формулу (3), получим спектр полиномиального сплайна произвольной степени m в виде

$$\dot{F}_{Sm}(\omega) = \sum_{i=-\infty}^{\infty} \left(\sum_{k=-(l-1)}^{(l-1)} C_{i+d\cdot k} \cdot B^m_{-d\cdot k} \right) \cdot e^{-j\omega i}. \quad (6)$$

Для определения полиномиальных сплайнов (1) применяются разные способы вычисления коэффициентов C_i, основанные на минимизации функционалов, характеризующих отклонение сплайна $S^m(t_i)$ от исходного сигнала $f(t_i)$ [1]. В общем случае вычисление коэффициентов C_i производится путём решения систем алгебраических уравнений и требует значительных затрат. На практике широко используются способы приблизительного вычисления коэффициентов с помощью явных формул [3]. В таких случаях значение коэффициента C_i вычисляется по значению исходного сигнала f_i в точке t_i и значениях сигнала в точках, близких к t_i.

Рассмотрим случай, когда коэффициент C_i, соответствующий *i*-му *B*-сплайну $B_i^m(t)$, определяется в виде взвешенной суммы произведений отсчётов исходного сигнала, которые попадают в интервал носитель i-го *B*-сплайна, на соответствующие значения *B*-сплайна (далее индекс m в обозначении *B*-сплайна опущен)

$$C_i = \frac{1}{d} \cdot \sum_{n=-(d\cdot l-q)}^{(d\cdot l-q)} B_n \cdot f_{i+n}.$$

где q – постоянная величина, значение которой зависит от m и d.

Для сплайнов нечётной степени m: $q=1$.

Для сплайнов чётной степени m:

$$q = \begin{cases} \dfrac{d+1}{2} & \text{- если } d \text{ - нечётное число,} \\ \dfrac{d}{2}+1 & \text{- если } d \text{ - чётное число.} \end{cases}$$

Число $2\cdot(d\cdot l-q)+1$ – определяет количество значений исходного сигнала, попадающих в интервал носитель соответствующего *B*-сплайна.

Таким образом

$$S^m(t_i) = \frac{1}{d} \cdot \sum_{k=-(l-1)}^{(l-1)} B_{-d\cdot k} \cdot \sum_{n=-(d\cdot l-q)}^{(d\cdot l-q)} B_n \cdot f_{i+d\cdot k+n}. \quad (7)$$

С учётом формулы (7) спектр полиномиального сплайна произвольной степени m, заданного на сетке t_i своими значениями $S^m(t_i)$, определяется выражением

$$\dot{F}_{Sm}(\omega) = \sum_{i=-\infty}^{\infty} \left\{ \frac{1}{d} \cdot \sum_{k=-(l-1)}^{(l-1)} B_{-d\cdot k} \cdot \sum_{n=-(d\cdot l-q)}^{(d\cdot l-q)} B_n \cdot f_{i+d\cdot k+n} \right\} \cdot e^{-j\omega i}. \quad (8)$$

Учитывая свойство симметрии базисных *B*-сплайнов

$$B_n = B_{-n}, \quad n \in [-(d\cdot l-q);+(d\cdot l-q)] \quad (9)$$

и группируя слагаемые в сумме по k, получим

$$\dot{F}_{Sm}(\omega) = \frac{1}{d} \cdot \sum_{i=-\infty}^{\infty} \left\{ B_{d\cdot(l-1)} \cdot \left[\sum_{n=-(d\cdot l-q)}^{(d\cdot l-q)} B_n \cdot f_{i-d\cdot(l-1)+n} + \sum_{n=-(d\cdot l-q)}^{(d\cdot l-q)} B_n \cdot f_{i+d\cdot(l-1)+n} \right] + \right.$$

$$+ B_{d\cdot(l-2)} \cdot \left[\sum_{n=-(d\cdot l-q)}^{(d\cdot l-q)} B_n \cdot f_{i-d\cdot(l-2)+n} + \sum_{n=-(d\cdot l-q)}^{(d\cdot l-q)} B_n \cdot f_{i+d\cdot(l-2)+n} \right] + \dots$$

$$\left. \dots + B_0 \cdot \left[\sum_{n=-(d\cdot l-q)}^{(d\cdot l-q)} B_n \cdot f_{i+n} \right] \right\} \cdot e^{-j\omega i}. \quad (10)$$

Выполнив преобразования в квадратных скобках формулы (10) с учётом свойства симметрии (9) базисных *B*-сплайнов и применив перестановку операций суммирования, получим

$$\dot{F}_{Sm}(\omega) = \frac{1}{d} \cdot \left\{ B_{d\cdot(l-1)} \cdot \left[\sum_{n=1}^{(d\cdot l-q)} B_n \cdot \sum_{i=-\infty}^{\infty} \left(f_{i-d\cdot(l-1)+n} + f_{i+d\cdot(l-1)-n} \right) \cdot e^{-j\omega i} + \right. \right.$$

$$+ \sum_{n=1}^{(d\cdot l-q)} B_n \cdot \sum_{i=-\infty}^{\infty} \left(f_{i-d\cdot(l-1)+n} + f_{i+d\cdot(l-1)-n} \right) \cdot e^{-j\omega i} +$$

$$\left. + B_0 \cdot \sum_{i=-\infty}^{\infty} \left(f_{i-d\cdot(l-1)} + f_{i+d\cdot(l-1)} \right) \cdot e^{-j\omega i} \right] +$$

$$B_{d\cdot(l-2)} \cdot \left[\sum_{n=1}^{(d\cdot l-q)} B_n \cdot \sum_{i=-\infty}^{\infty} \left(f_{i-d\cdot(l-2)+n} + f_{i+d\cdot(l-2)-n} \right) \cdot e^{-j\omega i} + \right.$$

$$+ \sum_{n=1}^{(d\cdot l-q)} B_n \cdot \sum_{i=-\infty}^{\infty} \left(f_{i-d\cdot(l-2)+n} + f_{i+d\cdot(l-2)-n} \right) \cdot e^{-j\omega i} +$$

$$\left. + B_0 \cdot \sum_{i=-\infty}^{\infty} \left(f_{i-d\cdot(l-2)} + f_{i+d\cdot(l-2)} \right) \cdot e^{-j\omega i} \right] + \dots$$

$$\left. \dots + B_0 \cdot \left[\sum_{n=1}^{(d\cdot l-q)} B_n \cdot \sum_{i=-\infty}^{\infty} \left(f_{i+n} + f_{i-n} \right) \cdot e^{-j\omega i} + B_0 \cdot \sum_{i=-\infty}^{\infty} f_i \cdot e^{-j\omega i} \right] \right\}. \quad (11)$$

Можно показать, что с учётом свойств преобразования Фурье для сигнала с задержкой [2], которое утверждает, что спектр сигнала, сдвинутого во времени на величину τ, соответствует спектру исходного сигнала, умноженному на комплексную экспоненту $e^{-j\omega\tau}$, формула (11) преобразуется к виду

$$\dot{F}_{Sm}(\omega) = \dot{F}(\omega) \cdot \frac{4}{d} \cdot \left[\frac{B_0}{2} + \sum_{k=1}^{(l-1)} B_{d\cdot k} \cdot COS(\omega \cdot d \cdot k) \right] \cdot \left[\frac{B_0}{2} + \sum_{n=1}^{(d\cdot l-q)} B_n \cdot COS(\omega \cdot n) \right]. \quad (13)$$

Таким образом, формула (13), устанавливает связь спектра исходного аппроксимируемого сигнала и спектра соответствующего ему полиномиального сплайна произвольной степени, который характеризуется произведением двух функций, определяемых в виде рядов Фурье.

Число *l* членов первого ряда пропорционально степени m полиномиального сплайна. Коэффициентами данного ряда являются значения функции $B_i^m(t)$, соответствующие узлам *B*-сплайна. Число $(d\cdot l-q)$ членов второго ряда зависит от степени m

сплайна и числа *d* периодов дискретизации, которые образуют интервал между узлами сплайна. Коэффициентами второго ряда являются значения *B*-сплайнов, соответствующие отсчётам $f(t_i)$ исходного сигнала.

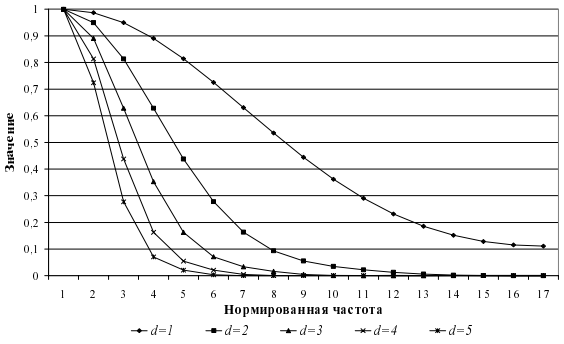


Рис. 2 Амплитудный спектр одиночной импульсной функции (14), приближённой сплайном степени m=3.

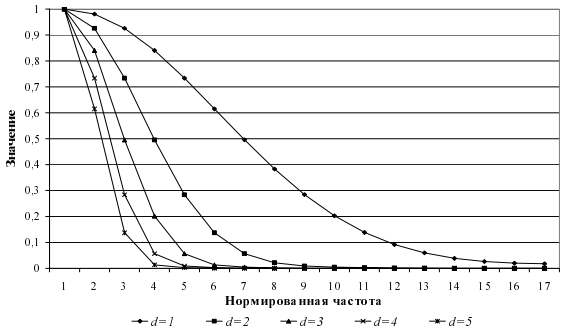


Рис. 3 Амплитудный спектр одиночной импульсной функции (14), приближённой сплайном степени m=5.

Для численного эксперимента в качестве исходного дискретного сигнала $f_1(t_i)$ взята одиночная импульсная функция [2], которая задана одним значением в центре интервала $t \in [t_0, t_N]$. Значение сигнала $f_1(t_i)$ выбрано из условия

$$\frac{1}{N} \cdot \sum_{0}^{N} f_1(t_i) = 1 \cdot$$

Следовательно,

$$f_1(t_i) = \begin{cases} N & \text{- при } i = \dfrac{N}{2}, \\ 0 & \text{- при } i \neq \dfrac{N}{2}. \end{cases} \qquad (14)$$

Такая функция имеет равномерный амплитудный спектр с амплитудой, равной единице для всех составляющих спектра.

Амплитудные спектры полиномиальных сплайнов (7) различных степеней m, приближающих одиночную импульсную функцию (14), определённые в соответствии с формулой (13) для $N=64$ и разных значений d, приведены на рис. 2-4.

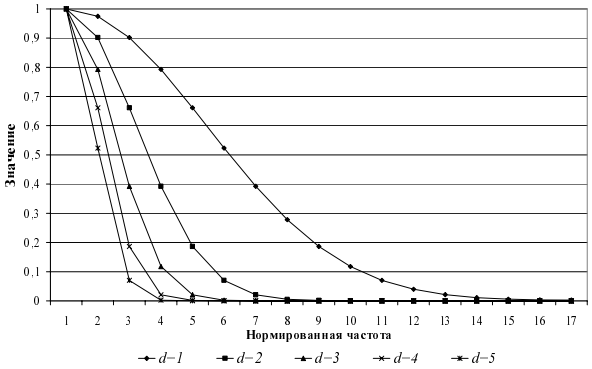


Рис. 4 Амплитудный спектр одиночной импульсной функции (14), приближённой сплайном степени m=7.

При этом амплитудные спектры соответствующих полиномиальных сплайнов, определённые с помощью дискретного преобразования Фурье, полностью совпадают со спектрами, приведёнными на рис. 2-4.

Таким образом, спектр полиномиального сплайна произвольной степени, представленного через базисные *B*-сплайны, коэффициенты которых определяются в виде взвешенной суммы произведений отсчётов исходного сигнала на соответствующие значения *B*-сплайнов, связан со спектром соответствующего исходного сигнала формулой (13), а амплитудный спектр сплайна в отличие от амплитудного спектра исходного сигнала, имеет затухающий характер, следовательно операция восстановления может рассматриваться как низкочастотная фильтрация сигнала.

Литература

1. Завьялов Ю.С., Квасов Б.И., Мирошниченко В.Л. Методы сплайн-функций. – М.: Наука, 1980. – 352 с.
2. Сергиенко А.Б. Цифровая обработка сигналов.– СПб.: Питер, 2002.– 608с.
3. Гребенников А.И. Метод сплайнов и решение некорректных задач теории приближений. – М.: МГУ, 1983. – 208 с.

УДК 621.317.08

ВЕСОВЫЕ ФУНКЦИИ ДЛЯ АДАПТИВНОГО ГАРМОНИЧЕСКОГО АНАЛИЗА СИГНАЛОВ С МНОГОМОДОВЫМ СПЕКТРОМ

Давыдочкин В.М., Давыдочкина С.В.

Предложен метод расчёта весовых функций, который позволяет адаптировать их параметры для получения одновременного минимума методической погрешности измерения частоты и амплитуды сигнала при спектральном анализе. Приведён каталог весовых функций с минимальным уровнем боковых лепестков при заданной ширине основного лепестка и заданной скорости уменьшения уровня боковых лепестков.

Введение

Один из наиболее эффективных и часто используемых методов обработки сигналов при измерении их параметров выполняется на основе преобразования Фурье (ПФ). Основные недостатки этого метода обусловлены значительной методической погрешностью из-за низкой разрешающей способности и искажающих действий боковых лепестков. Эффективным способом снижения методической погрешности является применение сглаживающих весовых функций (ВФ). Этому вопросу посвящено много публикаций, поток которых не ослабевает и в настоящее время [1, 2, 3]. Обработка данных с помощью ВФ позволяет ослабить влияние боковых лепестков, но лишь за счёт ухудшения спектрального разрешения. Считается, что в результате этих трудно разрешимых противоречий при использовании классического спектрального анализа погрешность оценки частоты сигнала, представленного коротким отрезком гармонического колебания не может быть низкой [4]. В [5] предложен метод минимизации методической погрешности оценки частоты отрезка моногармонического сигнала итерационной процедурой изменения формы ВФ. Для этого могут быть использованы ВФ с варьируемыми параметрами, к которым относятся ВФ Дольфа-Чебышева (ДЧ) и ВФ Кайзера-Бесселя (КБ). В то же время наиболее типичной является задача измерения частоты и амплитуды сигнала на фоне помех.

Цель настоящей работы – получение ВФ, позволяющих устранить или минимизировать методическую погрешность измерения частот и амплитуд слагаемых отрезка полигармонического сигнала, а также получение ВФ с минимальным уровнем боковых лепестков (УБЛ) спектральной плотности (СП) при заданной ширине основного лепестка и заданной скорости уменьшения УБЛ.

Метод расчёта адаптируемых весовых функций

ВФ, которые позволяют изменять их параметры для получения заданных спектральных свойств назовём адаптируемыми ВФ (АВФ).

При спектральном анализе за оценку частоты наиболее часто принимают частоту, соответствующую максимуму спектра сигнала [6]. В соответствии с таким определением для оценки частоты сигнала необходимо решить уравнение

$$\frac{d}{dx}\left|\sum_{i=1}^{N_c} e^{j(\Phi_i-\Phi_1)} S_i(x-x_i)+\sum_{i=1}^{N_c} e^{-j(\Phi_i+\Phi_1)} S_i(x+x_i)\right|^2 = 0, \quad (1)$$

где $e^{j(\Phi_i-\Phi_1)} S_i(x)$ - слагаемые СП

$$\dot{S}(x) = \int_{-\infty}^{\infty} w(t)\, u(t) \exp(-j2\pi x t) dt,$$

аргументами которых являются, соответственно, частоты $(x+x_i)$, $(x-x_i)$; $\dot{S}(x)$ – СП взвешенной выборки суммы гармонических сигнала и помехи

$$u(t) = \sum_{i=1}^{N} U_i(t) \cos\left[\varphi_i(t) + \Phi_i\right],$$

полученной на симметричном нормированном временном интервале $\lfloor -0{,}5T \ldots 0{,}5T \rfloor$; N_c - число гармонических слагаемых сигнала; $x = \omega T (2\pi)^{-1}$; $t = t_{абс}/T$ – нормированная частота сигнала и нормированное время; U_i, Φ_i – амплитуда и постоянная составляющая фазы i - го компонента сигнала; $w(t)$ – ВФ симметричная относительно середины временного интервала. Разумеется, результаты, полученные для симметричного временного интервала, будут справедливы для несимметричного временного интервала с учётом известных фазовых соотношений [7].

Будем считать, что $S_1(x_1)$ соответствует сигналу, частоту которого x_1 требуется определить.

Среди множества возможных решений уравнения (1) существует ряд частных, из которых следует, что для того чтобы оценка частоты сигнала x_1 была не смещена необходимо, чтобы на этой частоте первые производные всех мешающих слагаемых СП были равны нулю

$$\begin{cases} \dfrac{d}{dx} S_i(x-x_i) = 0, \quad i = \overline{2, N_c}; \\ \dfrac{d}{dx} S_i(x+x_i) = 0, \quad i = \overline{1, N_c}. \end{cases} \quad (2)$$

При использовании известных ВФ совокупность решений, определяемая системой уравнений (2), соответствует значениям частот, при которых совпадают положения максимума основного лепестка слагаемого спектра $S_1(x-x_1)$ с положениями экстремумов боковых лепестков слагаемых $S_1(x+x_1)$, $S_i(x-x_1)$ и $S_i(x+x_1)$. В этом случае частота x_1 для любой применяемой ВФ измеряется без погрешности независимо от фаз слагаемых сигнала. Однако погрешность оценки амплитуды сигнала определяется уровнем боковых лепестков и при использовании известных ВФ достигает экстремума, который, в частности, может совпадать с максимумом. Очевидно, что для того чтобы оценка амплитуды сигнала на частоте x_1 соответствовала истинному

значению, на этой частоте должны быть равны нулю все мешающие слагаемые СП.

В общем случае можно заключить, что если мешающие слагаемые СП на заданной частоте x_1 аппроксимируются нулем на основе формулы Тейлора, то оценки частоты и амплитуды сигнала не будут смещены. Таким образом, задача устранения методической погрешности измерения частот и амплитуд слагаемых отрезка полигармонического сигнала может быть решена, если существует ВФ, форма которой может задаваться варьируемыми параметрами таким образом, чтобы на частоте каждого из N_c слагаемых сигнала СП других слагаемых были равны нулю вместе с заданным количеством M их производных, т.е. если существует не тривиальное решение системы уравнений

$$\begin{cases} S(x)\big|_{x=b_1}=0, \\ \ldots\ldots\ldots\ldots\ldots \\ S(x)\big|_{x=b_{2N_c-1}}=0, \\ \frac{d}{dx}S(x)\Big|_{x=b_1}=0, \\ \ldots\ldots\ldots\ldots\ldots \\ \frac{d}{dx}S(x)\Big|_{x=b_{2N_c-1}}=0, \\ \ldots\ldots\ldots\ldots\ldots \\ \frac{d^{(M)}}{dx^{(M)}}S(x)\Big|_{x=b_1}=0, \\ \ldots\ldots\ldots\ldots\ldots \\ \frac{d^{(M)}}{dx^{(M)}}S(x)\Big|_{x=b_{2N_c-1}}=0. \end{cases} \quad (3)$$

где b_i - частота задаваемого нуля СП АВФ.

Так как форма спектра отрезка гармонического сигнала с точностью до постоянного сомножителя совпадает с формой спектра ВФ [7], достаточно получить решение для заданных частот спектра ВФ, при этом $b_1=2x_1$; $b_i=|x_1-x_i|$, $i=\overline{2,N_c}$ для слагаемых СП сигнала с положительной полуоси частот; $b_i=x_1+x_i$, $i=\overline{2,N_c}$ для слагаемых СП с отрицательной полуоси частот.

Для определения возможностей исключения методических погрешностей рассмотрим ВФ, представимые тригонометрическими рядами

$$w_{sN}(t,b_1,\ldots,b_{2N_c-1})= 1+\sum_{n=1}^{N}a_{sn}\left(b_1,\ldots,b_{2N_c-1}\right)\cos\left(2\pi nt\right), \quad (4)$$

$$w_{cN}(t,b_1,\ldots,b_{2N_c-1})= \frac{1}{K}\left\{\cos(\pi\cdot t)+\sum_{n=1}^{N}a_{cn}\left(b_1,\ldots,b_{2N_c-1}\right)\cos\left[\pi(2n+1)t\right]\right\}. \quad (5)$$

где $N=(2N_c-1)(M+1)$

Преобразование Фурье ВФ (4) и (5) представим, соответственно, выражениями

$$S_{sN}\left(x,b_1,\ldots,b_{2N_c-1}\right)= \frac{\sin(\pi x)}{\pi x}\left[1+\sum_{n=1}^{N}a_{sn}\left(b_1,\ldots,b_{2N_c-1}\right)\cos\left(n\pi\right)\frac{x^2}{x^2-n^2}\right], \quad (6)$$

$$S_{cN}(x,b_1,\ldots,b_{2N_c-1})= \frac{-\cos(\pi x)}{\pi K}\left[\frac{0{,}5}{x^2-0{,}25}+\sum_{n=1}^{N}a_{cn}(b_1,\ldots,b_{2N_c-1})\cos(n\pi)\frac{n+0{,}5}{x^2-(n+0{,}5)^2}\right]. \quad (7)$$

СП (6) и (7) представлены бесконечно дифференцируемым произведением сомножителей, первые из которых, соответственно, $\sin(\pi x)/\pi x$ или $\cos(\pi x)/\pi K$. Так как ни эти сомножители, ни их производные не равны нулю на произвольной частоте x, для выполнения условия равенства нулю уравнений системы (3) должны быть равны нулю вторые сомножители СП

$$S_{sN}\left(x,b_1,\ldots b_{2N_c-1}\right) \text{ и } S_{cN}(x,b_1,\ldots b_{2N_c-1})$$

вместе со своими производными. Тогда система уравнений (3) для ВФ (4) и (5) преобразуется, соответственно, к системам уравнений

$$\begin{cases} \sum_{n=1}^{N}a_{sn}\left(b_1,\ldots,b_{2N_c-1}\right)\cos(n\pi)\frac{b_1^2}{b_1^2-n^2}=-1, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \sum_{n=1}^{N}a_{sn}\left(b_1,\ldots,b_{2N_c-1}\right)\cos(n\pi)\frac{b_{2N_c-1}^2}{b_{2N_c-1}^2-n^2}=-1, \\ \sum_{n=1}^{N}a_{sn}\left(b_1,\ldots,b_{2N_c-1}\right)\cos(n\pi)\frac{d}{dx}\left(\frac{x^2}{x^2-n^2}\right)\Big|_{x=b_1}=0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \sum_{n=1}^{N}a_{sn}\left(b_1,\ldots,b_{2N_c-1}\right)\cos(n\pi)\frac{d}{dx}\left(\frac{x^2}{x^2-n^2}\right)\Big|_{x=b_{2N_c-1}}=0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \sum_{n=1}^{N}a_{sn}\left(b_1,\ldots,b_{2N_c-1}\right)\cos(n\pi)\frac{d^{(M)}}{dx^{(M)}}\left(\frac{x^2}{x^2-n^2}\right)\Big|_{x=b_1}=0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \sum_{n=1}^{N}a_{sn}\left(b_1,\ldots,b_{2N_c-1}\right)\cos(n\pi)\frac{d^{(M)}}{dx^{(M)}}\left(\frac{x^2}{x^2-n^2}\right)\Big|_{x=b_{2N_c-1}}=0. \end{cases} \quad (8)$$

$$\begin{cases} \sum_{n=1}^{N}a_{cn}\left(b_1,\ldots,b_{2N_c-1}\right)\cos(n\pi)\frac{(n-0{,}5)^2}{b_1^2-(n-0{,}5)^2}\Big|=-1, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \sum_{n=1}^{N}a_{cn}\left(b_1,\ldots,b_{2N_c-1}\right)\cos(n\pi)\frac{(n-0{,}5)^2}{b_{2N_c-1}^2-(n-0{,}5)^2}=-1, \\ \sum_{n=1}^{N}a_{cn}\left(b_1,\ldots,b_{2N_c-1}\right)\cos(n\pi)\frac{d}{dx}\left[\frac{(n-0{,}5)^2}{x^2-(n-0{,}5)^2}\right]\Big|_{x=b_1}=0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \sum_{n=1}^{N}a_{cn}\left(b_1,\ldots,b_{2N_c-1}\right)\cos(n\pi)\frac{d}{dx}\left[\frac{(n-0{,}5)^2}{x^2-(n-0{,}5)^2}\right]\Big|_{x=b_{2N_c-1}}=0, \\ \sum_{n=1}^{N}a_{cn}\left(b_1,\ldots,b_{2N_c-1}\right)\cos(n\pi)\frac{d^{(M)}}{dx^{(M)}}\left[\frac{(n-0{,}5)^2}{x^2-(n-0{,}5)^2}\right]\Big|_{x=b_1}=0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \sum_{n=1}^{N}a_{cn}\left(b_1,\ldots,b_{2N_c-1}\right)\cos(n\pi)\frac{d^{(M)}}{dx^{(M)}}\left[\frac{(n-0{,}5)^2}{x^2-(n-0{,}5)^2}\right]\Big|_{x=b_{2N_c-1}}=0. \end{cases} \quad (9)$$

При $N \geq 2$ решениями систем (8), (9) относительно коэффициентов, соответственно, $a_{sn}\left(b_1,...,b_{2N_c-1}\right)$ и $a_{cn}(b_1,...,b_{2N_c-1})$ являются:

$$a_{sn}\left(b_1,...,b_{2N_c-1}\right)=(-1)^{n+N}\cdot\prod_{\substack{k=1\\k\neq n}}^{N}\frac{k^2}{n^2-k^2}\cdot\left[\prod_{i=1}^{2N_c-1}\left(1-\frac{n^2}{b_i^2}\right)\right]^{M+1}, \quad (10)$$

$$a_{cn}(b_1,...,b_{2N_c-1})=$$

$$\frac{(-1)^{n+N}}{2n+1}\prod_{\substack{k=1\\k\neq n}}^{N}\frac{k^2+k}{(n^2+n)-(k^2+k)}\cdot\left[\prod_{i=1}^{2N_c-1}\left(\frac{b_i^2-(n+0{,}5)^2}{b_i^2-0{,}25}\right)\right]^{M+1}. \quad (11)$$

Для $N=1$ $a_{s1}(b)=\left(1-\frac{1^2}{b^2}\right)$ и $a_{c1}(b)=\frac{1}{3}\left(\frac{b^2-2{,}25}{b^2-0{,}25}\right)$.

В задачах одновременного оценивания частоты и амплитуды сигнала выражение для нормирующего множителя определяется из условия равенства единице нормированной СП ВФ на нулевой частоте.

Спектральные свойства АВФ

Отметим некоторые свойства АВФ, которые в наибольшей степени влияют на возможность снижения методической погрешности и свойства, использование которых позволяет получать ВФ с оптимальными параметрами.

1. Положение нулевых значений (нулей) СП АВФ определяется положением нулей двух сомножителей. Положение нулей первого сомножителя $\sin(\pi x)$ или $\cos(\pi x)$, соответственно, для СП (6) или (7) периодическое и не влияет на положение нулей второго сомножителя, которым определяется огибающая боковых лепестков СП (6) или (7).

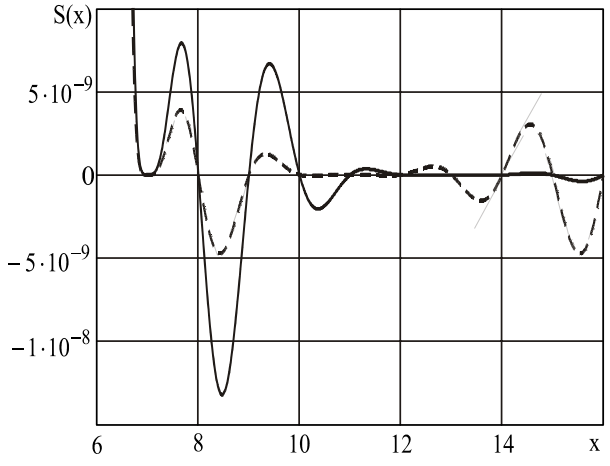


Рис. 1 – СП АВФ в области ближних боковых лепестков

Независимость положений нулей СП (6), (7) АВФ, определяемых, соответственно, сомножителем $\sin(\pi x)$ или $\cos(\pi x)$ от положения задаваемых нулей b_1, b_2,..., b_{2N_c-1}, позволяет для краткости называть эти положения нулей стационарными, а положения нулей вторых сомножителей – варьируемыми.

В частном случае системы уравнений (3) при $M=0$ для АВФ с N варьируемыми параметрами получаются решения, которые позволяют выполнить формирование нулевых значений СП на N задаваемых частотах b_1, b_2,..., b_N. Совмещение на одной частоте m нулей СП приводит к образованию на этой частоте нулей $m-1$ производных. В зависимости от решаемой задачи это свойство позволяет применять не равный порядок нулей на различных задаваемых частотах и использовать при этом более компактное выражение для коэффициентов $a_{sn}\left(b_1,...b_N\right)$ и $a_{cn}(b_1,...b_N)$

$$a_{sn}\left(b_1,...b_N\right)=(-1)^{n+N}\cdot\prod_{\substack{k=1\\k\neq n}}^{N}\frac{k^2}{n^2-k^2}\cdot\prod_{i=1}^{N}\left(1-\frac{n^2}{b_i^2}\right), \quad (12)$$

$$a_{cn}(b_1,...b_N)=$$

$$\frac{(-1)^{n+N}}{2n+1}\prod_{\substack{k=1\\k\neq n}}^{N}\frac{k^2+k}{(n^2+n)-(k^2+k)}\prod_{i=1}^{N}\left(\frac{b_i^2-(n+0{,}5)^2}{b_i^2-0{,}25}\right), \quad (13)$$

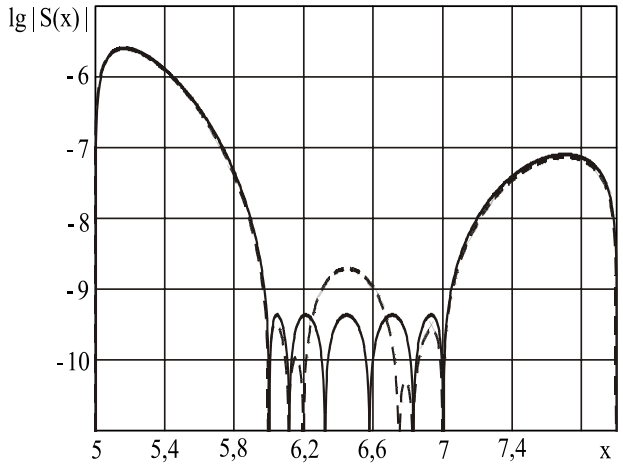


Рис. 2 – Область боковых лепестков СП АВФ

Возможность задания нулей СП на любых частотах иллюстрируется рис. 1 с СП АВФ (4) при $N=6$ и заданными нулями третьего порядка на двух относительных частотах: 7 и 13 для сплошной кривой, 7 и 11 для пунктирной кривой.

2. Уровень боковых лепестков зависит от распределения по частоте сформированных нулевых значений СП и их кратности. Увеличение интервала между частотами соседних нулей, определяющего ширину бокового лепестка в этом интервале, приводит к возрастанию УБЛ в этом интервале и влияет на уровень соседних боковых лепестков. Это свойство иллюстрируется рис. 2 для СП АВФ (4) с $N=4$, на котором для сплошной кривой, кроме нулей, на частотах $N+k$, где $k=1;2;....$, определяемых первым сомножителем $\sin(\pi x)$, заданы нули на частотах $b_1=6{,}1148135$, $b_2=6{,}319898$, $b_3=6{,}575851$, $b_4=6{,}828999$ и получен минимальный УБЛ в интервале частот [6, 7], а для пунктирной кривой нули второго сомножителя заданы на частотах $b_1=6{,}1148135$, $b_2=6{,}2$, $b_3=6{,}75$, $b_4=6{,}828999$. В результате, интервал частот между центральными нулями увеличен и уровень бокового лепестка в этом интервале также увеличен. Но при этом ширина соседних боковых лепестков и их уровень существенно уменьшились. Постепенно затухающее влияние изменения ширины бокового лепестка распространяется на соседние боковые лепестки (в интервалах от 5 до 6 и от 7 до 8).

3. Асимптотическая скорость уменьшения УБЛ СП (6) или (7) ($C_S = 20\lg\lim_{x\to\infty}|S(2x)/S(x)|$ в децибелах на октаву) составляет, соответственно, $(6+12N_\infty)$ дБ/окт и $(12+12N_\infty)$ дБ/окт, где N_∞ - число нулей, заданных на бесконечной частоте или на частоте $b >> N$.

Методика расчёта ВФ на основе АВФ с минимальным УБЛ при заданных ширине основного лепестка и скорости уменьшения УБЛ

Выбор параметров ВФ особенно важен при спектральном анализе для измерения параметров отдельных тонов в сигнале, содержащем несколько гармоник. Для того, чтобы динамический диапазон измеряемых сигналов был максимален, ПФ применяемой ВФ должно иметь узкий главный лепесток и низкий УБЛ [1]. Применение только такого критерия к выбору ВФ неизбежно приводит к использованию ВФ ДЧ. Однако из-за когерентного суммирования боковых лепестков его СП, имеющих постоянный уровень, она не подтверждает своих высоких характеристик при измерении параметров нескольких сигналов различных частот. Поэтому кроме минимального УБЛ СП при заданной ширине основного лепестка важным параметром ВФ является скорость уменьшения УБЛ СП.

Однозначная связь формы СП АВФ с положением задаваемых нулей позволяет на основе АВФ получать ВФ с оптимальными параметрами. Считаем, что задаются: ширина ΔF_0 основного лепестка СП на нулевом уровне, скорость C_S уменьшения УБЛ и количество N слагаемых ВФ. Необходимо определить положения варьируемых нулей АВФ таким образом, чтобы УБЛ при этом был минимален.

Учитывая, что при $b_1 < N$ один из нулей определяет ширину основного лепестка СП и его положение неизменно, а число N_∞ определяется заданной C_S, число варьируемых положений нулей АВФ будет составлять $N_{вар} = N - (1 + N_\infty)$. Тогда для минимизации УБЛ, целевую функцию можно записать в виде:

$$\max S_N\left(x, b_2, b_3, \ldots b_{N_{вар}}, N\right) \Rightarrow \min_{b_i} \qquad (14)$$

При решении этой задачи в качестве эталона по ширине основного лепестка и УБЛ можно использовать СП ВФ ДЧ, для которой предельные значения соотношений ширины основного лепестка и УБЛ определены фундаментальным свойством полиномов Чебышева.

Для поиска оптимальных параметров могут быть использованы стандартные программы, например, программа многомерной оптимизации fminsearch в системе программирования МАТЛАВ.

В таблицах 1, 2 приведены ВФ с оптимизированными параметрами большинство из которых не имеют аналогов. При оптимизации задавался ряд значений исходных параметров - ΔF_0 = 2,8; 3,2; 3,6; 4; 5; 6; 7; 8; 9, скорость уменьшения УБЛ (6 дБ/окт для СП АВФ (4) и 12 дБ/окт для СП АВФ (5)) и от двух до четырёх варьируемых параметров АВФ. Отметим две закономерности в свойствах СП оптимальных ВФ. Во-первых, УБЛ СП ВФ снижается при увеличении числа слагаемых ВФ и асимптотически стремится к УБЛ СП ВФ ДЧ. Во-вторых, ширина основного лепестка СП на любом уровне, отличном от нуля, уменьшается с ростом числа слагаемых ВФ и асимптотически стремится к ширине основного лепестка СП ВФ ДЧ. При увеличении C_S обе закономерности проявляются сильнее.

Таблица 1.

ВФ $w_{SN}(*)$, $C_S = 6$ дБ/окт, $N_\infty = 0$

N	Частоты b_i	УБЛ, дБ	ΔF_6
2	1,4; 2,0632	-27,09	1,486
4	1,4; 2,0444;2,94; 3,9281	-27,637	1,4794
2	1,6; 2,2083	-32,9795	1,6058
4	1,6; 2,1788; 3,0232; 3,9739	-33,4899	1,6021
2	1,8; 2,4041	-38,2672	1,7165
4	1,8; 2,3216; 3,1144; 4,0246	-39,387	1,7123
2	2; 2,6491	-43,1875	1,8188
4	2; 2,4758; 3,2257; 4,1180	-45,1383	1,8153
2	2,25; 3,0493	-49,421	1,9368
4	2,25; 2,6825; 3,3926; 4,3057	-52,0279	1,9371
2	2,5; 3,7226	-56,6824	2,0518
4	2,5; 2,8995; 3,5797; 4,5672	-58,6678	2,0521
3	3; 3,3524; 5,3290	-71,4828	2,2638
3	3,5; 4,0150; 6,0405	-85,2222	2,4257
4	4; 4,2803; 5,0893; 8,7183	-98,1739	2,6243
4	4,5; 4,858; 6,093; 10,661	-110,903	2,7799

Таблица 2.

ВФ $w_{CN}(*)$, $C_C = 12$ дБ/окт, $N_\infty = 0$

N	Частоты b_i	УБЛ, дБ	ΔF_6
3	1,4; 2,1577; 3,1993	-24,1852	1,5265
4	1,4; 2,1212; 3,1065; 4,2018	-25,1503	1,5138
2	1,6; 2,3355	-28,3002	1,6731
4	1,6; 2,2518; 3,1872; 4,2475	-30,6118	1,6249
2	1,8; 2,4514;	-34,1185	1,7843
4	1,8; 2,3911; 3,2759; 4,2984	-36,1187	1,7591
2	2; 2,2573	-40,2450	1,8816
4	2; 2,5377; 3,3720; 4,3540	-41,657	1,8647
2	2,25; 2,7549	-47,9658	1,9905
4	2,25; 2,7292; 3,5013; 4,4294	-48,7074	1,9845
2	2,5; 3,0309	-54,3019	2,0959
3	2,5; 2,9301; 3,6503	-55,8631	2,0936
4	2,5; 2,9281; 3,6396; 4,5107	-55,9594	2,0929
2	3; 3,8862	-67,8151	2,2880
3	3; 3,3765; 4,1783	-69,1854	2,2941
4	3; 3,3540; 3,9742; 4,8074	-70,2846	2,2912
4	3,5; 3,8116; 4,3845; 5,4297	-83,5472	2,4769
3	4; 4,4703; 6,3214	-95,412	2,6368
4	4; 4,2709; 4,7753; 6,6762	-97,3896	2,6480

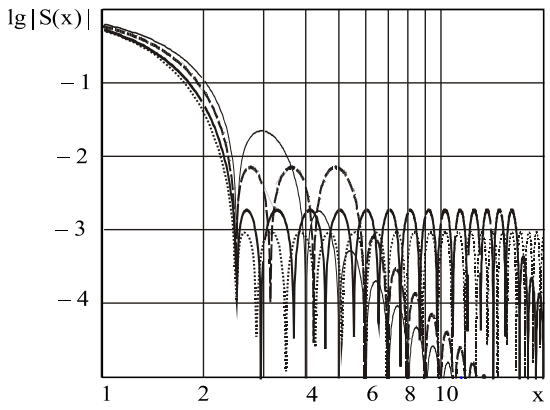


Рис. 3 – СП ВФ Дольфа-Чебышева и СП АВФ

Эти закономерности на рис.3 иллюстрируются СП оптимальных ВФ, полученных из АВФ (4) с заданными: нормированной шириной основного лепестка по нулевому уровню 5 бин, скоростью уменьшения УБЛ $C_S = 30$ дБ/окт и $N = 3$ (тонкая сплошная кривая), $N = 5$ (пунктир), $N = 16$ (жирная сплошная кривая). ВФ для $N = 3$, $N = 5$ и $N = 16$ описываются параметрами w_{S3}(t; 2,5; ∞ ; ∞), w_{S5} = (t; 2,5; 3,1604; 4,2408; ∞ ; ∞) и w_{S16}(t; 2,5; 2,944937; 3,677263; 4,561547; 5,526585; 6,53828; 7,579867; 8,642864; 9,723159; 10,819388; 11,932493; 13,066362; 14,230791; 15,454889; ∞ ; ∞) соответственно. Видно снижение УБЛ СП ВФ и уменьшение ширины основного лепестка при увеличении N. Пределом являются УБЛ и ширина основного лепестка СП ВФ ДЧ (точечная кривая). Но существенные недостатки ВФ ДЧ – наличие δ-функций на краях временного интервала и, соответственно, незатухающий УБЛ СП в приведённых ВФ исключены.

Характерный вид СП оптимальных ВФ по приведённому критерию – постоянный УБЛ в интервале нормированных частот $x \le N$ и уменьшающийся с заданной скоростью C_s УБЛ при $x > N$. Однако оптимизация распределения нулей СП АВФ возможна не только по минимуму УБЛ, но и по минимуму отклонения от заданного закона изменения УБЛ.

Сравнение спектральных свойств АВФ со свойствами лучших известных ВФ [1, 2, 3, 8] показывает, что АВФ с оптимальными параметрами имеют существенно меньший УБЛ. Разница доходит до 14 дБ. В качестве примера на рис. 4 приведены логарифмы модулей СП лучшей из известных ВФ для $\Delta F_0 = 4{,}744$ и $C_S = 6$ дБ/окт [8] (сплошная линия) и оптимальной ВФ $w_{C3}(t;2{,}372;2{,}829;3{,}578)$ (пунктир).

Приведённые в таблицах частоты задаваемых нулей и получаемые параметры СП ВФ округлены до 4-х знаков. В общем случае такие округления увеличивают минимально достижимый УБЛ не более чем на 0,01 дБ.

Коэффициенты (12), (13) по заданным в таблицах частотам могут быть пересчитаны в числовые.

Отметим, что большинство традиционных ВФ являются частными случаями полученных АВФ при определённых фиксированных значениях параметров последних. Это утверждение легко иллюстрируется на ряде широко используемых традиционных ВФ. Подстановка нулевых значений b_i СП любой известной ВФ, представленной тригонометрическим рядом в выражение $w_S(t, b_1, ...b_2, N)$, приводит к совпадению СП на любой частоте и совпадениям известной ВФ с АВФ.

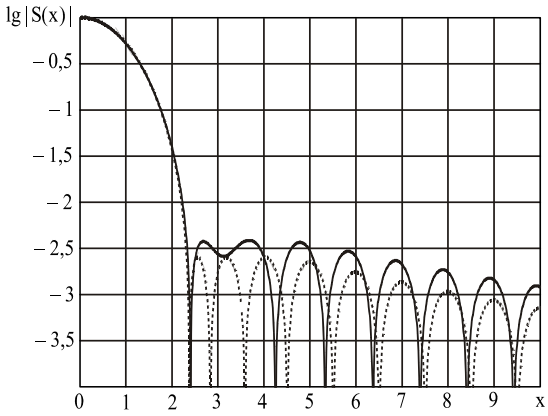


Рис.4 – СП алгебраической ВФ [8] и СП АВФ

И наоборот, использование коэффициентов при соответствующих слагаемых ряда позволяет получить значения частот b_i. Например, для популярной ВФ Блэкмана [1] (вида $w_{S2}(t, b_1, b_2)$) с коэффициентами $a_1 = 0{,}5/0{,}42$ и $a_2 = 0{,}08/0{,}42$ из (4) и (12) нетрудно получить, что $b_1 = \sqrt{28/3}$ и $b_2 = \infty$. Весовую функцию Наттолла

$$w(t) = 0{,}363581 + 0{,}4891775\cos 2\pi t + \\ +0{,}1365995\cos 4\pi t + 0{,}0106411\cos 6\pi t$$

можно получить, задав b_1=4,28200081517, b_2=5,082904286136, b_3=8,73773653982.

Выводы

Результаты исследования показывают, что по основным параметрам, определяющим погрешность оценки частоты и амплитуды сигнала, АВФ и ВФ с оптимизированными параметрами, полученные на основе АВФ, имеют преимущество перед известными ВФ.

АВФ и ВФ с оптимизированными параметрами могут использоваться не только в спектральном анализе, но и в других прикладных областях: в задачах синтеза различных измерительных устройств, интерполяции сигналов и других приложениях, для которых представленная работа может стать отправным пунктом.

Литература

1. Хэррис Дж. // ТИИР – 1978. – Т. 66. – №1. – С. 60.
2. Дворкович А.В. // Цифровая обработка сигналов. – 2001. – № 2. – С. 49.
3. Кравченко В.Ф. Лекции по теории атомарных функций и некоторым их приложениям. - М.: Радиотехника, 2003.
4. Иванов Ю.Е. // Проблемы управления и информатики. – 1998. – № 2. – С. 102.
5. Давыдочкин В.М., Езерский В.В. // Цифровая обработка сигналов. –2005. – № 3. – С. 22.
6. Тихонов В.И., Харисов В.Н. Статистический анализ и синтез радиотехнических устройств и систем. - М.: Радио и связь, 1991.
7. Гоноровский И.С. Радиотехнические цепи и сигналы. - М.: Радио и связь, 1986.
8. Кириллов С.Н., Соколов М.Ю., Стукалов Д.Н. // Радиотехника. –1996. – № 6. – С. 36.

УДК 681.518

АДАПТИВНАЯ ФИЛЬТРАЦИЯ ИЗОБРАЖЕНИЙ СО СТРУКТУРНЫМИ ИСКАЖЕНИЯМИ

Костров Б.В., Саблина В.А.

Рассмотрена проблема восстановления аэрокосмических изображений со структурными искажениями. Исследовано применение метода адаптивной фильтрации в целях решения поставленной задачи. Предложен способ моделирования изучаемого шума. Разработана методика выбора коэффициента адаптации фильтра. Интерпретированы результаты проведенных экспериментов.

Введение

Процесс регистрации аэрокосмических изображений сопровождается внесением искажений, обусловленных различными факторами, в том числе несовершенством самой формирующей оптико-электронной системы. Например, вследствие неравномерности чувствительности приемных элементов на изображениях появляются структурные искажения в виде характерной «полосатости» [1].

Устранение таких структурных искажений является разновидностью задачи восстановления изображения, которая формулируется как задача воссоздания исходного изображения $f(x,y)$ по наблюдаемому искаженному изображению $g(x,y)$. В результате получается восстановленное изображение $\hat{f}(x,y)$ [2].

Статья посвящена решению задачи устранения смоделированного аддитивного шума исследуемого типа, накладываемого на аэрокосмические изображения земной поверхности, методом адаптивной фильтрации, описана разработанная методика подбора коэффициента адаптации фильтра. В целях снижения требований к вычислительным ресурсам для перехода в спектральную область и обратно использованы алгоритмы быстрого преобразования Уолша (БПУ) [3-5].

Математическая постановка задачи

В настоящей статье рассматриваются искажения, вызванные наличием аддитивного шума:

$$g(x,y)=f(x,y)+n(x,y). \qquad (1)$$

где $g(x,y)$ – искаженное изображение,
$f(x,y)$ – исходное изображение,
$n(x,y)$ – аддитивный шум.

Модель процессов искажения и восстановления изображения путем частотной фильтрации представлена на рис. 1.

При восстановлении в частотной области осуществляется разложение искаженного изображения $g(x,y)$ в спектр $G(u,v)$ с последующей фильтрацией. При этом получается спектр восстановленного изображения $\hat{F}(u,v)$:

$$\hat{F}(u,v)=G(u,v)H(u,v), \qquad (2)$$

где $H(u,v)$ – передаточная функция фильтра, $G(u,v)$ – спектр искаженного изображения.

Затем производится возврат в пространственную область, в результате чего возникает восстановленное изображение $\hat{f}(x,y)$, которое является приближением к неискаженному изображению $f(x,y)$.

В рассмотренной модели особую важность имеет проблема оценки эффективности произведенной фильтрации. В реальных условиях шум обычно носит случайный характер, что мешает абсолютно точному воссозданию исходного изображения, поэтому необходимо решить задачу минимизации отличий восстановленного изображения от исходного.

Количественно для оценки степени отличия изображения от эталонного часто используются такие метрики, как среднеквадратическое отклонение (СКО) и пиковое отношение сигнал/шум (ПОСШ). СКО задается следующим соотношением:

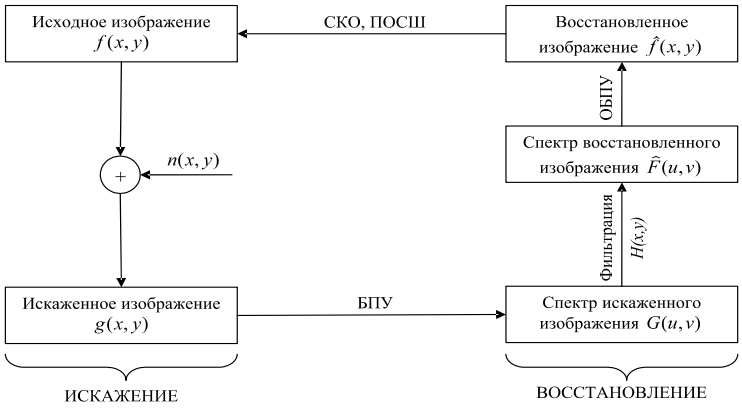


Рис. 1 – Модель процессов искажения/восстановления изображения

$$\sigma = \sqrt{\frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}(\hat{f}(x,y)-f(x,y))^2}\ , \qquad (3)$$

где $M \times N$ – размер рассматриваемого изображения, σ – среднеквадратическое отклонение.

Полученное значение СКО используется для расчета ПОСШ:

$$PSNR = 20\log_{10}\frac{255}{\sigma}, \qquad (4)$$

где $PSNR$ - пиковое отношение сигнал/шум.

Следует отметить, что метрики СКО и ПОСШ не всегда хорошо определяют визуально воспринимаемое качество изображения. В последнее время предпринимается большое количество попыток разработать новые метрики объективной оценки качества изображений, учитывающие свойства зрительной системы человека, что само по себе является довольно сложной задачей. Поэтому на рассматриваемом этапе исследований ограничимся использованием критерия минимизации СКО в целях адаптации создаваемого фильтра.

Для получения спектра искаженного изображения представляется целесообразным использовать прямое дискретное двумерное преобразование Уолша, осуществляемое следующим образом [6]:

$$G(u,v)=\frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1}g(x,y)wal_{(u,v)}(x/M,y/N), \qquad (5)$$

где $wal_{(u,v)}(x/M, y/N)$ - двумерная функция Уолша.

Соответственно, возврат в пространственную область после фильтрации выполняется с помощью обратного дискретного двумерного преобразования Уолша:

$$\hat{f}(x,y)=\sum_{u=0}^{M-1}\sum_{v=0}^{N-1}\hat{F}(u,v)wal_{(u,v)}(x/M,y/N). \qquad (6)$$

Из пары преобразований (5) и (6) видно, что для прямого и обратного преобразования Уолша используются одни и те же базисные функции с точностью до нормировочного коэффициента $1/MN$. Таким образом, матрицы прямого и обратного преобразования Уолша совпадают, что удобно для программной реализации алгоритма. Расчет спектра Уолша требует только многократного выполнения операций сложения и вычитания, что также удобно для программной реализации алгоритма.

В свою очередь, быстрое преобразование Уолша (БПУ), описанное в научной литературе [7,8], позволяет понизить сложность алгоритма с N^4 операций сложения, необходимых для вычисления обычного двумерного дискретного преобразования Уолша (ДПУ) до всего лишь $2N^2\log_2 N$ операций сложения для изображений размерами $N \times N$, где N является степенью двойки. Аналогично вычисляется и обратное быстрое преобразование Уолша (ОБПУ).

Однако использование преобразования Уолша имеет не только преимущества, но и некоторые недостатки. Трудности проведения фильтрации в спектре Уолша вытекают из отсутствия соотношений, позволяющих производить эффективное вычисление циклической свертки с переходом в спектральное пространство [9]. В случае преобразования Уолша, в отличие от преобразования Фурье, известны соотношения для вычисления диадической свертки. Тем не менее указанные трудности не препятствуют осуществлению описанных в настоящей статье исследований.

В соответствии с вышесказанным основной задачей, возникающей при восстановлении изображения с помощью фильтрации, является нахождение такой передаточной функции фильтра $H(u,v)$, которая обеспечивает наилучшее приближение восстановленного изображения к исходному, что при выбранном критерии подбора фильтра эквивалентно минимизации значения среднеквадратического отклонения σ.

Модель шума

На изображении, сформированном при помощи линейки ПЗС-элементов, часто возникают искажения, связанные с неравномерностью чувствительности приемников излучения. Визуально такие искажения представляются в виде полосовых помех на изображении.

В статье рассматривается модель аддитивного шума $n(x,y)$, искажающего одинаково каждый столбец изображения:

$$n(x,y)=n(x,0). \qquad (7)$$

В простейшем случае можно рассмотреть неслучайный шум, представляющий собой попеременное чередование полос, осветляющих либо затемняющих исходное изображение.

Такой шум математически описывается следующим образом:

$$n(x,y)=\begin{cases} A \text{ при } (2l-2)s \le x < (2l-1)s, \\ -A \text{ при } (2l-1)s \le x < 2ls, \end{cases} \qquad (8)$$

где A – амплитуда шума, l – натуральное число, s – ширина одной полосы в пикселях.

Путем варьирования амплитуды шума A можно получить шум различного уровня.

Например, при $A=0$ шум вообще отсутствует. Чтобы приблизить модель к реальным типам шума, необходимо исследовать шум, имеющий случайный характер.

Такой шум задается выражением:

$$n(x,y)=\xi_l \text{ при } (l-1)s \le x < ls, \qquad (9)$$

где ξ_l – l-ая реализация случайной величины ξ, l – натуральное число, s – ширина одной полосы в пикселях.

Исследования проводились на примере аддитивного равномерного по строке шума, для которого функция плотности распределения вероятностей имеет вид:

$$p(\xi)=\begin{cases} \dfrac{1}{2A} \text{ при } -A \le \xi \le A, \\ 0 \text{ в остальных случаях.} \end{cases} \qquad (10)$$

Здесь значение A задает диапазон возможных амплитуд шума, поэтому в отличие от неслучайного шума при таком же значении A получается в среднем более низкий уровень шума.

Результаты экспериментальных исследований

Исследования проводились с помощью специально разработанного программного средства, позволяющего моделировать указанные типы шума, накладывать его на исходные изображения, рассчитывать БПУ, проводить фильтрацию в спектре Уолша, а также оценивать СКО и ПОСШ.

Так как изучаемые типы шума оказывают влияние только на спектральные составляющие, для которых $u=0$, то в качестве передаточной функции фильтра была выбрана следующая функция:

$$H(u,v)=\begin{cases}1/k(v) \text{ при } u=0,\\ 1 \text{ при } u\neq 0.\end{cases} \quad (11)$$

где $k(v)$ – функция адаптации фильтра.

Далее подробно описан выбор фильтра для случая $s=1$, так как модель шума с полосами шириной в один пиксель наиболее приближена к реальным типам шума, когда такой искаженный столбец изображения формируется одним и тем же элементом из линейки ПЗС-элементов [1].

Поскольку при ширине полос в один пиксель при случайном шуме для его спектральных составляющих не наблюдается зависимости от v, то можно рассматривать в качестве $k(v)$ просто некоторый коэффициент K:

$$k(v)=\begin{cases}1 \text{ при } v=0,\\ K \text{ при } v\neq 0,\end{cases} \quad (12)$$

где K - коэффициент адаптации фильтра.

При $v=0$ оставляем $k(v)=1$, то есть $H(0,0)=1$, чтобы сохранить среднюю яркость изображения неизменной.

Необходимо найти оптимальный коэффициент адаптации фильтра, обеспечивающий наилучшее качество результатов фильтрации.

Очевидно, что искомое значение коэффициента K будет зависеть от амплитуды шума A: чем сильнее шум, тем больше его необходимо подавлять.

Таким образом, задача исследования сводится к разработке методики выбора коэффициента $K=K_{opt}$ при заданном значении $A=const$, который минимизирует значение СКО:

$$\sigma(K,A)\rightarrow \min_{K}, \quad (13)$$

Аналитическое нахождение зависимости оптимального $K_{opt}(A)$ представляется затруднительным, проведенное экспериментальное исследование позволяет легко построить ее эмпирически.

Основные этапы экспериментов описаны на примере двух тестовых изображений, представляющих собой фрагменты изображений земной поверхности размерами 128×128 пикселей (рисунки 2 и 3) и имеющие 256 градаций яркости.



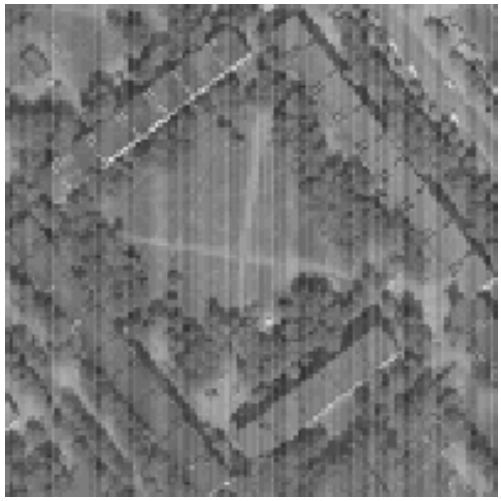
Рис. 2 – Тестовое изображение 1



Рис. 3 – Тестовое изображение 2

На тестовые изображения накладывался шум с различной амплитудой A. Рассматривались варианты случайного и неслучайного шума. Пример наложения случайного шума с $A=16$ приводится на рис. 4.

Для заданного значения A проводилась фильтрация в спектре Уолша с различным коэффициентом адаптации K, начиная от $K=1$, когда обрабатываемое изображение не изменяется. Логично предположить, что при слишком малых значениях K еще остается шум, а при слишком больших значениях K теряется информация о самом изображении. Это должно приводить к наличию точки минимума функции $\sigma(K)$, что в дальнейшем и подтвердилось экспериментально. Таким образом, при построении исследуемых зависимостей имеет смысл увеличивать значение K от 1 и до начала возрастания $\sigma(K)$.



Рису. 4 – Изображение 1, искаженное случайным шумом

В результате были получены таблицы $\sigma(K,A)$ зависимостей среднеквадратического отклонения от коэффициента адаптации фильтра K и амплитуд A от 0 до 39.

Для шума, который носит случайный характер, значения СКО были усреднены по 100 реализациям шума.

Фрагмент таблицы СКО для изображения 1 со случайным шумом приводится в таблице 1.

Таблица 1

Пример таблицы значений СКО

K\A	0	1	2	3	4	5	6	7
1	0	0,71	1,2205	1,7647	2,3343	2,9072	3,4962	4,0761
1,5	1,6979	1,8203	1,9485	2,1389	2,3495	2,6255	2,9204	3,2368
2	2,5587	2,63	2,6898	2,7621	2,86	3,0066	3,1458	3,3295
2,5	3,0835	3,1329	3,1512	3,22	3,2621	3,351	3,44	3,5377
3	3,4709	3,4763	3,4984	3,525	3,5571	3,6015	3,666	3,7321

По таким таблицам, считая $A = const$, можно построить семейства графиков зависимостей $\sigma(K, A)$, где $A = 0, 3, ..., 39$. По этим графикам можно оценить вид соответствующих зависимостей, например, для изображения 1 со случайным шумом такое семейство представлено на рис. 5.

Эмпирический характер построенных зависимостей, а также случайный характер шума обуславливают появление неровностей на графиках, особенно заметных для сильных шумов. Перед дальнейшим исследованием можно произвести аппроксимацию функций зависимостей $\sigma(K)$.

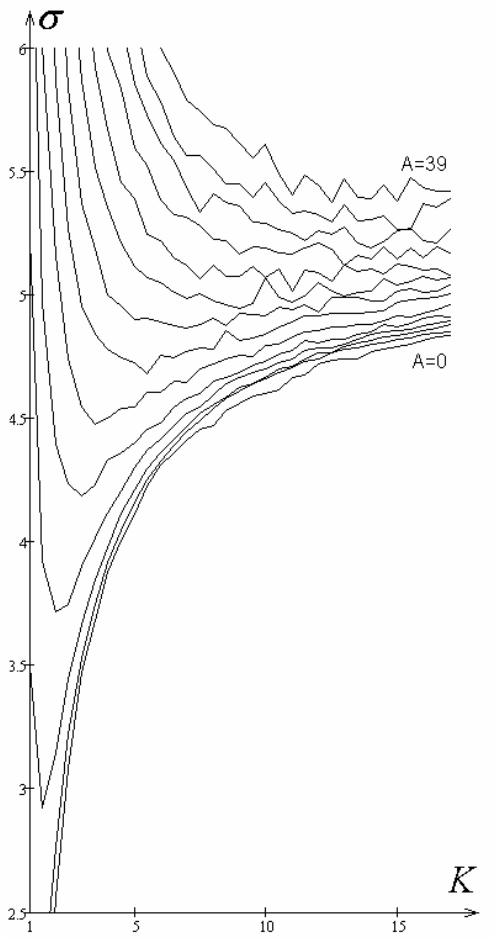


Рис. 5 – Семейство графиков зависимостей СКО от коэффициента адаптации фильтра K

Тогда можно заметить, что рассматриваемые функции, начиная с некоторого значения A, имеют точку экстремума, которая с увеличением A сдвигается в сторону увеличения K.

Это означает, что для совсем малых шумов не имеет смысла проводить фильтрацию, а для сильного шума следует выбирать большее значение коэффициента адаптации фильтра K.

Также можно построить графики зависимостей $\sigma(K_{opt}(A), A)$ минимальных СКО от A для оптимальных значений коэффициента адаптации фильтра $K = K_{opt}$.

На рис. 6 представлены такие графики для различных рассмотренных вариантов изображений и шумов.

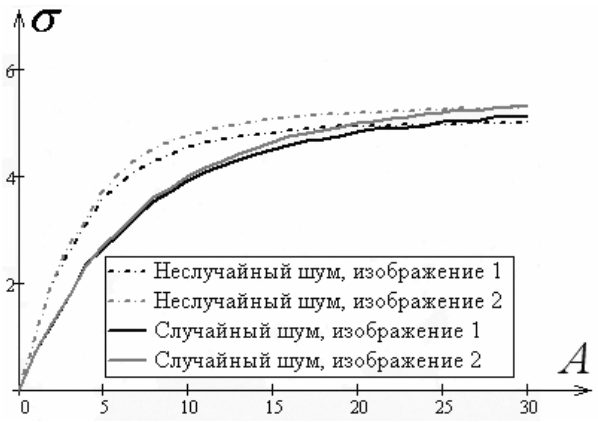


Рис. 6 – Графики зависимостей минимально достижимого значения СКО от амплитуды шума A

Видно, что при использовании описанной методики восстановления изображения, искаженного моделируемым шумом, СКО восстановленного изображения $\hat{f}(x, y)$ лежит в пределах 1,5÷2% при изменении амплитуды шума A от ±10 до ±30 градаций яркости.

Соответственно найденным значениям минимальных СКО для $A = const$ определяются и оптимальные значения коэффициента $K = K_{opt}$.

Таким образом, находятся решения задачи $\sigma(K, A) \to \min_K$ для каждого A.

Отсюда можно построить аппроксимированные графики зависимостей $K_{opt}(A)$, которые представлены на рис. 7.

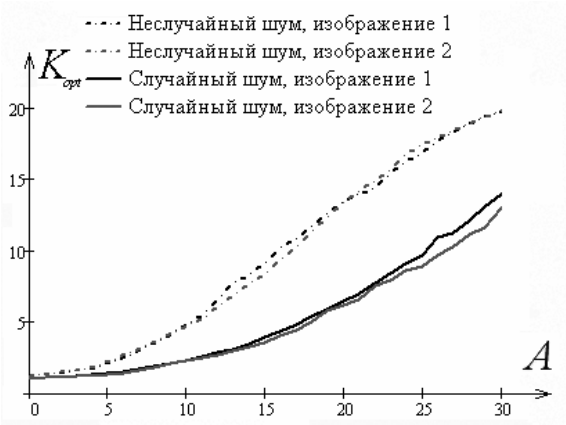


Рис. 7 – Аппроксимированные графики зависимостей оптимального значения коэффициента адаптации фильтра K_{opt} от амплитуды шума A

Пример изображения, восстановленного по предлагаемой методике, приводится на рис. 8.



Рис. 8 – Восстановленное изображение 1

Заключение

Предложена методика выбора фильтра для восстановления изображения со структурными искажениями. Условием применения данной методики является наличие априорной информации о характере шума, позволяющей отнести его к одному из рассмотренных типов.

Фильтрация осуществляется в спектре Уолша, при этом для построения фильтра, обеспечивающего наилучшее приближение восстановленного изображения к оригиналу, можно использовать зависимости, представленные на рис. 7. Для других типов шума нужно предварительно построить аналогичные зависимости, провести дополнительные исследования.

Описанная методика обладает таким преимуществом как высокая скорость вычислений, достижимая за счет использования алгоритма БПУ.

Литература

1. В.К. Злобин, В.В. Еремеев Обработка аэрокосмических изображений. – М.: ФИЗМАТЛИТ, 2006. – 288 с.
2. Р. Гонсалес, Р. Вудс Цифровая обработка изображений. - М.: «Техносфера», 2006. – 1072 с.
3. Б.В. Костров Основы цифровой передачи и кодирования информации. – М.: «ТехБук», 2007. –192 с.
4. Л.А. Залманзон Преобразования Фурье, Уолша, Хаара и их применение в управлении, связи и других областях. – М.: «Наука», 1989. – 496 с.
5. Проектирование специализированных информационно-вычислительных систем. Под ред. Ю.М. Смирнова. – М: Высшая школа, 1984. – 359 с.
6. В.А. Саблина, Б.В. Костров О возможностях использования преобразования Уолша для спектрального анализа изображений //Проблемы передачи и обработки информации в сетях и системах телекоммуникаций: Материалы 15-й Международной научно-технической конференции. Часть 2. - Рязань: Рязанский государственный радиотехнический университет. 2008. С. 129-132.
7. H.F. Harmuth Sequency Theory. Foundations and Applications. - New York, San Francisco, London: Academic Press, Inc., 1977. – 526 p.
8. Н. Ахмед, К.Р. Рао Ортогональные преобразования при обработке цифровых сигналов. – М: Связь, 1980. – 248 с.
9. Э. Прэтт Цифровая обработка изображений. – М.: Мир, 1982. – Кн. 1. – 312 с.

НОВЫЕ КНИГИ

Цифровая обработка сигналов. Практическое руководство для инженеров и научных работников (+CD) / Стивен Смит; пер. с англ. А. Ю. Линовича, С.В. Витязева., И.С. Гусинского. — М. : Додэка-XXI, 2008. — 720 с. : ил. — (Серия «Схемотехника»).
ISBN 978-5-94120-145-7
Заказ: www.dokabooks.ru

В книге изложены основы теории цифровой обработки сигналов. Акцент сделан на доступности изложения материала и объяснении методов и алгоритмов так, как они понимаются при практическом использовании. Цель книги — практический подход к цифровой обработке сигналов, позволяющий преодолеть барьер сложной математики и абстрактной теории, характерных для традиционных учебников. Изложение материала сопровождается большим количеством примеров, иллюстраций и текстов программ (которые вы также можете найти на прилагаемом CD). Книга предназначена научным работникам и инженерам, желающим применять методы цифровой обработки в различных технических сферах. Рекомендуется аспирантам и студентам, изучающим цифровую обработку сигналов.

Методы автоматического обнаружения и сопровождения объектов. Обработка изображений и управление / Б.А. Алпатов, П.В. Бабаян, О.Е. Балашов, А.И. Степашкин – М.: Радиотехника, 2008. – 176 с.

Освещены вопросы пространственно-временной обработки и анализа последовательностей изображений применительно к задачам обнаружения и сопровождения объектов. Значительное внимание уделено алгоритмам управления приводами поворотных устройств систем сопровождения объектов. Книга предназначена для научных и инженерных работников, аспирантов, студентов старших курсов, интересующихся задачами, связанными с разработкой систем обнаружения и сопровождения объектов по данным видеонаблюдений. www.radiotec.ru

Логическое проектирование цифровых систем на основе программируемых логических интегральных схем / В.В. Соловьев, А.М. Климович – М.: Горячая линия – Телеком, 2008. – 376 с.

Книга посвящена проблемам логического проектирования отдельных цифровых устройств на основе программируемых логических интегральных схем (ПЛИС). Предлагаются методы синтеза комбинационных схем, конечных и микропрограммных автоматов, позволяющие эффективно использовать архитектурные особенности ПЛИС, а также учитывать системные требования. Представлены новые модели конечных автоматов, позволяющие значительно снизить стоимость реализации и повысить быстродействие проектируемых устройств. Изложение материала сопровождается большим числом примеров. Предназначена для разработчиков цифровых систем, может быть использована в качестве учебного пособия для аспирантов и студентов. www.techbook.ru/solovjev.html

УДК 621.396.96

СПОСОБ ПОДАВЛЕНИЯ ДИСКРЕТНЫХ КОРРЕЛИРОВАННЫХ ПОМЕХ АДАПТИВНЫМ РЕШЕТЧАТЫМ ФИЛЬТРОМ

Бартенев В.Г.

Предложено построение адаптивного цифрового решетчатого фильтра, работающего в условиях воздействия дискретных коррелированных помех. Рассмотрен пример использования предложенного устройства.

Применение решетчатых фильтров для подавления мешающих отражений известно давно[1,2]. Адаптивный решетчатый фильтр включает в себя формирование коэффициентов отражения и использование их в качестве весовых коэффициентов для обеления коррелированной помехи, минимизируя ошибки предсказания вперед и назад. Многозвенный решетчатый фильтр позволяет осуществлять эффективную фильтрацию многокомпонентных помех с многомодовыми спектрами. Хотя способ решетчатой фильтрации характеризуется высокой эффективностью и возможностью режекции помех с многомодовыми спектрами, его использование применительно к дискретным помехам наталкивается на большие трудности. Если коррелированная помеха дискретная и занимает один дискрет дальности, то при наличии полезного сигнала в этом же дискрете дальности сформированные коэффициенты отражения будут учитывать и параметры полезного сигнала, поэтому при фильтрации он будет режектироваться так же, как и помеха. Таким образом, классический режекторный фильтр может работать только по классифицированной выборке наблюдений. Главный вопрос не в том, как формировать коэффициенты отражения и каким для этого алгоритмом следует воспользоваться, а в том как исключить влияние полезного сигнала на формируемые коэффициенты отражения. Для протяженной помехи, занимающей много элементов разрешения по дальности, когда оценки коэффициентов отражения формируются с усреднением по элементам дальности, протяженность коррелированной помехи значительно превышает протяженность полезного сигнала, что дает основание рассматривать входную выборку наблюдений классифицированной и влиянием полезного сигнала на коэффициенты отражения можно пренебречь. Совсем другое дело, когда помеха дискретна и когда для оценки коэффициентов отражения производится усреднение только по обрабатываемой пачке импульсов, а формируемые коэффициенты отражения учитывают в том числе и полезный сигнал.

С целью исключения подавления полезного сигнала при фильтрации неклассифицированной выборки наблюдений предлагается способ, который включает в себя формирование коэффициентов отражения решетчатого фильтра по двум выборкам наблюдения, отличающимся частотой повторения или несущей частотой. При этом коэффициенты отражения, сформированные по одной выборке наблюдений, используют в качестве весовых коэффициентов для фильтрации другой выборки наблюдений и наоборот, а результаты фильтрации одной и другой выборок наблюдения объединяют на выходе. Предлагаемый способ отличается тем, что разнос частот повторения или несущих частот двух выборок наблюдения выбирают так, чтобы разность соответствующих им доплеровских смещений частоты была бы значительно меньше доплеровских смещений частоты для каждой из выборок наблюдений. Предлагается использовать существенные различия в доплеровской скорости наблюдаемых объектов: медленно движущихся источников дискретных помех и быстро движущихся полезных целей. Разницу в несущих частотах или частоте повторения предлагается выбирать так, чтобы коэффициенты отражения для медленно движущейся помехи практически совпадали, а вот для полезного сигнала существенно различались. Поэтому, если для фильтрации сигналов одной выборки наблюдений использовать коэффициенты отражения другой выборки наблюдений, помеха будет компенсироваться, а полезный сигнал нет. Блок схема предлагаемого решетчатого фильтра с перекрестным вводом коэффициентов отражения приведена на рис.1 (где ЛЗ цифровая линия задержки на период повторения, К-блок вычисления коэффициентов отражения, Х-умножители, ∑- сумматоры,D-детекторы).

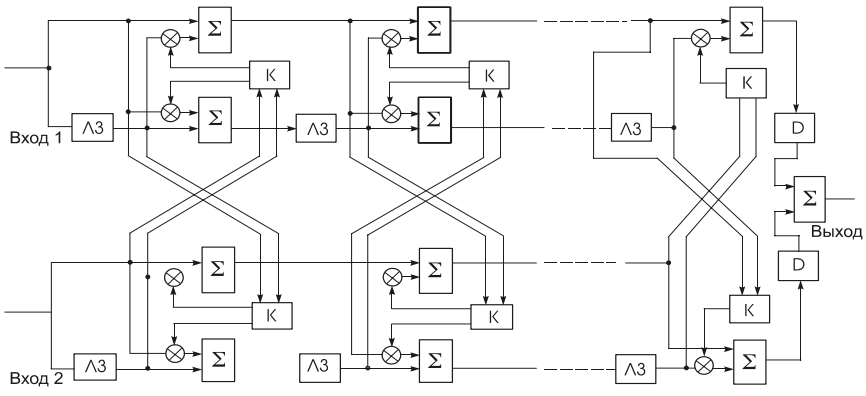


Рис.1 Адаптивный решетчатый фильтр, работающий по неклассифицированной выборке наблюдений.

Проиллюстрируем работу предлагаемого способа на конкретном примере, прибегнув к моделированию устройства с помощью системы MATLAB [3]. Выполним фильтрацию двух выборок наблюдений на разных несущих частотах 500 МГц и 550МГц с постоянным периодом повторения 0,001 с, задавшись параметрами полезного сигнала и двух компонентной помехи для каждой выборки наблюдений:

```
as=1;%амплитуда полезного сигнала
fs1=400;%доплер полезного сигнала, Гц для первой выборки наблюдений
am=1;%амплитуда 1 компоненты помехи
fm1=10;%доплер 1 компоненты помехи, Гц для первой выборки наблюдений
l1=300/500;%длина волны для первой выборки наблюдений с несущей 500МГц
l2=300/550;%длина волны для второй выборки наблюдений с несущей 550МГц
fc1=20;%доплер 2 компоненты помехи, Гц для первой выборки наблюдений
ac=2;%амплитуда 2 компоненты помехи
vm=fm1*l1/2; fm2=2*vm/l2; %доплер 1 компоненты помехи, Гц для второй выборки %наблюдений
vc=fc1*l1/2; fc2=2*vc/l2; %доплер 2 компоненты помехи, Гц для второй выборки %наблюдений
vs=fs1*l1/2;% fs2=2*vs/l2;% доплер полезного сигнала, Гц для второй выборки %наблюдений
%ФОРМИРОВАНИЕ ВЫБОРОК НАБЛЮДЕНИЙ
t = 0:0.001:0.256;%Время наблюдения 256 мсекунд с тактом 1мсек
% входной сигнал для первой выборки наблюдений
z1 = ac*exp(2*fc1*pi*t*sqrt(-1))+am*exp(2*fm1*pi*t*sqrt(-1))+as*exp(2*fs1*pi*t*sqrt(-1))+.1*randn(1,length(t));
% входной сигнал для второй выборки наблюдений
z2 = 2*exp(2*fc2*pi*t*sqrt(-1))+am*exp(2*fm2*pi*t*sqrt(-1))+as*exp(2*fs2*pi*t*sqrt(-1))+.1*randn(1,length(t));
figure(2)%Спектральное представление входных сигналов первого и второго частотных каналов
Z1 = fft(z1,256);% для 1канала
f = 256*(0:127)/256;
plot(f,10*log10(abs(Z1(1:128))),'K')
hold on
Z2 = fft(z2,256);% для 2 канала
f = 256*(0:127)/256;
plot(f,10*log10(abs(Z2(1:128))),'K')
    %ФОРМИРОВАНИЕ КОЭФФИЦИЕНТОВ ОТРАЖЕНИЯ
    for i = 2:p+1,%p-порядок авторегрессии
      ep2 = ef2(i:N);% ef2-ошибка предсказания вперед второго канала
      em2 = eb2(i-1:N-1);
    % K2 – коэффициент отражения во втором частотном канале
    K2(i-1) = 2 * ep2' * em2 / (ep2'*ep2 + em2'*em2);
      a2 = [a2;0] - K2(i-1) * [0;flipud(a2)];
       for j = N:-1:i,
        ef2_old = ef2(j);
        ef2(j) = ef2(j) - K2(i-1) * eb2(j-1);
    eb2(j) = eb2(j-1) - K2(i-1) * ef2_old;
      end
       E2(i) = (1 - K2(i-1)'*K2(i-1)) * E2(i-1);
    End
```

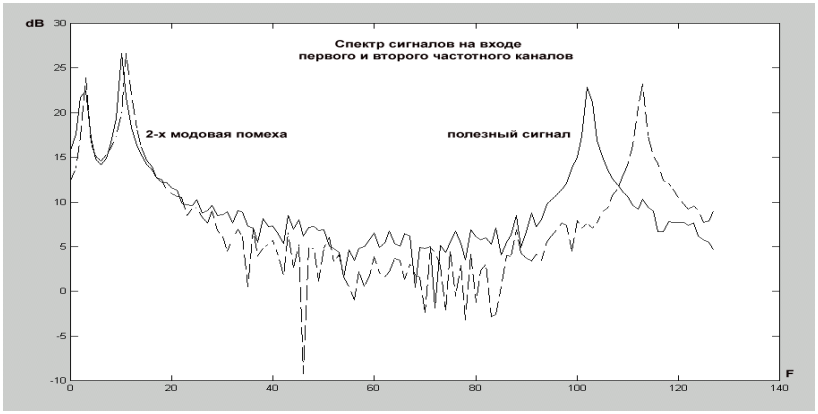


Рис.2 Спектральное представление входных сигналов первого и второго частотных каналов

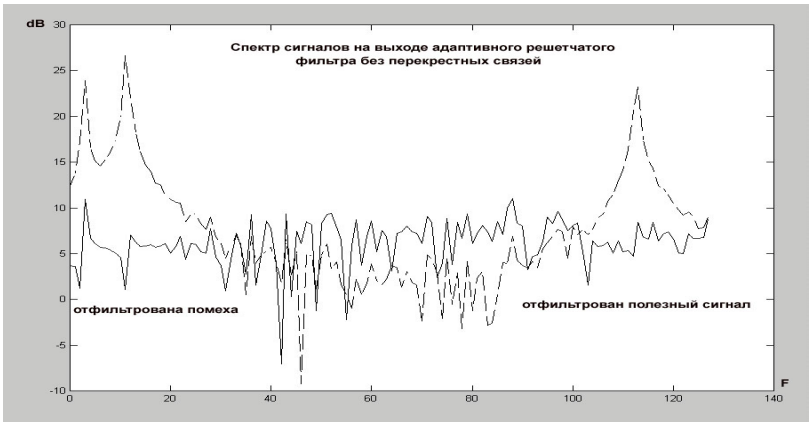


Рис.3 Спектральное представление сигналов на выходе решетчатого фильтра без перекрестных связей

На рис. 2 представлен спектр сигналов на входе решетчатого фильтра для сигналов на разных несущих частотах: спектр двухмодовой медленно движущейся коррелированной помехи практически совмещен для двух несущих частот. В то же время, спектр сигнала для быстро движущегося объекта имеет существенные различия для разных несущих частот.

figure(3) % Спектральное представление сигналов на выходе решетчатого фильтра без перекрестных связей

```
%plot(real(ef2(1:128)),'R')
EZ2 = fft(ef2',256);% для 2 канала
fe = 256*(0:127)/256;
plot(fe,10*log10(abs(EZ2(1:128))),['K','-'])
hold on
   plot(f2,10*log10(abs(Z2(1:128))),['K','--'])
```

На рис. 3 представлен спектр сигналов на выходе решетчатого фильтра без перекрестных связей. Отметим, что при этом скомпенсированы как помеха, так и полезный сигнал. Показан только один частотный канал, в другом канале полезный сигнал также подавлен.

figure(4) % Спектральное представление сигналов на выходе решетчатого фильтра с %перекрестными связями

```
EZ = fft(ef',256);% для 2канала
fe = 256*(0:127)/256;
plot(fe,10*log10(abs(EZ(1:128))),['K','-'])
hold on
plot(f2,10*log10(abs(Z1(1:128))),['K','--'])
```

На рис. 4 представлен спектр сигналов на выходе решетчатого фильтра с перекрестными связями. При этом, как видно из рисунка, помеха скомпенсирована, а полезный сигнал сохранен.

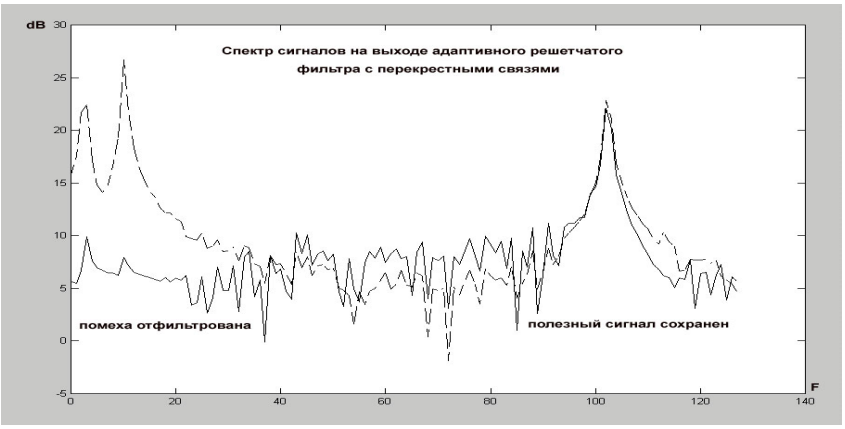


Рис.4 Спектральное представление сигналов на выходе решетчатого фильтра с перекрестными связями

```
%ПЕРЕКРЕСТНОЕ ИСПОЛЬЗОВАНИЕ
КОЭФФИЦИЕНТОВ ОТРАЖЕНИЯ
for i = 2:p+1,
  ep = ef(i:N); ef-ошибка предсказания вперед первого канала
  em = eb(i-1:N-1);
  K(i-1) = 2 * ep' * em / (ep'*ep + em'*em);
  a = [a;0] - K(i-1) * [0;flipud(a)];
% ПЕРЕКРЕСТНЫЙ ВВОД КОЭФФИЦИЕНТА
ОТРАЖЕНИЯ ВТОРОГО КАНАЛА К2
  % ПРИ ФОРМИРОВАНИИ ОШИБКИ
ПРЕДСКАЗАНИЯ ef ПЕРВОГО КАНАЛА
 for j = N:-1:i,
    ef_old = ef(j);
    ef(j)  = ef(j)   - K2(i-1) * eb(j-1);
              eb(j)  = eb(j-1) - K2(i-1) * ef_old;
end
  E(i) = (1 - K(i-1)'*K(i-1)) * E(i-1);
End
```

Аналогичным образом можно показать работу предложенного способа при использовании вобуляции периода повторения импульсов от пачки к пачке. Главный вывод - проведенное исследование в системе MATLAB полностью подтверждает положительный эффект от применения предложенного способа фильтрации дискретных помех по неклассифицированной выборке наблюдений.

Литература

1. Haykin S., Currie B., Kesler S. “Maximum Entropy spectral analysis of radar clutter”, Proc. IEEE Trans.vol.70, September 1982, pp.953-962.
2. Barbarossa B., Picardi G. “Predictive Adaptive Moving Target Indicator”, Signal Processing, №10, 1986, pp.83-97.
3. Потемкин В.Г. “Справочник по MATLAB” Анализ и обработка данных. http://matlab.exponenta.ru/ml/book2/chapter8/

ДИСКОВЫЕ ПОДСИСТЕМЫ: ДОСТИЖЕНИЕ МАКСИМАЛЬНОЙ СКОРОСТИ ПРИ НАИМЕНЬШЕМ КОЛИЧЕСТВЕ ДИСКОВ

Аминев Д.А.

Приводится обзор различных дисковых подсистем с наименьшим количеством дисков, устанавливаемых в ПЭВМ контроллеров. Рассматривается способ их тестирования. Выявляется максимальная скорость записи на дисковые подсистемы.

Введение

Во многих приложениях, в том числе при работе с видео- и фотоизображениями большего объема, возникает необходимость записи потока данных со скоростями выше 500 Мбит/с на носители информации емкостью более 500 Гбайт. При этом предъявляются требования к количеству дисков не более двух. Таким образом, в общем варианте задача сводится к достижению максимальной скорости записи потока данных при наименьшем количестве жестких дисков.

Наиболее скоростными из распространенных зарекомендовали себя накопители на жестком магнитном диске (далее НЖМД) ST3750640AS типа Barracuda фирмы Seagate. Внешний вид диска представлен на рис. 1, а его характеристики описаны в таблице 1.

Далее представлены результаты тестирований как одного данного типа НЖМД, так и массива типа RAID-0, составленного из нескольких НЖМД. Выбор массива НЖМД типа RAID-0 обусловлен достижением максимальной скорости записи, так как запись происходит параллельно на все НЖМД массива.

Тестирование производилось с помощью специальной программы, которая осуществляет запись данных в файл на дисковой подсистеме посредством функции посекторной записи. Эта функция дает максимальный выигрыш в скорости по сравнению со стандартными функциями операционной системы ПЭВМ. Специальная программа отображает график зависимости скорости записи от времени и объема записанной на дисковую подсистему информации.



Рис. 1. НЖМД ST3750640AS

Таблица 1 – характеристики НЖМД ST3750640AS

Характеристика	Описание
Интерфейс	SATA 2.0
Размер жесткого диска, Гб	750
Скорость передачи данных, МБит/с	300
Объем буфера, МБ	16
Скорость вращения жесткого диска, об/мин	7200
Среднее время доступа, мс	12
Количество головок	8
Количество пластин	4
Среднее время наработки на отказ (MTBF), тыс. часов	500
Уровень шума, дБ	30
Потребление, Вт	13
Вес, кг	0.72
Размеры, Ш х Д х В	102 мм х 147 мм х 26 мм

Таблица 2 – характеристики контроллера Intel®ICH8R

Характеристика	Описание
Уровни RAID	RAID 0, 1, 5, 10
Количество подключаемых устройств	До шести накопителей с интерфейсом SATA
Скорость передачи данных, МБит/с	3000

Результаты тестирований

Тестирование одного НЖМД

Тестирование НЖМД осуществлялось через контроллер Intel®ICH8R, который встроен в “южный мост” системной платы GA-965P-DS4 фирмы GIGABYTE. Характеристики контроллера описаны в таблице 2, а на рис. 2 представлен график процесса записи на один НЖМД.

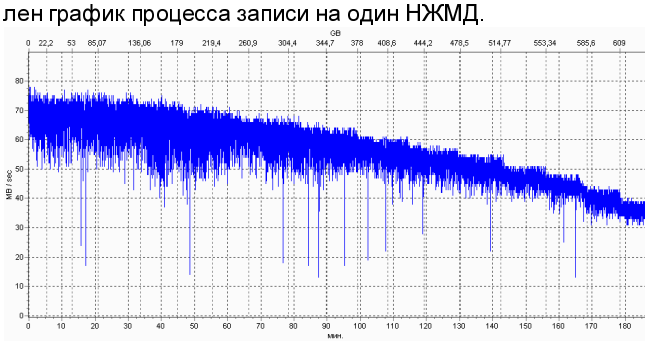


Рис. 2. Запись на 1 НЖМД.

Как видно из рис. 2, средняя скорость записи на 1 НЖМД изменяется от 70 до 35 Мбайт/с, что в целом не достигает 500 Мбит/с (62,5 Мбайт/с).

Тестирование двух, трех и четырех НЖМД в массиве RAID-0

При объединении двух, трех и четырех НЖМД в массив RAID-0 посредством контроллера Intel®ICH8R, встроенного в “южный мост” системной платы GA-965P-DS4 фирмы GIGABYTE, были получены следующие результаты (рис. 3, 4, 5).

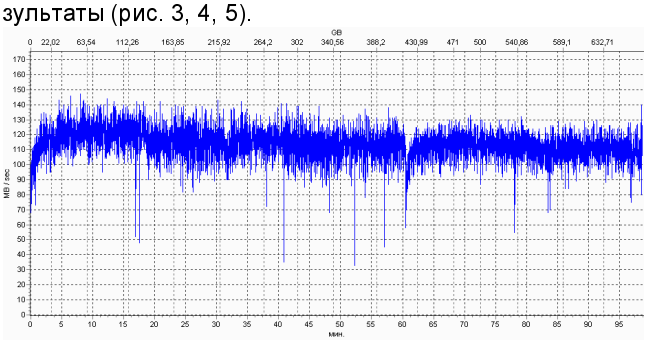


Рис. 3. Запись на 2 НЖМД в массиве RAID-0.

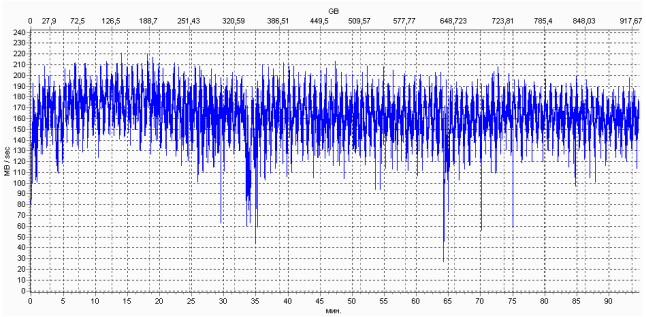


Рис. 4. Запись на 3 НЖМД в массиве RAID-0.

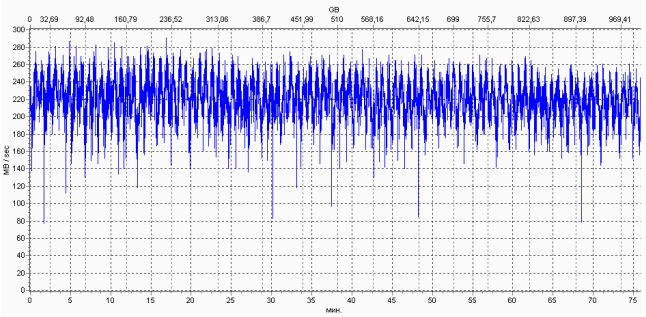


Рис. 5. Запись на 4 НЖМД в массиве RAID-0.

По графикам процесса записи можно сделать вывод, что скорость записи значительно превышает 500 Мбит/с и увеличивается пропорционально количеству НЖМД массиве RAID-0.

Особенностью такой дисковой подсистемы является ее непосредственное расположение в корпусе компьютера.

Тестирование двух НЖМД в массиве RAID-0 на дисковой подсистеме Thecus

Используя внешнюю дисковую подсистему Thecus N2050 с двумя установленными НЖМД в массиве RAID-0, подключенную к компьютеру посредством контроллера 1225SA фирмы Adaptec, можно достичь результатов, представленных на рис. 6. Характеристики контроллера и дисковой подсистемы описаны в таблицах 3 и 4 соответственно.

Таблица 3 – характеристики дисковой подсистемы Thecus N2050.

Характеристика	Описание
Внешний интерфейс	eSATA
Уровни RAID	RAID 0, 1
Количество устанавливаемых НЖМД	До двух накопителей с интерфейсом SATA
Форм фактор НЖМД	3,5”
Скорость передачи данных, МБит/с	3000
Размеры, Ш х Д х В	160 мм х 200 мм х 85 мм

Таблица 4 – характеристики контроллера 1225SA фирмы Adaptec.

Характеристика	Описание
Тип интерфейса системной шины	PCI Express x1
Уровни RAID	RAID 0, 1 и JBOD (отдельный диск)
Количество подключаемых устройств	До двух накопителей с интерфейсом eSATA
Тип адресации	48-битная логическая адресация (LBA), поддержка дисков большого объема
Скорость передачи данных, МБит/с	3000
Размеры, В х Д	64 мм х 85 мм

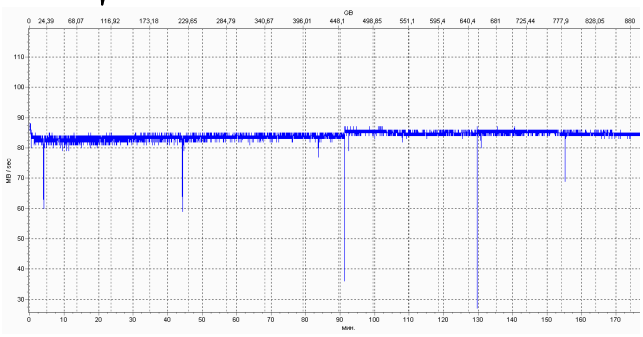


Рис. 8. Запись на дисковую подсистему Thecus N2050.

По графику на рис. 8 видно, что, несмотря на среднюю скорость записи 83 Мбайт/с (664 Мбит/с), что меньше по скорости, чем в массиве RAID-0 на контроллере Intel®ICH8R, диапазон разброса скорости значительно меньше.

Заключение

Таким образом, проведен анализ скоростей записи на различных дисковых подсистемах, и наиболее подходящим вариантом для реализации требований по скорости записи выше 500 Мбит/с на носители информации емкостью более 500 Гбайт является массив RAID-0 составленный из двух НЖМД ST3750640AS на:

1) контроллере Intel®ICH8R, который встроен в системную плату GA-965P-DS4;

2) внешней дисковой подсистеме Thecus N2050, подключенной к компьютеру посредством контроллера 1225SA фирмы Adaptec.

Литература

1. Barracuda 7200.10 Serial ATA Product Manual. – Seagate Technology LLC. 2007, 68 с.
2. GA-965P-DS4 User's Manual. – G.B.T. Technology. 2006, 104 с.
3. Adaptec SATA RAID 1225SA. – Adaptec, Inc. 2007, 2с.
4. www.adaptec.com/ru-RU/products/sata_tech/Adaptec - Adaptec RAID 1225SA.htm – описание контроллера 1225SA фирмы Adaptec
5. Quick Installation Guide N2050. – Thecus Technology Corp. 2007, 12 с.
6. ru.wikipedia.org/wiki/RAID – описание уровней RAID
7. ru.wikipedia.org/wiki/JBOD – описание массива JBOD

ИНФОРМАЦИОННОЕ ПИСЬМО

Уважаемый коллега!
Приглашаем Вас принять участие в работе 11-ой Международной научно-технической конференции
«Цифровая обработка сигналов и ее применение – DSPA'2009», которая состоится в Москве 24-26 марта 2009 года

ОРГАНИЗАТОРЫ:

- Российское научно-техническое общество радиотехники, электроники и связи им. А.С. Попова
- IEEE Signal Processing Society
- Российская секция IEEE
- Институт радиотехники и электроники РАН
- Институт проблем управления РАН
- Институт проблем передачи информации РАН
- Московский научно-исследовательский телевизионный институт
- Компания AUTEX Ltd. (АВТЭКС)

ОСНОВНЫЕ НАПРАВЛЕНИЯ РАБОТЫ КОНФЕРЕНЦИИ:

- Теория сигналов и систем
- Теория и методы ЦОС
- Цифровая обработка многомерных сигналов
- Цифровая обработка речевых и звуковых сигналов
- Цифровая обработка изображений
- ЦОС в системах телекоммуникаций
- ЦОС в радиотехнических системах
- ЦОС в системах управления и робототехники
- Цифровая обработка измерительной информации
- Нейрокомпьютерная обработка сигналов и изображений
- Цифровое телевидение и телевещание
- Цифровое радиовещание
- ЦОС в системах защиты информации
- Проектирование и техническая реализация систем ЦОС
- ЦОС в открытых системах
- Проблемы подготовки специалистов в области ЦОС

НОВОСТИ УНИВЕРСИТЕТСКОЙ ПРОГРАММЫ TEXAS INSTRUMENTS INC.

Университетская программа организована фирмой Texas Instruments Inc. (TI) – мировым лидером в области технологий цифровой обработки сигналов, аналоговых устройств и полупроводниковой техники – с целью оказания образовательным учреждениям поддержки в создании учебных и научных лабораторий цифровой обработки сигналов и информационных технологий реального времени. В рамках университетской программы TI предоставляет оборудование на льготных условиях, проводит обучение и технические консультации преподавателей и ученых, поощряет работу студентов. Подробнее об университетской программе TI можно узнать на сайте www.dspa.ru/cosvuz или из статей, опубликованных в журналах «Цифровая обработка сигналов» № 4 за 2007 год и № 1 за 2008 год. Данная статья является кратким обзором новостей 2008 года, касающихся университетской программы.

Фирма TI объявляет о начале функционирования в сети Internet русскоязычного сайта, посвященного университетской программе TI. Сайт доступен в рамках общего русскоязычного ресурса TI: www.ti.com/ru. Появление данного раздела сайта можно считать знаменательным в том смысле, что оно утверждает намерения фирмы укреплять свои позиции на российском рынке и в целом в Восточной Европе. Сайт будет содержать регистрационную форму с возможностью заполнения на русском языке. TI ожидает, что это позволит большему числу преподавателей и ученых подключиться к программе и по достоинству оценить предоставляемые ею возможности. Надеемся, что и раздел сайта журнала «ЦОС» www.dspa.ru/cosvuz, который пока еще находится в разработке, сможет существовать параллельно сайту TI и освещать события, связанные с развитием ЦОС-лабораторий в российских ВУЗах на более близком для российских специалистов уровне.

Другой новостью университетской программы TI следует назвать появление обучающего диска (teaching ROM) под названием “From Matlab and Simulink to real-time with TI DSP” (От Matlab и Simulink к обработке сигналов в реальном времени на ЦСП фирмы TI). Как следует из названия, диск представляет собой учебник по использованию системы Matlab для быстрой и эффективной разработки программных решений, реализуемых на сигнальных процессорах фирмы TI. Привлекательность данного инструмента понятна, но чтобы научиться им правильно пользоваться, необходимо хорошее руководство. Таким руководством и должен стать предлагаемый в рамках университетской программы TI Обучающий диск. По словам авторов, материалы диска организованы так, что упор в обучении сделан не на реализацию отдельных типовых алгоритмов ЦОС, таких как БПФ и КИХ-фильтр, а на процесс разработки программного обеспечения для заданного приложения. Вы начинаете с разработки и моделирования алгоритма и постепенно доводите проект до реализации и оценки эффективности на аппаратной платформе. Примеры приложений: аудио- и видеосистемы, проводная и беспроводная связь, системы управления.

Другим новым продуктом является программное обеспечение для оценки эффективности и выбора аналогово-цифровых преобразователей Analog-to-Digital Converter Evaluation Software – ADCPro™. ADCPro может использоваться автономно для анализа набора данных, сформированных при тестировании АЦП, или совместно с аппаратными модулями оценки АЦП от фирмы TI. В последнем случае вы получаете мощную систему оценки АЦП. На оценочный модуль устанавливается одна из сменных карт оцифровки, и ADCPro позволяет анализировать эффективность выбранного преобразователя. Данный продукт может быть очень полезен в учебном процессе. Он позволяет исследовать и иллюстрировать различия в топологиях разных преобразователей и организации предварительной аналоговой обработки сигнала.

Еще одним полезным для обучения продуктом является демонстрационный набор eZ430 RF2480. Он построен на основе связки микроконтроллера MSP430F2274 и модуля беспроводной передачи данных про протоколу ZigBee CC2480 2.4 ГГц. Набор может послужить отличной стартовой аппаратно-программной платформой для обучения построению сетей на основе ZigBee.

Кроме перечисленных выше новых продуктов и услуг TI рекомендует также в рамках университетской программы использовать новые инструментальные средства и комплекты типовых программных кодеков для платформы DaVinci. Фирма IAR Systems разработала новейшие средства для разработки программного обеспечения встраиваемых систем обработки сигналов, поддерживающие языки ассемблера, Си и Си++. Средства ориентированы на использование с микроконтроллерами семейства MSP430. Фирмы Blackhawk и Spectrum Digital представляют новые эмуляторы для повышения эффективности работы с устройствами на базе контроллеров семейства TMS320C2000. Поддержка платформы DaVinci продолжает усиливаться. Могут быть бесплатно получены комплекты некоторых кодеков наиболее распространенных стандартов представления видео- и аудиоданных.

Университетская программа TI – это мощный инструмент для научных и образовательных коллективов, способный дать реальные возможности по освоению и практическому использованию самых современных информационных технологий реального времени. Рекомендуем ознакомиться с предложениями программы на сайте www.dspa.ru/cosvuz и поддержать планы редакции журнала «Цифровая обработка сигналов» по созданию сайта, объединяющего российские учебные и научные лаборатории цифровой обработки сигналов на основе обмена опытом и информацией.