

ПЕРЕДАЧА И ОБРАБОТКА ИНФОРМАЦИИ

УДК 681.3.06

С.Н. Кириллов, Д.С. Семин

ИМИТАЦИОННАЯ МОДЕЛЬ НАБЛЮДЕНИЙ ЗА КОСМИЧЕСКИМ ОБЪЕКТОМ В ВИДИМОМ И ИК-ДИАПАЗОНАХ ПРИ НАЛИЧИИ МЕШАЮЩИХ ФАКТОРОВ

Предложена имитационная модель наблюдений за космическими объектами в видимом и инфракрасном диапазонах при воздействии на него мешающих факторов. Представлены результаты имитационного моделирования полученной модели в виде траекторий движения космического объекта при наблюдении с наземной станции и фотоснимка наблюдаемого тела в диапазоне длин волн от 570 до 800 нм.

Ключевые слова: имитационное моделирование, модель возмущенного движения, модель проведения траекторных измерений, оптический диапазон, ИК-диапазон.

Введение. Проблема контроля космического пространства является весьма актуальной в связи с быстрым ростом населенности околоземной орбиты космическим мусором, связанным как с запуском новых космических аппаратов, так и с выводом из эксплуатации старых [1]. Для прогноза столкновения действующих космических аппаратов с другими телами необходим мониторинг и контроль околоземного космического пространства. Оптические наблюдения в видимом и инфракрасном (ИК) диапазонах позволяют собрать достаточно информации не только о параметрах движения, но и температуре, материале и характере космического объекта (КО). При этом важным является построение адекватных имитационных моделей наблюдения и движения КО в околоземном пространстве при наличии воздействия мешающих факторов, позволяющих прогнозировать траекторию движения.

Цель работы. Построение имитационной модели наблюдения за КО в видимом и ИК диапазонах, учитывающей различные особенности движения объекта по околоземной траектории, а также особенности формирования снимка объекта с использованием ПЗС матрицы.

Имитационная модель наблюдений. Структурная схема разработанной имитационной модели наблюдения за КО в видимом и ИК-диапазонах при наличии мешающих факторов

(рисунок 1) включает блоки модели движения КО, пересчёта координат, модели проведения измерений и 3D-модели КО.

В разработанной модели в качестве исходных данных используются начальные элементы Кеплеровской орбиты $\Omega_0, i_0, \omega_0, e_0, p_0, t_{I0}$ (долгота восходящего узла, наклонение орбиты, аргумент перигея, эксцентриситет, параметр орбиты, время прохождения перигея соответственно), информация о внешнем облике КО и оптических свойствах его поверхности (конфигурация КО), координаты пункта наблюдения (долгота и широта), параметры атмосферы, фотоприёмника и телескопа.

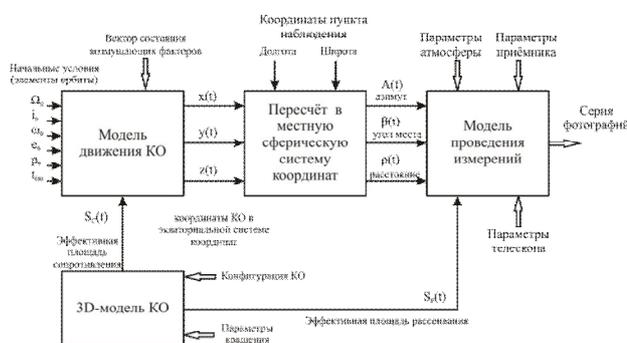


Рисунок 1 – Структурная схема имитационной модели наблюдений за КО

Результатом работы модели проведения

наблюдений является серия фотографий КО, получаемых со станции наблюдения из заданной точки земного шара.

Модель движения КО. Движение КО под действием возмущающих факторов описывается дифференциальным уравнением [2, 3]:

$$d^2 \mathbf{r}_{КО} / dt^2 = -(\mu / r_{КО}^2) \mathbf{r}^0 + \Delta \mathbf{g}, \quad (1)$$

где $\mathbf{r}_{КО}$ - радиус-вектор КО в геоцентрической экваториальной системе координат, μ - гравитационный параметр Земли, $r_{КО}$ - расстояние от начала координат до КО, $\mathbf{r}^0 = \mathbf{r}_{КО} / r_{КО}$ - единичный вектор, $\Delta \mathbf{g}$ - вектор ускорения, создаваемого действием возмущающих факторов.

Отклонение от Кеплеровской траектории движения КО может быть вызвано следующими возмущающими факторами:

- нецентральность поля тяготения Земли ($\Delta \mathbf{g}_Г$);
- сопротивление атмосферы ($\Delta \mathbf{g}_А$);
- давление света ($\Delta \mathbf{g}_С$);
- солнечный ветер ($\Delta \mathbf{g}_{СВ}$);
- гравитационные поля других небесных тел ($\Delta \mathbf{g}_{НТ}$);
- взаимодействия электромагнитного поля Земли ($\Delta \mathbf{g}_М$).

При этом суммарное возмущающее ускорение $\Delta \mathbf{g}$ определяется по принципу суперпозиции векторов [2, 3]:

$$\Delta \mathbf{g} = \Delta \mathbf{g}_А + \Delta \mathbf{g}_С + \Delta \mathbf{g}_{СВ} + \Delta \mathbf{g}_М + \Delta \mathbf{g}_{НТ} + \Delta \mathbf{g}_Г.$$

Описание возмущенного движения в модели производилось методом оскулирующих элементов орбиты [3]. Изменения элементов $\Omega, i, \omega, e, p, t_f$ под действием возмущающего ускорения $\Delta \mathbf{g}$ описываются дифференциальными уравнениями [3]:

$$\begin{aligned} d\Omega / dt &= \Delta g_{\omega} r_{КО} \sin u / (\sqrt{\mu p} \sin i), \\ di / dt &= \Delta g_{\omega} r_{КО} \cos u / \sqrt{\mu p}, \\ d\omega / dt &= e^{-1} \sqrt{p / \mu} \cdot (-\Delta g_r \cos \vartheta + \\ &+ \Delta g_n (1 + r_{КО} / p) \sin \vartheta - \\ &- \Delta g_{\omega} e \cdot (r_{КО} / p) \operatorname{ctg} i \sin u), \\ dp / dt &= 2 \Delta g_n r_{КО} \sqrt{p / \mu}, \\ de / dt &= \sqrt{p / \mu} \cdot (\Delta g_r \sin \vartheta + \\ &+ \Delta g_n e r_{КО} / p + \Delta g_n (1 + r_{КО} / p) \cos \vartheta), \end{aligned} \quad (2)$$

$$dt_{П} / dt = (r_{КО}^2 / (e\mu)) \cdot (\Delta g_r (eN_{П} \sin \vartheta - \cos \vartheta) \Delta g_n N_{П} r_{КО} / p),$$

где $N_f = (1 - e^2)^{-1} ((2 + e \cos \vartheta) \sin \vartheta - 3e \sqrt{\mu p} \times \times (t - t_{П}) / r_{КО}^2)$, $\Delta g_r, \Delta g_n, \Delta g_{\omega}$ - радиальная, тангенциальная и бинормальная составляющие возмущающего ускорения соответственно, ϑ - истинная аномалия, $u = \vartheta + \omega$ - аргумент широты.

Для выражений (2) в разработанной модели применяются численные методы интегрирования с шагом времени 10...60 секунд [4].

На рисунке 2 приведена структурная схема разработанной статистической модели движения КО под действием возмущающих факторов.

При моделировании возмущений, вызванных неравномерностью гравитационного поля Земли, учитывалась только вторая зональная гармоника (обладающая наибольшей амплитудой).

Влияние неравномерности рельефа поверхности моделируется воздействием случайных радиальной и меридиональной составляющих $\Delta g_r^Г$ и $\Delta g_{\varphi}^Г$, распределённых в силу центральной предельной теоремы Ляпунова по нормальному закону с среднеквадратическим отклонением (СКО) на уровне $1.67 \cdot 10^{-4}$ м/с².

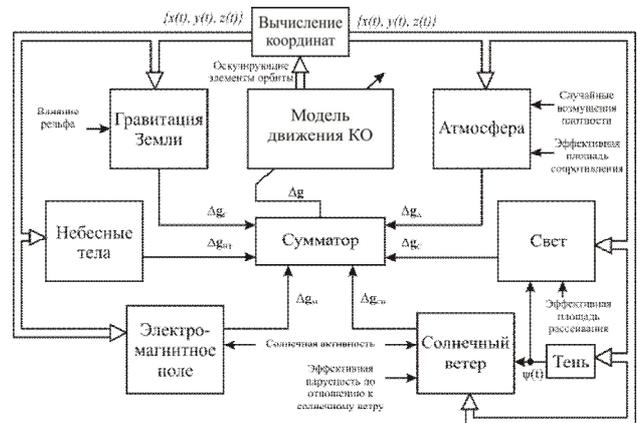


Рисунок 2 – Статистическая модель движения КО

Модель атмосферы описывает сопротивление движению КО верхних слоев, учитывающая изменение возмущающего воздействия из-за смены времени суток, полугодового эффекта, а также солнечной активности, при этом принимаются во внимание конструктивные особенности КО, полученные из 3D-модели. Стохастические флуктуации атмосферы представляются небольшим гауссовским процессом, СКО которого составляет 5 %, а время корреляции 0,05 от периода обращения КО вокруг Земли.

Действие солнечного ветра определяется эф-

фективной площадью освещенной поверхности, взаимным положением КО и Солнца, а также возможностью затенения возмущающего потока Землей с помощью введения теневой функции:

$$\psi(t) = \begin{cases} 1, & \text{КО вне поля тени от Земли;} \\ 0, & \text{КО в поле тени от Земли.} \end{cases}$$

Сила влияния солнечного ветра определяется случайным процессом со временем корреляции равной 27 суткам (период обращения Солнца вокруг своей оси). СКО процесса определяется 11-летним циклом Солнечной активности. Направление приложения силы определяется взаимным положением КО и Солнца.

Взаимодействие с электромагнитным полем Земли учитывается в модели аналогично гравитационному взаимодействию.

Составляющие слагаемого возмущающего ускорения, соответствующие воздействию гравитационного поля от постороннего небесного тела, определяются взаимным расположением Земли и рассматриваемого небесного тела, а также положением КО на околоземной орбите.

Величина слагаемых возмущающего ускорения неодинакова и в общем случае зависит от орбиты, геометрии и свойств поверхности КО. Поэтому помимо модели движения КО по орбите, разработана 3D-модель КО, описывающая движение объекта вокруг центра масс. Эти две модели считаются независимыми друг от друга.

На выходе модели движения КО получаются координаты в прямоугольной экваториальной

системе координат, которые далее в блоке пере-счёта переводятся в местную сферическую систему, ассоциированную с наблюдательным пунктом.

Модель проведения измерений при действии мешающих факторов. Наблюдения за КО в оптическом и ИК-диапазонах проводятся с помощью телескопов различных модификаций [6]. Регистрация получаемых изображений производится на ПЗС-матрицу фотоприёмника (для видимого диапазона размер матрицы 1040×1160, а для ИК-диапазона - 320×256). На основе вышеизложенного была создана имитационная модель проведения измерений в оптическом и ИК-диапазонах спектра, учитывающая следующие мешающие факторы:

- неточность системы слежения;
- вносимые турбулентной атмосферой искажения волнового фронта;
- дифракционные искажения, вызванные ограниченной апертурой телескопа;
- шум, обусловленный пространственной дискретностью ПЗС-матрицы;
- фотонный и темновой шум ПЗС-матрицы;
- шум квантования АЦП.

Структурная схема модели проведения измерений приведена на рисунке 3. Для различных диапазонов (видимого и ИК) меняются характеристики фазового шума, относительный размер апертуры, также изменяются параметры ПЗС-матрицы - количество пикселей и геометрические размеры.

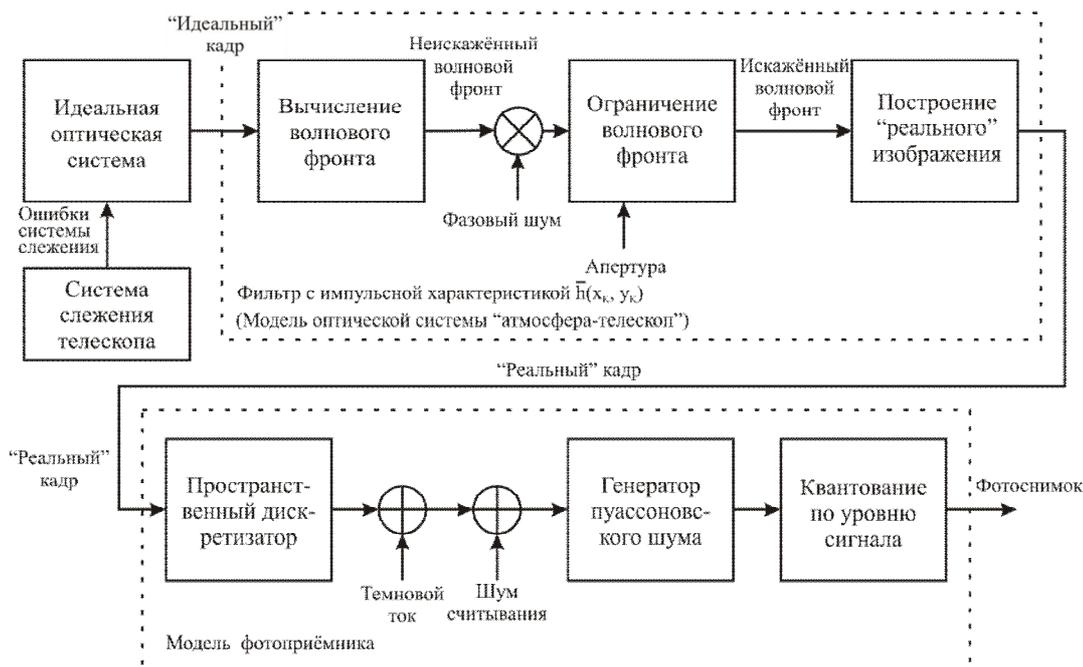


Рисунок 3 – Статистическая модель проведения измерений

Ошибки системы слежения в разработанной модели представляются в виде белого гауссовского шума, приводящего к отклонению от истинного положения центра системы координат $x_K O y_K$, связанной с фокальной плоскостью телескопа. Данные ошибки объясняются неидеальностью механических узлов системы гидирования, ветровыми нагрузками, а также прочими воздействиями.

Влияние турбулентной атмосферы в данной модели соответствует действующему фазовому шуму $\varphi(x_a, y_a, t)$ (x_a, y_a – координаты в плоскости апертуры телескопа) [5, 6]. Данный шум зачастую считается нормальным шумом с интервалом корреляции r_0 и временем корреляции τ [7]:

$$r_0 = \left(0.432 \left(\frac{2\pi}{\lambda} \right)^2 \sec z \int_0^L C_n^2(h) \left(\frac{h}{L} \right)^{5/3} dh \right)^{-3/5},$$

$$\tau = \frac{1}{v_{вет}} \left(3 \int_0^1 (1-\xi)^2 (e^{-1} a_s^{-1/3}(\xi L) + (1-e^{-1})(a_s^2(\xi L) + 2 \left(\frac{h_0}{5\pi} \right)^2)^{-1/6}) d\xi \right)^{-3},$$

где λ – длина волны, h – высота над уровнем океана, L – протяженность турбулентного слоя атмосферы, z – зенитное расстояние наблюдаемого объекта, $C_n^2(h)$ – структурный коэффициент температуры поля, $v_{вет}$ – скорость ветра в плоскости перпендикулярной к оси наблюдения, $a_s(\xi L)$ – эффективный радиус гауссовского пучка в турбулентной атмосфере, h_0 – высота трассы над подстилающей поверхностью.

Эти величины зависят от длины волны выбранного рабочего диапазона. В случае наиболее распространенных систем, время накопления которых значительно больше величины τ , турбулентная атмосфера рассматривается как телескоп с апертурой r_0 .

Совокупная оптическая система «атмосфера-телескоп» ввиду этих соображений описывалась воздействием фильтра с усреднённой импульсной характеристикой $\bar{h}(x_K, y_K)$ на изображение «идеального» кадра (показано пунктиром). Импульсная характеристика $\bar{h}(x_K, y_K)$ также будет определяться длиной волны выбранного рабочего диапазона.

Блок «Пространственный дискретизатор» моделирует вносимые искажения, обусловленные матричной структурой ПЗС-матрицы. Ввиду квантованности электронных пакетов ПЗС-матрицы, сигналы на выходе фото-

приёмника являются случайными, распределёнными по Пуассоновскому закону.

Блок «Квантование по уровню сигнала» имитирует процессы, происходящие в АЦП (для видимого диапазона разрядность АЦП равна 16, для ИК – не хуже 10).

3D-модель КО. Наблюдение за КО с помощью ПЗС-матрицы позволяет получить кривую блеска наблюдаемого объекта в диапазоне частот, воспринимаемых матрицей (для видимого и ИК-излучения). Для получения большей информации о КО на станциях наблюдения устанавливают фильтры, позволяющие разделить диапазон частот, воспринимаемый ПЗС матрицей, на несколько фотометрических полос. На рисунке 4 показаны спектральные чувствительности $Q(\lambda)$ различных фильтров в зависимости от длины волны. Как видно из приведенных зависимостей наблюдение может производиться в видимом, инфракрасном и ультрафиолетовом диапазонах. Таким образом, на станции наблюдения получается информация не только об отражающей способности КО, но и о средней степени нагрева его деталей. По полученной информации можно судить о состоянии аппарата при его известных рабочих характеристиках.

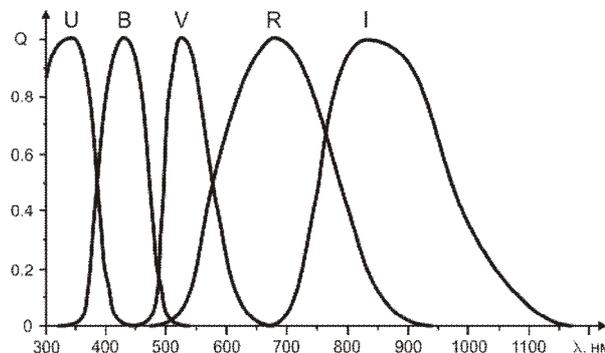


Рисунок 4 – Нормированная спектральная чувствительность ПЗС-матрицы для различных фильтров

Создана математическая 3D-модель существующих КО, позволяющая имитировать блеск КО в указанных полосах частот. Данная модель учитывает положение и поворот аппарата относительно наблюдателя, положение солнца, отражательную способность и степень нагрева отдельных деталей.

Экспериментальные исследования.

На рисунке 5, а,б показана траектория движения КО в потенциальной области видимости пункта наблюдения, находящегося в точке с координатами 37° долготы 51° широты, и области видимости телескопа (направление оптической оси телескопа – угол места 85° ,

азимут 52° , диаметр апертуры 80 см, фокусное расстояние объектива 6 м). Траектория КО рассчитана на период 1 суток с дискретным временем 20 с. Характеристики КО: масса 90 кг, средняя площадь $9,2 \text{ м}^2$; исходная орбита КО: наклонение орбиты 60° , аргумент перигея 30° , эксцентриситет 0,1, параметр орбиты $7,5 \cdot 10^3 \text{ км}$, время перигея 0. Снимок, показанный на рисунке 5, в, сформирован с помощью модели измерений с использованием фильтра R (рисунок 4), с полосой от 570 до 800 нм, по уровню 0,5, время экспозиции 10 с.

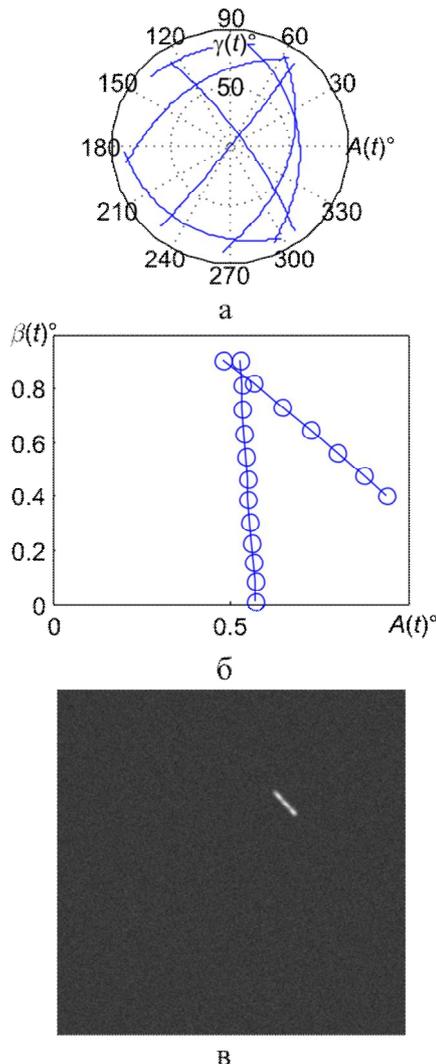


Рисунок 5 – Траектория движения КО:
а - верхняя полусфера видимости с пункта наблюдения; б - объектив телескопа; в - фотоснимок с ПЗС матрицы

Выводы. Разработанная имитационная модель позволяет моделировать возмущенное движение КО, учитывать его конструктивные особенности, а также имитировать процесс наблюдения за космическим телом с наземных измерительных пунктов с пассивным режимом работы в видимом и/или ИК-диапазонах спектра. Предложенная модель позволяет учитывать

факторы, действующие на траекторию КО: неравномерность поля тяготения Земли, сопротивление атмосферы, давление света, солнечный ветер, гравитационные поля других небесных тел, взаимодействие с электромагнитным полем Земли, а также факторы, влияющие на процесс формирования фотоснимков: неточность системы слежения; искажения волнового фронта, вносимые турбулентной атмосферой; дифракционные искажения, вызванные ограниченной апертурой телескопа; шум, обусловленный пространственной дискретностью ПЗС-матрицы; фотонный и темновой шум ПЗС-матрицы, что отличает предложенную модель от существующих [8]. Также модель обеспечивает точность описания флуктуаций силы тяготения не менее $10^{-3} \%$ за счет учета до 21-й составляющей ряда разложения гравитационного поля Земли. Позволяет проводить наблюдения из любой точки Земного шара в диапазонах длин волн 390-470 нм, 500-600 нм, 600-790 нм и 760-980 нм.

Работа выполнена в рамках реализации ФЦП «Научные и научно-педагогические кадры инновационной России» на 2009 – 2013 годы (контракт №П529 от 14 мая 2010 г.) при поддержке гранта Президента Российской Федерации для государственной поддержки молодых российских ученых МК-5000.2010.9 (договор №02.120.11.5000-МК), гранта Президента Российской Федерации для государственной поддержки молодых российских ученых МК-1581.2009.8 (договор №02.120.11.1581-МК).

Библиографический список

1. Молотов И.Е., Агапов В.М., Куприянов В.В., и др. Научная сеть оптических инструментов для астрометрических и фотометрических наблюдений // "Известия ГАО РАН" Режим доступа: <http://lfnv.astronomer.ru/report/0000042/index.htm>.
2. Аксенов Е.П. Теория движения искусственных спутников Земли. – М.: Наука, 1977. – 360 с.
3. Охоцимский Д.Е., Сихарулидзе Ю.Г. Основы механики космического полёта: учеб. пособие. - М.: Наука. Гл. ред. Физ.-мат. Лит., 1990. - 448 с.
4. Иванов Н.М., Лысенко Л.Н. Баллистика и навигация космических аппаратов. М.: Дрофа, 2004. - 544 с.
5. Свиридов К.Н. Оптическая локация космического мусора. – М.: Знание, 2006. – 488 с.
6. Щеглов П.В. Проблемы оптической астрономии. – М.: Наука. Главная редакция физико-математической литературы, 1980. – 272 с.
7. Зуев В.Е., Банах В.А., Покасов В.В. Оптика турбулентной атмосферы. Том 5. Л.: Гидрометеоздат, 1988. – 272 с.
8. Pradipta S., Lawrence D. System of on board prediction of trajectories for autonomous navigation of GPS satellites. Patent № US19970802620. Режим доступа: <http://www.patentstorm.us/>.

УДК 675.019.5

*П.В. Бабаян, Н.В. Серёгина***СЕГМЕНТАЦИЯ ИЗОБРАЖЕНИЙ ПОЛУФАБРИКАТОВ КОЖИ**

Работа посвящена задаче сегментации изображений полуфабрикатов кожи с помощью системы технического зрения. Рассмотрены два метода сегментации и их модификации, приведены результаты сравнительных экспериментальных исследований этих методов.

Ключевые слова: сегментация, сегментация по водоразделам с использованием маркеров, поиск пороков, пороговая обработка, кожа.

Введение. В настоящее время системы технического зрения (СТЗ) широко применяются в различных отраслях промышленности. Настоящая работа посвящена разработке СТЗ для автоматического определения сорта кожи, предназначенной для использования в кожевенном производстве. Обработка кожи является трудоемким процессом, включающим порядка десяти основных этапов. На трех из них происходит сортировка полуфабриката кожи. В настоящее время эта сортировка выполняется оператором. Разрабатываемая СТЗ должна «уметь» автоматически выделять пороки на изображении полуфабриката кожи, распознавать их, анализировать и в режиме реального времени принимать решение о сортности кожи. Здесь и далее под «пороками» кожи понимаются дефекты, вызванные процессом жизнедеятельности животного, нарушением правил хранения, консервации и обрядки или нарушением технологии обработки.

СТЗ для определения сорта кожи должна решать следующие основные задачи [1]:

1. Обнаружение пороков.
2. Классификация пороков с оценкой площади поражения.
3. Идентификация зон нахождения пороков.
4. Оценка площади полуфабриката.
5. Принятие решения о сорте кожи.

Решение перечисленных задач осуществляется на основе применения различных методов сегментации.

Процедура обнаружения пороков затрудняется тем, что внешний вид порока может сильно варьироваться даже для пороков одного типа, так как в каждом конкретном случае различаются площадь, интенсивность, освещенность и характер поражения. Таким образом, невозможно сформировать эталонные изображения пороков. По этой причине обнаружение пороков

должно проводиться по характерным признакам (яркостным, текстурным или другим), свойственным каждой группе пороков.

Цель работы. Основной целью является разработка алгоритмов сегментации изображений, ориентированных на использование в СТЗ для обнаружения и измерения пороков полуфабрикатов кожи. В дальнейшем будут рассмотрены алгоритмы, основанные на следующих двух методах:

- метод пороговой обработки;
- метод сегментации по водоразделам с использованием маркеров.

Метод пороговой обработки. В соответствии с методом пороговой обработки сегментация наблюдаемого изображения $l(i, j)$ производится по правилу:

$$r = \{(i, j) : l(i, j) \geq T\}, \quad (1)$$

где r – бинарное изображение, заданное множеством точек с координатами (i, j) ; T – порог.

Существует несколько подходов к вычислению порога T . В рамках данной работы используется метод Отсу, основанный на максимизации межклассовой дисперсии [2]. Данный метод опробован на ряде изображений полуфабрикатов кожи. Результаты экспериментов показали, что сегменты, соответствующие порокам, оказываются сильно фрагментированными. Для решения этой проблемы было решено использовать морфологическую обработку полученного в результате пороговой обработки бинарного изображения.

Морфологическая обработка состоит из нескольких операций, проводимых над полученным в результате пороговой обработки бинарным изображением r :

- 1). Операция эрозии со структурообразующим элементом в форме круга с заданным

радиусом. (Формируется изображение r_e);

2). Морфологическая реконструкция с маркером r_e и маской r . Структурообразующий элемент в форме квадрата размерами 3×3 . (Формируется изображение r_2);

3). Дилатация изображения r_2 . (Формируется изображение r_d);

4). Морфологическая реконструкция с маркером r_d и маской r_2 . Структурообразующий элемент в форме квадрата размерами 3×3 .

Рекомендуемые структурообразующие элементы для операции эрозии – круг с радиусом 10 пикселей, для дилатации – круг с радиусом 4 пикселя.

Реконструкция является морфологическим преобразованием [3], в котором участвуют два изображения (маркер и маска) и один структурообразующий элемент. Используемый структурообразующий элемент определяет связность.

Если g – это маска, f – маркер, B – структурообразующий элемент, причем $f \subseteq g$, то реконструкция g по f , которая обозначается $R_g(f)$, определяется следующей итеративной процедурой.

1. Инициализация: присвоить h_1 маркерное изображение f .
2. Построить структурообразующий элемент B .
3. Повторять:

$$h_{k+1} = (h_k \oplus B) \cap g, \quad (2)$$

где \oplus – операция дилатации, до тех пор, пока не станет $h_{k+1} = h_k$.

Вследствие неравномерности средней яркости наблюдаемого изображения, описанный подход не всегда обеспечивает требуемую точность сегментации, поэтому разработана модификация алгоритма, предполагающая разбиение изображения на области. Для каждой из областей ищется порог по методу Отсу, проводится пороговая сегментация и морфологическая обработка. Затем изображение восстанавливается из отдельных, уже обработанных, частей. Для полуфабрикатов размером около $3 \times 2 \text{ м}^2$ приемлемое качество сегментации достигается при разбиении изображения на 36 прямоугольных областей равной площади.

Сегментация по водоразделам с использованием маркеров. Понятие водораздела основано на представлении изображения как трехмерной поверхности, заданной двумя прост-

ранственными координатами и уровнем яркости в качестве высоты поверхности (рельефа) [4]. В такой интерпретации рассматриваются точки трех видов: а) точки локального минимума; б) точки, находящиеся на склоне (с которых вода скатывается в один и тот же локальный минимум); в) точки, находящиеся на пике (с которых вода с равной вероятностью скатывается более чем в один такой минимум). Применительно к конкретному локальному минимуму набор точек, удовлетворяющих условию (б), назовем бассейном или водосбором этого минимума. Множество точек, удовлетворяющих условию (в), образуют линии гребней и поверхности рельефа и называются линиями водораздела.

Непосредственное применение алгоритма сегментации по водоразделам обычно приводит к повышенной фрагментации сегментов изображения. Решение этой проблемы возможно путем включения в состав процедуры шага предварительной обработки, состоящего из морфологической обработки полученного изображения и построения маркеров. Под «маркером» здесь понимается подмножество точек, гарантированно принадлежащих искомой области изображения (неповрежденному участку кожи либо пороку). Маркеры играют роль стартовых точек, вокруг которых «выращиваются» сегменты. Различают внутренние маркеры, относящиеся к порокам, и внешние маркеры, соответствующие фону изображения.

Можно сформулировать следующие основные этапы предлагаемого метода сегментации.

1. Морфологическая обработка исходного изображения.
2. Построение внутренних маркеров.
3. Сегментация по водоразделам с тем ограничением, что в качестве локальных минимумов рассматриваются только внутренние маркеры.
4. Построение внешних маркеров, разграничивающих изображение на области, каждая из которых содержит единственный внутренний маркер и часть фоновой области.
5. Сегментация полученных областей по водоразделам.
6. Морфологическая обработка.

В основе **морфологической обработки** (этап 1) лежат операции морфологического размывания и замыкания. Итоговый результат выполнения двух операций состоит в удалении или ослаблении как темных, так и светлых артефактов шума. Над полутоновым изображением $I(i, j)$ проводится следующий ряд операций:

- размывание изображения $I(i, j)$ [получаем изображение $I_2(i, j)$];

- эрозия изображения $l(i, j)$ [получаем изображение $l_3(i, j)$];

- реконструкция. Маркер – $l_3(i, j)$, маска – $l(i, j)$ [получаем изображение $l_4(i, j)$];

- операция замыкания изображения $l_2(i, j)$ [получаем изображение $l_5(i, j)$];

- операция дилатации над изображением $l_4(i, j)$ [получаем изображение $l_6(i, j)$];

- размыкание реконструкции. Маркер – замыкание реконструкции $l_6(i, j)$, маска – замыкание реконструкции $l_4(i, j)$ [получаем изображение $l_7(i, j)$];

- замыкание реконструкции $l_7(i, j)$.

Для всех морфологических операций используется структурообразующий элемент в форме круга радиусом 5 пикселей.

Определения основных используемых морфологических операций приведены в [3].

Процедура **построения внутренних маркеров** (этап 2) состоит в поиске областей, окруженных точками с большей «высотой»; таких, что их точки образуют компоненту связности площадью не менее заданной; и все точки областей имеют одинаковое значение яркости.

Для выполнения **сегментации по водоразделам** (этап 3) целесообразно перейти к рассмотрению градиентного изображения $g(x, y)$, для вычисления которого используется модуль градиента Собеля.

Алгоритм сегментации по водоразделам предполагает нахождение множества C^* , включающего все точки изображения, кроме линий водораздела. Поиск множества C^* выполняется итерационно. Пусть

$$C(n) = \bigcup_{i=1}^R C_n(M_i), \quad (3)$$

где $C_n(M_i)$ – множество точек бассейна, соответствующих маркеру M_i , которые оказались заполненными на n -ом шаге алгоритма; R – количество маркеров.

Обозначим $T(n)$ как множество точек, в которых поверхность $g(x, y)$ лежит ниже плоскости $g(x, y) = n$.

Выполнение алгоритма начинается с инициализации, где g_{\min} – это наименьшее значение яркости градиентного изображения. После этого алгоритм выполняется итерационно в предположении, что на n -м шаге множество $C(n-1)$ уже построено. При построении множества $C(n)$ учитывается, что каждая компонента связности

множества $C(n-1)$ содержится ровно в одной связной компоненте множества $T(n)$. Для получения множества $C(n)$ из множества $C(n-1)$ применяется следующая процедура. Пусть $Q(n)$ – множество компонент связности множества $T(n)$. Тогда для каждой связной компоненты $q \in Q(n)$ есть три исхода:

а) $q \cap C(n-1)$ – пустое множество;

б) $q \cap C(n-1)$ содержит единственную компоненту связности множества $C(n-1)$;

в) $q \cap C(n-1)$ содержит более одной компоненты связности множества $C(n-1)$.

Характер дальнейших действий зависит от того, какое из трех условий имеет место. Условие (а) означает, что встретился локальный минимум, не соответствующий ни одному маркеру, и никаких действий не производится. Условие (б) имеет место, когда q лежит внутри бассейна некоторого маркера, в этом случае для построения $C(n)$ компонента q добавляется к $C(n-1)$. Условие (в) выполняется, когда встретились точки водораздела двух или более бассейнов. В этом случае дальнейшее заполнение приводит к слиянию этих бассейнов, поэтому из множества q следует исключить линию (или линии) водораздела и оставшиеся элементы добавить в $C(n-1)$.

На последнем шаге алгоритма (при $n = g_{\max} + 1$, где g_{\max} – максимальная яркость градиентного изображения) получится искомое множество C^* .

Для **построения внешних маркеров** (этап 4) используются пиксели, которые расположены точно посередине между внутренними маркерами.

Повторная сегментация (этап 5), но уже не всего изображения, а полученных после выполнения третьего этапа областей, осуществляется по водоразделам. Как правило, в результате сегментации по водоразделам с использованием маркеров полученные сегменты, разделенные линиями водораздела, можно четко разделить на два класса – области, относящиеся к фону, и области, относящиеся к объекту. Классификацию можно провести по следующему алгоритму.

1. Определение средней яркости каждого сегмента.

2. Определение средней яркости всего изображения.

3. Сравнение средней яркости каждого сегмента со средней яркостью всего изображения. Если яркость сегмента меньше средней яркости

всего изображения, то данный сегмент классифицируется как объект.

Морфологическая обработка бинарного изображения (этап 6) аналогична обработке бинарного изображения, полученного в результате пороговой обработки (п. 2).

5. Сравнение, используемых алгоритмов сегментации

Для адекватного оценивания и сравнения работоспособности алгоритмов сегментации друг с другом требуется ввести количественные характеристики работоспособности (КХР).

Рассмотрим две КХР [5]:

- процент неверно отсегментированных пикселей (P_r);

- частоты верного и ложного выделения ($\hat{P}_{лв}$ и $\hat{P}_{пв}$).

Для видеопоследовательности результат усредняется по всем изображениям.

При оценке по частоте верного и ложного выделения используются следующие формулы:

$$\hat{P}_{лв} = E \left[\frac{N(\hat{r} \cap r^c)}{N(r^c)} \right], \quad (4)$$

$$\hat{P}_{пв} = E \left[\frac{N(\hat{r} \cap r)}{N(r)} \right], \quad (5)$$

где $\hat{P}_{лв}$ - частота ложного выделения;

$\hat{P}_{пв}$ - частота правильного выделения;

E – оператор усреднения;

N – операция вычисления количества элементов множества;

\hat{r} - множество точек бинарного изображения, полученного при сегментации;

r - множество точек эталонной бинарной маски;

r^c - дополнение множества r .

Отметим, что в (4) и (5) при выполнении операций дополнения в качестве универсального множества рассматривается множество всех точек изображения.

Сравнение двух методов сегментации с помощью этого критерия должно осуществляться при одинаковом показателе частоты ложного выделения.

При проведении экспериментальных исследований использовалось полутоновое изображение полуфабриката кожи размером 1340x1127 пикселей, пораженного пороком типа «свищи». Результаты сравнения методов сегментации приведены в таблице. Используются следующие сокращения: ПО – пороговая обработка, СВМ – сегментация на основе водоразделов с использо-

ванием маркеров, I-III – модификации метода пороговой обработки (I – пороговая обработка всего изображения без использования морфологической обработки; II – пороговая обработка всего изображения с последующей морфологической обработкой сегментированного изображения; III – разбиение изображения на части, сегментация этих частей на основе пороговой обработки с последующей морфологической обработкой).

По результатам экспериментальных исследований можно сделать следующие выводы.

1. Сегментация изображений, имеющих бимодальные гистограммы, показывает наилучшие результаты при применении метода пороговой обработки.

2. Для изображений, имеющих полимодальные гистограммы, лучшие результаты показывает сегментация по водоразделам с использованием маркеров.

3. При обнаружении мелких пороков (менее 0,5 см²) сегментация по водоразделам с использованием маркеров осуществляется с высокой степенью фрагментации.

4. При обработке двух одинаковых изображений сегментация на основе водоразделов занимает больше времени, чем пороговая обработка. Например, для изображения размером 1134x1127 сегментация на основе водоразделов занимает 33 с, а пороговая обработка – 7 с.

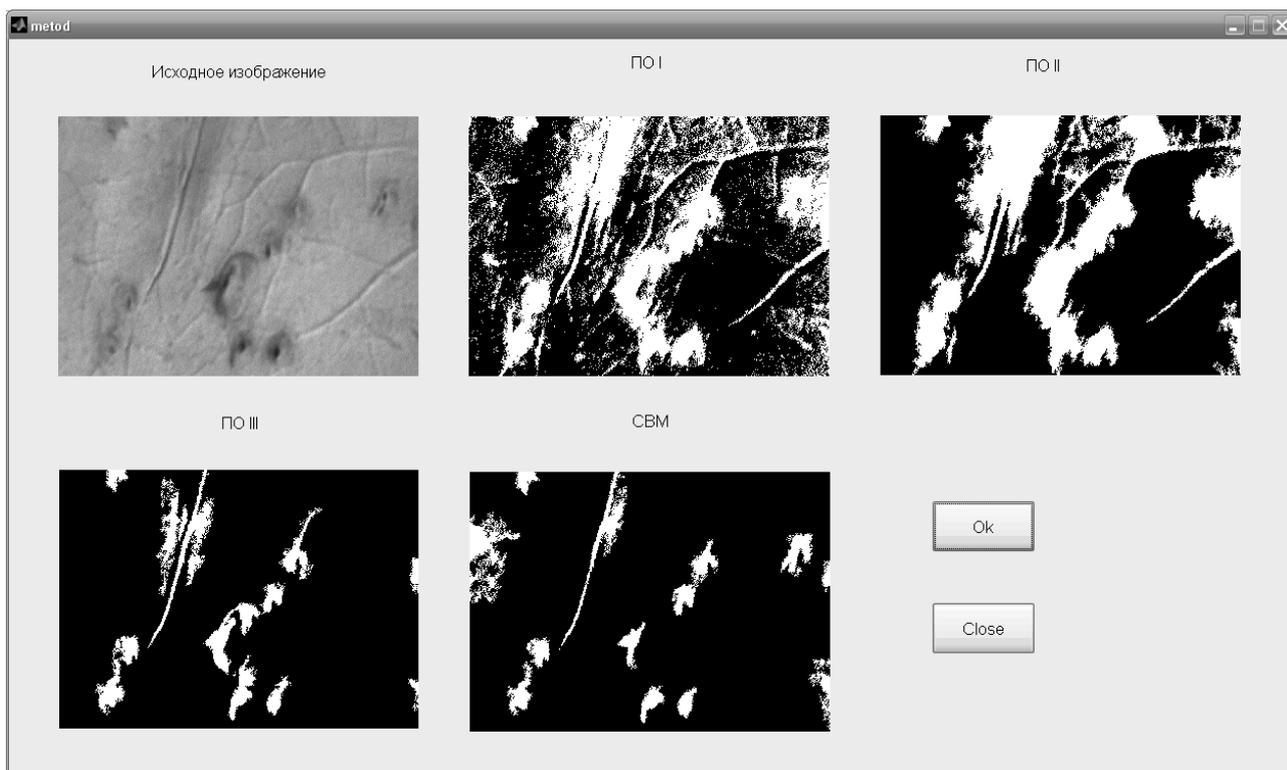
На рисунке приведен пример сегментации изображения полуфабриката кожи, пораженного пороком типа «свищи», различными методами. Основные сокращения на рисунке соответствуют сокращениям в таблице.

Сравнение методов сегментации

Метод сегментации	$\hat{P}_{лв}$	$\hat{P}_{пв}$	P_r , %
ПО I	0,056	0,70	9,5
ПО II		0,82	9,1
ПО III		0,89	8,1
СВМ		0,85	7,95

Заключение. Таким образом, с точки зрения решения задач сегментации изображений полуфабрикатов кожи, наилучшими показателями качества обладает модифицированный алгоритм пороговой обработки (ПО III), использующий разбиение наблюдаемого изображения на области, а также морфологическую фильтрацию бинарных изображений.

Результаты данной работы ориентированы на использование при разработке СТЗ для кожевенной промышленности.



Изображения, сегментированные различными методами

Работа проводилась при финансовой поддержке Министерства образования и науки РФ (госконтракт № 02.740.11.0002).

Библиографический список

1. Серегина Н.В. Разработка алгоритмов определения сорта полуфабриката кожи на основе анализа изображений // Новые информационные технологии в научных исследованиях и в образовании: XIV Всероссийская научно-техническая конф. студентов, молодых учёных и специалистов. – Рязань, 2009.

2. N. Otsu. A threshold selection method from gray-level histograms. IEEE Trans. Sys., Man., Cyber. 1979.

3. Гонсалес Р., Вудс Р., Эддингс С. Цифровая обработка изображений в среде MATLAB. М.: ТЕХНОСФЕРА, 2006.

4. Гонсалес Р., Вудс Р., Цифровая обработка изображений. М.: ТЕХНОСФЕРА, 2006.

5. Алпатов Б.А., Бабаян П.В., Смирнов С.А. Автоматическое слежение за объектами при отсутствии априорных сведений о фоноцелевой обстановке // Цифровая обработка сигналов. – 2009. – №3.

УДК 621.391

С.Н. Бузыкканов

ВЕЙВЛЕТ-ОБРАБОТКА ИЗОБРАЖЕНИЙ В ВЕСОВОМ ПРОСТРАНСТВЕ СОБОЛЕВА

Предложен алгоритм получения вейвлет-функций, ортогональных в весовом пространстве Соболева W_2^1 . Обоснован алгоритм применения полученных вейвлет-функций для обработки видеоизображений. Показано, что при обработке видеоизображений с помощью предложенных вейвлет-функций удастся повысить концентрацию их энергии с 91 % в пространстве L_2 до 95-96 % в весовом пространстве Соболева в зависимости от весового коэффициента α и начального уровня разложения.

Ключевые слова: ортогональные вейвлет-функции, весовое пространство Соболева W_2^1 , цифровая обработка изображений.

Введение. Одним из эффективных современных методов анализа сигналов является вейвлет-анализ, который позволяет провести разложение сигналов с помощью локализованных в частотно-временной области базисных функций. Вейвлеты представляют альтернативу анализу Фурье; их свойство локализации и их иерархическая конструкция - основные причины почему они предоставляют новый инструмент для решения многих проблем. В частности, при рассмотрении вейвлет-базисов для обработки видеоизображений подразумевают базис в пространстве L_2 , сформированный с помощью масштабирования и сдвига одной функции [1], т.е.

$$\psi_{j,k}(x) = 2^{j/2} \psi(2^j x - k), \quad j, k \in Z. \quad (1)$$

Однако, как это было показано в [1...5], более естественным для радиотехнических сигналов является представление в весовом пространстве Соболева W_2^1 , которое накладывает дополнительные ограничения на скорость изменения сигналов, т.е. их производную. В связи этим представляет интерес рассмотрение алгоритмов обработки видеоизображений при построении вейвлет-базисов в весовом пространстве Соболева W_2^1 .

Цель работы. Разработка процедуры построения вейвлет-функций, ортогональных в весовом пространстве Соболева W_2^1 , и исследование их свойств при обработке видеоизображений.

Теоретические исследования. Теоретические исследования возможности построения вейвлет-базисов в пространстве Соболева проводятся с конца девяностых годов прошлого века как в нашей стране [6,7], так и за рубежом [8...15]. Опираясь на данные исследования, рассмотрим возможности вейвлет-преобразования в весовом пространстве Соболева W_2^1 . Весовым пространством Соболева W_2^1 называется пространство, в котором норма функции и скалярное произведение определяются следующими выражениями [3,4]:

$$\|f(t)\|_1^2 = (1-\alpha) \int_{-\infty}^{\infty} f(t)f(t)dt + \alpha \int_{-\infty}^{\infty} f'(t)f'(t)dt, \quad (2)$$

$$\langle f(x), g(x) \rangle_W = (1-\alpha) \int_{-\infty}^{\infty} f(x)g(x)dx + \alpha \int_{-\infty}^{\infty} f'(x)g'(x)dx, \quad (3)$$

или в спектральной области

$$\|f(t)\|_1^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} (1-\alpha + \alpha w^2) |\hat{f}(w)|^2 dw, \quad (4)$$

$$\langle f(x), g(x) \rangle_W = \frac{1}{2\pi} \int_{-\infty}^{\infty} (1-\alpha + \alpha w^2) \hat{f}(w) \overline{\hat{g}(w)} dw, \quad (5)$$

где $\hat{f}(w), \hat{g}(w)$ - спектры соответствующих функций.

Согласно [1], вейвлет-разложение сигнала может быть записано в виде:

$$f(t) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \langle f(t), \psi_{j,k}(t) \rangle_W \tilde{\psi}_{j,k}(t), \quad (6)$$

где $\langle \psi_{j,k}, \tilde{\psi}_{j',k'} \rangle_W = \delta_{kk'} \delta_{jj'}$. Зачастую для простоты вычислений стараются применить ортогональные функции, удовлетворяющие более строгому условию $\langle \psi_{j,k}, \tilde{\psi}_{j',k'} \rangle_W = \delta_{kk'}$. В этом случае для разложения и восстановления функции используется один и тот же базис. На практике наибольшее распространение получил алгоритм Малла [16] (быстрое вейвлет-преобразование), когда производят разложение на каждом уровне на высокочастотную и низкочастотную составляющие:

$$C\varphi_{j,k}(x) = \langle f(x), \varphi_{j,k}(x) \rangle, \quad (7)$$

$$C\psi_{j,k}(x) = \langle f(x), \psi_{j,k}(x) \rangle. \quad (8)$$

С учетом (3) в весовом пространстве Соболева выражения (7) и (8) соответственно принимают вид:

$$C\varphi_{j,k}(x) = (1-\alpha) \int_{-\infty}^{\infty} f(x)\varphi_{j,k}(x)dx + \alpha \int_{-\infty}^{\infty} f'(x)\varphi'_{j,k}(x)dx, \quad (9)$$

$$C\psi_{j,k}(x) = (1-\alpha) \int_{-\infty}^{\infty} f(x)\psi_{j,k}(x)dx + \alpha \int_{-\infty}^{\infty} f'(x)\psi'_{j,k}(x)dx. \quad (10)$$

Для упрощения дальнейших преобразований введем следующие обозначения:

$$a_{j,k} = \int_{-\infty}^{\infty} f(x)\varphi_{j,k}(x)dx, \quad (11)$$

$$b_{j,k} = \int_{-\infty}^{\infty} f'(x)\varphi'_{j,k}(x)dx = - \int_{-\infty}^{\infty} f(x)\varphi''_{j,k}(x)dx, \quad (12)$$

$$c_{j,k} = \int_{-\infty}^{\infty} f(x)\psi_{j,k}(x)dx, \quad (13)$$

$$d_{j,k} = \int_{-\infty}^{\infty} f'(x)\psi'_{j,k}(x)dx = - \int_{-\infty}^{\infty} f(x)\psi''_{j,k}(x)dx. \quad (14)$$

Также для проведения преобразований необходимо использовать условия масштабирования, известные из теории кратномасштабного анализа [1]:

$$\varphi_{j-1,k}(x) = \sqrt{2^{j-1}} \varphi(2^{j-1}x - k) =$$

$$= \sqrt{2^j} \sum_m h_m \varphi(2^j x - 2k - m), \quad (15)$$

$$\begin{aligned} \psi_{j-1,k}(x) &= \sqrt{2^{j-1}} \psi(2^{j-1} x - k) = \\ &= \sqrt{2^j} \sum_m g_m \psi(2^j x - 2k - m). \end{aligned} \quad (16)$$

В этом случае можно записать:

$$\begin{aligned} a_{j,k} &= \int_{-\infty}^{\infty} f(x) \varphi_{j,k}(x) dx = \\ &= \sqrt{2^{j+1}} \sum_m h_m \int_{-\infty}^{\infty} f(x) \varphi(2^j x - 2k - m) dx = \\ &= \sqrt{2^{j+1}} \sum_m h_m a_{j+1,2k+m}, \end{aligned} \quad (17)$$

$$\begin{aligned} c_{j,k} &= \int_{-\infty}^{\infty} f(x) \psi_{j,k}(x) dx = \\ &= \sqrt{2^{j+1}} \sum_m h_m \int_{-\infty}^{\infty} f(x) \psi(2^j x - 2k - m) dx = \\ &= \sqrt{2^{j+1}} \sum_m h_m c_{j+1,2k+m}. \end{aligned} \quad (18)$$

Для вычисления коэффициентов $b_{j,k}$ и $d_{j,k}$ приведем вспомогательные выкладки.

$$\varphi_{j-1,k}(x) = \sqrt{2^j} \sum_m h_m \varphi(2^j x - 2k - m), \quad (19)$$

$$\begin{aligned} \varphi'_{j-1,k}(x) &= \sqrt{2} \sum_m h_m \varphi'_{j,2k+m}(x) = \\ &= 2^j \sqrt{2^j} \sum_m h_m \varphi'(2^j x - 2k - m), \end{aligned} \quad (20)$$

$$\varphi''_{j-1,k}(x) = 2^{2j} \sqrt{2^j} \sum_m h_m \varphi''(2^j x - 2k - m), \quad (21)$$

$$\begin{aligned} &\sum_m h_m \varphi''(2^j x - 2k - m) = \\ &= -\sum_m h_m \int_{-\infty}^{\infty} 2^{-j} (2^{-j} w)^2 \Phi(2^{-j} w) e^{-iw(2k+m)} e^{iwx} dw = \\ &= \left| \begin{aligned} Hp(w) &= w^2 H(w) \\ Hp(w) &= \sum_m h p_m e^{-iwm} \end{aligned} \right| = \\ &= -2^{-2j} \int_{-\infty}^{\infty} \sum_m h p_m e^{-iwm} 2^{-j} \Phi(2^{-j} w) e^{iw(x-2k)} dw = \\ &= -2^{-2j} \sum_m h p_m \int_{-\infty}^{\infty} 2^{-j} \Phi(2^{-j} w) e^{-iw(2k+m)} e^{iwx} dw = \\ &= -2^{-2j} \sum_m h p_m \varphi(2^j x - 2k - m), \end{aligned} \quad (22)$$

т.е.

$$\varphi''_{j-1,k}(x) = -\sqrt{2^j} \sum_m h p_m \varphi(2^j x - 2k - m). \quad (23)$$

Аналогично получаем:

$$\psi_{j-1,k}(x) = \sqrt{2^j} \sum_m g_m \psi(2^j x - 2k - m), \quad (24)$$

$$\begin{aligned} \psi'_{j-1,k}(x) &= \sqrt{2} \sum_m g_m \psi'_{j,2k+m}(x) = \\ &= 2^j \sqrt{2^j} \sum_m g_m \psi'(2^j x - 2k - m), \end{aligned} \quad (25)$$

$$\psi''_{j-1,k}(x) = 2^{2j} \sqrt{2^j} \sum_m g_m \psi''(2^j x - 2k - m), \quad (26)$$

$$\psi''_{j-1,k}(x) = -\sqrt{2^j} \sum_m g p_m \psi(2^j x - 2k - m), \quad (27)$$

где

$$h p_m = \int_{-\infty}^{\infty} w^2 H(w) e^{iwm} dw, \quad g p_m = \int_{-\infty}^{\infty} w^2 G(w) e^{iwm} dw,$$

$$h_m = \int_{-\infty}^{\infty} H(w) e^{iwm} dw, \quad g_m = \int_{-\infty}^{\infty} G(w) e^{iwm} dw.$$

Используя (23), преобразуем (12)

$$\begin{aligned} b_{j,k} &= -\int_{-\infty}^{\infty} f(x) \varphi''_{j,k}(x) dx = \\ &= \int_{-\infty}^{\infty} f(x) \sqrt{2^{j+1}} \sum_m h p_m \varphi''(2^{j+1} x - 2k - m) dx = \\ &= \sqrt{2^{j+1}} \sum_m h p_m \int_{-\infty}^{\infty} f(x) \varphi(2^{j+1} x - 2k - m) dx = \\ &= \sqrt{2^{j+1}} \sum_m h p_m a_{j+1,2k+m}. \end{aligned} \quad (28)$$

С учетом полученных результатов перепишем (9) в виде:

$$C \varphi_{j,k}(x) = \sqrt{2^{j+1}} \sum_m ((1-\alpha) h_m + \alpha h p_m) a_{j+1,2k+m}. \quad (29)$$

Введем новое обозначение:

$$h W_m = (1-\alpha) h_m + \alpha h p_m \quad (30)$$

и произведем следующие преобразования в частотной области:

$$h W_m = \int_{-\infty}^{\infty} (1-\alpha + \alpha w^2) H(w) e^{iwm} dw. \quad (31)$$

В итоге окончательно можем записать:

$$C \varphi_{j,k}(x) = \sqrt{2^{j+1}} \sum_m h W_m a_{j+1,2k+m},$$

где

$$h W_m = \int_{-\infty}^{\infty} (1-\alpha + \alpha w^2) H(w) e^{iwm} dw. \quad (32)$$

С помощью аналогичных преобразований имеем

$$C \psi_{j,k}(x) = \sqrt{2^{j+1}} \sum_m g W_m a_{j+1,2k+m}, \quad (33)$$

где

$$\begin{aligned} g W_m &= (1-\alpha) g_m + \alpha g p_m = \\ &= \int_{-\infty}^{\infty} (1-\alpha + \alpha w^2) G(w) e^{iwm} dw. \end{aligned} \quad (34)$$

Рассмотрим более подробно выражения (15) и (16). В проведенных выкладках мы использовали привычные выражения

$$H(w) = \sum_m h_m e^{-iwm}, \quad h_m = \int_{-\pi}^{\pi} H(w) e^{iwm} dw. \quad (35)$$

Однако данные выражения соответствуют пространству L_2 . Проведем сравнительные преобразования в пространствах L_2 и W_2^1 .

Пространство L_2 :

$$\varphi_{j-1,k}(x) = \sqrt{2^j} \sum_m h_m \varphi_{j,2k+m}(x), \quad (36)$$

$$h_m = \frac{1}{\sqrt{2}} \int_{-\infty}^{\infty} \varphi_{j-1,k}(x) \varphi_{j,2k+m}(x) dx, \quad (37)$$

$$H(w) = \frac{1}{\sqrt{2}} \Phi_{j-1,k}(w) \Phi_{j,2k+m}(w). \quad (38)$$

Пространство W_2^1 :

$$\varphi_{j-1,k}(x) = \sqrt{2^j} \sum_m h w_m \varphi_{j,2k+m}(x), \quad (39)$$

$$h w_m = \frac{1}{\sqrt{2}} \left[(1-\alpha) \int_{-\infty}^{\infty} \varphi_{j-1,k}(x) \varphi_{j,2k+m}(x) dx + \alpha \int_{-\infty}^{\infty} \varphi'_{j-1,k}(x) \varphi'_{j,2k+m}(x) dx \right], \quad (40)$$

$$H W(w) = \frac{1}{\sqrt{2}} (1-\alpha + \alpha w^2) \Phi_{j-1,k}(w) \Phi_{j,2k+m}(w). \quad (41)$$

Таким образом, сравнивая выражения (38) и (41), получаем

$$H W(w) = (1-\alpha + \alpha w^2) H(w), \quad (42)$$

что точно соответствует выражению (31). Окончательно записываем:

$$C \varphi_{j-1,k}(x) = \langle f(x), \varphi_{j-1,k}(x) \rangle_W = \sum_m h_m \langle f(x), \varphi_{j,2k+m}(x) \rangle_W = \sum_m h_m C \varphi_{j,2k+m}(x), \quad (43)$$

$$C \psi_{j-1,k}(x) = \langle f(x), \psi_{j-1,k}(x) \rangle_W = \sum_m h_m \langle f(x), \psi_{j,2k+m}(x) \rangle_W = \sum_m h_m C \psi_{j,2k+m}(x). \quad (44)$$

Это доказывает, что в весовом пространстве Соболева правомерно использование алгоритма Малла для вычисления быстрого вейвлет-преобразования. При этом изменяется выражение для определения коэффициентов преобразования $h w_m$, а сам алгоритм вычислений не изменяется.

Для получения компактных вейвлет-функций используем отношение масштабирования и условие плотности [1]:

$$\begin{aligned} \bar{\varphi}^{(j)}(w) &= m_0^{(j+1)} \left(\frac{w}{2} \right) \bar{\varphi}^{(j+1)} \left(\frac{w}{2} \right) = \\ &= (1-\alpha + \alpha 2^{2j} w^2)^{-1/2} \prod_{p=1}^{+\infty} m_0^{(j+p)} (2^{-p} w). \end{aligned}$$

Характерной особенностью обработки сигналов в весовом пространстве Соболева W_2^1 является наличие сингулярности в точках $w = i 2^{-j} \sqrt{\frac{1-\alpha}{\alpha}}$. При построении компактных вейвлетов в весовом пространстве Соболева W_2^1 необходимо использовать такую функцию фильтра $m_0^{(j)}(w)$, которая компенсирует сингулярность, образующуюся в результате наличия в выражении для $\bar{\varphi}^{(j)}(w)$ множителя $(1-\alpha + \alpha 2^{2j} w^2)^{-1/2}$ из пространства W_2^1 . С этой целью по аналогии с [8] введем систему фильтров

$$m_{0,0}^{(j)}(w) = \frac{e^{-i w/2}}{\sqrt{ch \left(2^{-j} \sqrt{\frac{1-\alpha}{\alpha}} \right)}} \times \cos \left(\frac{w + i 2^{-j} \sqrt{\frac{1-\alpha}{\alpha}}}{2} \right). \quad (45)$$

Для целого $N \geq 1$ построим последовательность вида:

$$m_{0,N}^{(j)}(w) = m_{0,0}^{(j)}(w) \left(\frac{1 + e^{-i w}}{2} \right)^N L_N^{(j)}(w), \quad (46)$$

где $L_N^{(j)}(w)$ - тригонометрический полином с действительными коэффициентами. Второй множитель, как и в [1], введен для повышения регулярности вейвлет-функций, третий множитель выбирается так, чтобы выполнялось условие $|m_{N,m}^{(j)}(w)|^2 + |m_{N,m}^{(j)}(w + \pi)|^2 = 1$.

Используя метод конструирования через фильтры [1], определяем

$$\begin{aligned} \bar{\varphi}_0^{(j)}(w) &= \frac{1}{\sqrt{1-\alpha + i\sqrt{\alpha} 2^j w}} \prod_{p=1}^{+\infty} m_{0,0}^{(j+p)} (2^{-p} w) = \\ &= \sqrt[4]{\alpha(1-\alpha)} \frac{2^{1+\frac{j}{2}}}{(1-\alpha + \alpha 2^{2j} w^2)} \times \\ &\times \frac{e^{-i w/2}}{i \sqrt{sh \left(2^{-j} \sqrt{\frac{1-\alpha}{\alpha}} \right)}} \sin \left(\frac{w + i 2^{-j} \sqrt{\frac{1-\alpha}{\alpha}}}{2} \right), \quad (47) \\ \bar{\varphi}_N^{(j)}(w) &= \frac{1}{\sqrt{1-\alpha + i\sqrt{\alpha} 2^j w}} \prod_{p=1}^{+\infty} m_{0,N}^{(j+p)} (2^{-p} w) = \end{aligned}$$

$$= \tilde{\varphi}_0^{(j)}(w) e^{-iNw/2} \left(\frac{\sin(w/2)}{w/2} \right)^N \prod_{p=1}^{+\infty} L_N^{(j+p)}(2^{-p}w), \quad (48)$$

$$\begin{aligned} \widehat{\psi}_N^{(j)}(2w) &= -e^{-i\pi w} \overline{m_{0,N}^{(j+1)}(w+\pi)} \widehat{\varphi}_N^{(j+1)}(w) = \\ &= \frac{2^{1+j} \sqrt[4]{\alpha(1-\alpha)} e^{-i\pi w} i^{-N}}{(1-\alpha + \alpha 2^{2(j+1)} w^2)^N} \left(\frac{\sin^2(w/2)}{w/2} \right)^N \times \\ &\quad \times \frac{ch \left(2^{-j-1} \sqrt{\frac{1-\alpha}{\alpha}} \right) - \cos(w)}{\sqrt{sh \left(2^{-j} \sqrt{\frac{1-\alpha}{\alpha}} \right)}} \times \\ &\quad \times \overline{L_N^{(j+1)}(w+\pi)} \prod_{p=1}^{+\infty} L_N^{(j+p+1)}(2^{-p}w). \quad (49) \end{aligned}$$

Как показано в [8], при соответствующем выборе множителя $L_N^{(j)}(w)$, функции $\widehat{\varphi}_N^{(j)}(w)$ определяют кратномасштабный анализ в весовом пространстве Соболева W_2^1 .

Следующим шагом является получение функции $L_N^{(j)}(w)$. Для этого введем следующее обозначение [1]:

$$\left| L_N^{(j)}(w) \right|^2 = P_N^{(j)}(\sin^2(w/2)). \quad (50)$$

В этом случае выражение

$$\left| m_{0,N}^{(j)}(w) \right|^2 + \left| m_{0,N}^{(j)}(w+\pi) \right|^2 = 1, \quad (51)$$

при обозначении $y = \sin^2(w/2)$, преобразуется к виду

$$\begin{aligned} \left| m_{0,N}^{(j)}(w) \right|^2 + \left| m_{0,N}^{(j)}(w+\pi) \right|^2 &= \\ &= \left| m_{0,0}^{(j)}(w) \right|^2 P_N^{(j)}(y)(1-y)^N + \\ &+ \left| m_{0,0}^{(j)}(w+\pi) \right|^2 P_N^{(j)}(1-y)y^N. \quad (52) \end{aligned}$$

Для дальнейших преобразований рассмотрим более подробно выражение (44)

$$\begin{aligned} \left| m_{0,0}^{(j)}(w) \right|^2 &= \left| \frac{e^{-i\pi w/2}}{\sqrt{ch \left(2^{-j} \sqrt{\frac{1-\alpha}{\alpha}} \right)}} \cos \left(\frac{w + i 2^{-j} \sqrt{\frac{1-\alpha}{\alpha}}}{2} \right) \right|^2 = \\ &= \frac{ch^2 \left(2^{-j-1} \sqrt{\frac{1-\alpha}{\alpha}} \right) - y}{2ch^2 \left(2^{-j-1} \sqrt{\frac{1-\alpha}{\alpha}} \right) - 1}. \quad (53) \end{aligned}$$

Откуда получаем:

$$\left| m_{0,0}^{(j)}(w+\pi) \right|^2 = \frac{ch^2 \left(2^{-j-1} \sqrt{\frac{1-\alpha}{\alpha}} \right) - 1 + y}{2ch^2 \left(2^{-j-1} \sqrt{\frac{1-\alpha}{\alpha}} \right) - 1}. \quad (54)$$

Введя обозначение $x_j = ch^2 \left(2^{-j-1} \sqrt{\frac{1-\alpha}{\alpha}} \right)$,

окончательно записываем

$$\begin{aligned} \frac{x_j - y}{2x_j - 1} P_N^{(j)}(y)(1-y)^N + \\ + \frac{x_j - 1 + y}{2x_j - 1} P_N^{(j)}(1-y)y^N = 1. \quad (55) \end{aligned}$$

Данное выражение в точности соответствует выражению, рассмотренному в [8]. Для получения явного выражения, определяющего $L_N^{(j)}(w)$, воспользуемся полученными ранее [1] результатами.

При любых N и j существует тригонометрический полином степени N $L_N^{(j)}(w) = \sum_{k=0}^N b_{k,N}^{(j)} e^{-ikw}$ с действительными коэффициентами такой, что выполняется условие

$$\left| L_N^{(j)}(w) \right|^2 = P_N^{(j)}(\sin^2(w/2)). \quad (56)$$

Этот полином определяется единственным образом, если мы потребуем, чтобы корни выражения $\sum_{k=0}^N b_{k,N}^{(j)} z^k$ удовлетворяли условию $|z| \geq 1$ и

$$L_N^{(j)}(0) = \sqrt{P_N^{(j)}(0)} = \sqrt{2-1/x_j}. \quad (57)$$

Данный тригонометрический полином используется при определении $\varphi_N^{(j)}$ и $\psi_N^{(j)}$.

Пусть $N=1$, тогда

$$\begin{aligned} P_1^{(j)}(y) &= 2 + \frac{2y-1}{x_j} = \\ &= 2 - \frac{\cos(w)}{ch^2 \left(2^{-j-1} \sqrt{\frac{1-\alpha}{\alpha}} \right)} = \left| a_j + b_j e^{-i\pi w} \right|^2, \quad (58) \end{aligned}$$

где

$$a_j = \frac{\sqrt{ch \left(2^{-j} \sqrt{\frac{1-\alpha}{\alpha}} \right)} + \sqrt{2 + ch \left(2^{-j} \sqrt{\frac{1-\alpha}{\alpha}} \right)}}{\sqrt{2 \left(1 + ch \left(2^{-j} \sqrt{\frac{1-\alpha}{\alpha}} \right) \right)}}, \quad (59)$$

$$b_j = \frac{\sqrt{ch\left(2^{-j}\sqrt{\frac{1-\alpha}{\alpha}}\right)} - \sqrt{2+ch\left(2^{-j}\sqrt{\frac{1-\alpha}{\alpha}}\right)}}{\sqrt{2\left(1+ch\left(2^{-j}\sqrt{\frac{1-\alpha}{\alpha}}\right)\right)}}, \quad (60)$$

$$L_1^{(j)}(w) = a_j + b_j e^{-iw}. \quad (61)$$

Приведем вид скейлинг и вейвлет-функций для различных параметров α и j .

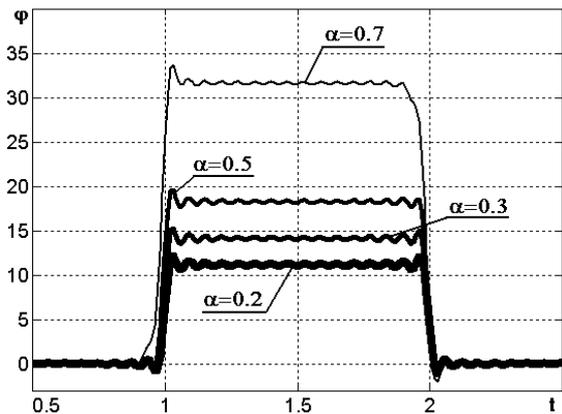


Рисунок 1 — Вид функций $\hat{\varphi}_1^{(-7)}$ для разных значений α

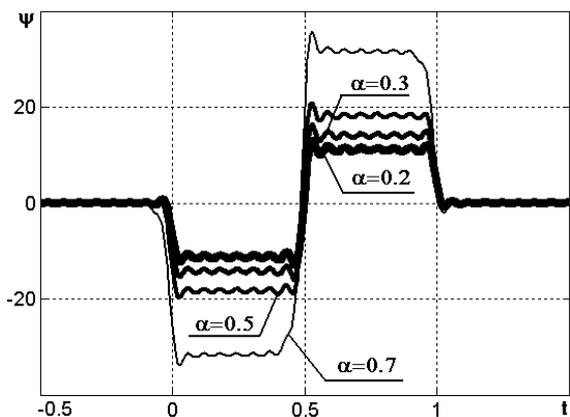


Рисунок 2 — Вид функций $\hat{\psi}_1^{(-7)}$ для разных значений α

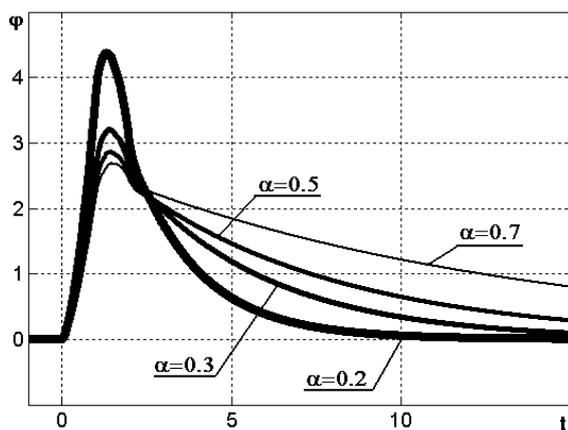


Рисунок 3 — Вид функций $\hat{\varphi}_1^{(2)}$ для разных значений α

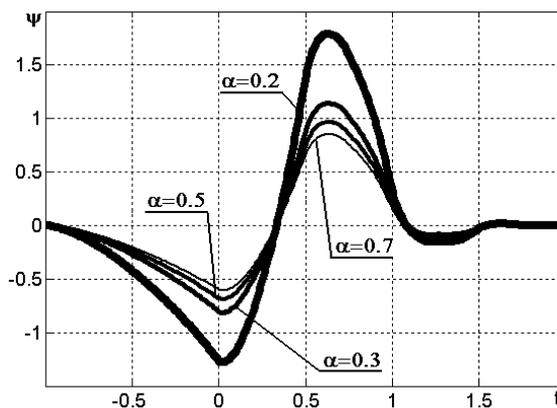


Рисунок 4 — Вид функций $\hat{\psi}_1^{(2)}$ для разных значений α

Экспериментальные исследования. Наиболее известным применением вейвлет-разложения является компактное представление видеоизображений, что нашло свое подтверждение в виде стандарта JPEG2000. Применим алгоритм быстрого вейвлет-преобразования (алгоритм Малла) с использованием полученных вейвлет-функций, ортогональных в весовом пространстве Соболева.

Пусть в качестве ортогональных вейвлет-функций используются функции (46)...(48) при $N=1$. В пространстве L_2 такими функциями являются вейвлеты Хаара (вейвлет Добеши-1). В пространстве Соболева вейвлет определяется выражениями (48) и представлен на рисунках 1...4. В качестве исходного изображения для проведения исследований возьмем стандартное изображение, представленное в математическом пакете Matlab (рисунок 5).



Рисунок 5 — Тестовое изображение

Доказано [8], что при использовании обработки в весовом пространстве Соболева вид

вейвлет-функции при переходе от одного масштаба к другому меняется, что приводит к адаптации к масштабу обрабатываемого изображения. Поэтому необходимо провести дополнительные исследования по выбору начального уровня разложения.

Сравнительным критерием будет являться степень концентрации энергии в низкочастотной области после нескольких уровней разложения. Большая степень концентрации определяет возможность получения более высокого качества изображений при одинаковой степени сжатия. Зависимость концентрации энергии от значения коэффициента α и начального уровня разложения для тестируемого изображения представлена на рисунке 6, где нижняя плоскость соответствует пространству L_2 , в котором результат преобразования не зависит от остальных параметров, а верхний график соответствует преобразованию в пространстве W_2^1 .

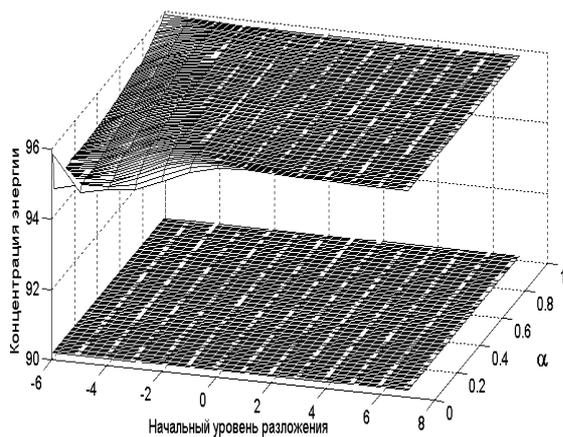


Рисунок 6 — Результат обработки тестового изображения

Анализ рисунка 6 позволяет сделать вывод, что применение ортогональных в весовом пространстве Соболева вейвлет-функций позволяет добиться роста концентрации энергии в низкочастотных отсчетах с 91 % в пространстве L_2 до 95-96 % в пространстве W_2^1 в зависимости от весового коэффициента α и начального уровня разложения.

Для визуальной демонстрации полученного результата приведем результат сжатия тестового изображения при обнулении 99 % отсчетов при обработке сигналов с помощью вейвлет-функций в пространстве L_2 (рисунок 7) и с помощью синтезированных вейвлет-функций (рисунок 8).

Сравнение рисунков 7 и 8 показывает, что при одной степени сжатия обработка в весовом пространстве Соболева обеспечивает лучшее качество изображения.



Рисунок 7 — Результат сжатия при обработке в пространстве L_2



Рисунок 8 — Результат сжатия при обработке в пространстве W_2^1

Выводы. В ходе проведенных исследований получены выражения для вычисления ортогональных в весовом пространстве Соболева W_2^1 вейвлет-функций и показано, что в данном пространстве правомерно использование алгоритма Малла для проведения быстрых вейвлет-преобразований. Экспериментальные исследования доказали целесообразность использования полученных вейвлет-функций для сжатия изображения. Для заданного изображения степень концентрации энергии удалось повысить с 91 % в пространстве L_2 до 96 % в весовом пространстве Соболева.

Работа выполнена при поддержке гранта Президента Российской Федерации для государственной поддержки молодых российских ученых МК-1581.2009.8 (договор № 02.120.11.1581-МК).

Библиографический список

1. Добеши И. Десять лекций по вейвлетам. Ижевск: НИЦ "Регулярная и хаотическая динамика", 2001. 464 с.
2. Соболев С.Л. Некоторые применения функционального анализа в математической физике: монография, изд. Ленингр. Гос. универ., 1950 (1988). 336с.
3. Кириллов С.Н., Бузыкканов С.Н. Двухканальная система передачи и обработки сигналов в пространстве Соболева // Вестник РГРТА. Вып. N.10. 2002. С. 6-8.
4. Кириллов С.Н., Бузыкканов С.Н. Оценка спектральной плотности мощности сигналов в модифицированном пространстве Соболева// Радиоэлектроника, 2002.- Т.45.-№12- С.46-51. (Известия высших учебных заведений).
5. Кириллов С.Н., Бузыкканов С.Н. Алгоритм дискретного спектрального анализа сигналов в модифицированном пространстве Соболева// Автоматика. 2003. №1. - С.88-94.
6. Субботин Ю.Н., Черных Н.И. Конструкция всплесков в $W_m^2(\mathbb{R})$ и их аппроксимативные свойства в разных метриках// Теория функций. Сборник научных трудов. Тр. ИММ. 11. № 2. 2005. - С.131-167.
7. Субботин Ю.Н., Черных Н.И. Всплески, ортонормированные относительно скалярного произведения в Соболевском пространстве W_m^2 периодических функций// Труды ИММ. 7:1 (2001). - С. 217-230.
8. F. Bastin, P. Laubin. Regular compactly supported Wavelets in Sobolev Spaces, Duke Mathematical Journal, Volume 87, 3, (April 1997), - P.481-508.
9. F. Bastin and P. Laubin. Compactly supported wavelets in Sobolev spaces of integer order. Applied and Computational Harmonic Analysis, Vol. 4, 1, (1997), - P.51-57.
10. F. Bastin and P. Laubin. A walk in the theory of wavelets from $L_2(\mathbb{R})$ to $H_s(\mathbb{R})$. Proceedings de "Third International Conference in Functional Analysis and Approximation Theory" Acquafredda di Maratea held in September 1996; Supplemento ai Rendiconti del Circolo Matematico di Palermo, Serie II, Numero 52, 1998, - P.239-252.
11. F. Bastin and C. Boigelot. Biorthogonal wavelets in $H_m(\mathbb{R})$. Journal of Fourier Analysis and Applications, Vol. 4, Issue 6, (1998), - P.749-768.
12. F. Bastin. Introduction, Construction and Applications of Wavelets in Sobolev spaces. Rev. Ciencias Matematicas, Habana, Vol 18, 2, (2000), 145-177 (Text based on the conferences of F. Bastin held at the occasion of the 4 April 26 - May 7 1999's workshop in Havana).
13. F. Bastin, C. Boigelot and P. Laubin. Wavelets in Sobolev spaces and applications. Coherent States, Quantization and Gravity, XVII Workshop on geometric methods in physics, Bialowieza, Poland, July 98, edited by M. Schlichenmaier, A. Strasburger, S. T. Ali, A. Odziejewicz, (2001), - P.113-127.
14. F. Bastin, C. Boigelot and P. Laubin. Spline wavelets in periodic Sobolev spaces and application to high order collocation methods. Revista de la Union Matematica Argentina, 44, 1, (2003), - P.53-74.
15. Mark A. Kon, Louise Arakelian Raphael. A Characterization of Wavelet Convergence in Sobolev Spaces.// Applicable Analysis: An International Journal, 1563-504X, Volume 78, Issue 3, 2001, - P.271 - 324
16. Малла С. Вэйвлеты в обработке сигналов. М.: Мир, 2005.

УДК 616.12

В.Г. Андреев, А.А. Кирьяков

ВЕКТОРНЫЙ АНАЛИЗ ПРОЦЕССОВ РЕГУЛЯЦИИ ФИЗИОЛОГИЧЕСКИХ ФУНКЦИЙ ЧЕЛОВЕКА

Рассмотрена задача векторного спектрального анализа диагностических данных, поступающих одновременно от трех датчиков: ритма сердца, объема воздуха в легких и капиллярного кровотока. Показано, что использование переопределенной векторной авторегрессионной модели даёт возможность по короткой (5...10 секунд) экспериментальной выборке выявить и оценить вариабельность сердечного ритма. В задаче экспресс-диагностирования процессов регуляции основных физиологических функций предложенный метод позволяет в 1,5...4 раза уменьшить относительную ошибку оценки частоты сокращений сердца по сравнению с обычной векторной авторегрессией того же порядка $p=2...4$. Выигрыш достигается за счет учета ошибок линейного предсказания, выходящих за длину p лага.

Ключевые слова: векторная линейная авторегрессия, спектр, векторное моделирование временных рядов, переопределённая система уравнений Юла-Волкера.

Введение. Проблема экспресс-контроля состояния здоровья актуальна для широкого круга людей. Например, для спортсменов, представителей ответственных профессий — пилотов, водителей, операторов сложных технических установок, а также для инвалидов, лиц с ограниченными функциональными возможностями и т.д.

Известно [1], что колебания длительности сердечного цикла, называемые также вариабельностью ритма сердца (ВСР), наблюдаемые у человека, могут служить основой неинвазивной (без исследования биологических жидкостей) экспресс-диагностики состояния процессов регуляции его основных физиологических функций. ВСР вызвана сложным взаимодействием гемодинамических, электрофизиологических и химических процессов, происходящих в организме. В норме амплитуда колебаний составляет 5...10 % от средней длительности кардиоцикла и зависит от состояния регуляторных систем организма [1, 2]. Так как сердце обладает собственным автоматизмом, то влияние вегетативной нервной системы (ВНС) на пульс является модулирующим. Сердце иннервируется как симпатическими, так и парасимпатическими отделами ВНС. Симпатические воздействия приводят к повышению скорости сердечных сокращений, а парасимпатические — к понижению [3].

В настоящее время широкое распространение получили методы спектрального и корреляционного анализа ритма сердца [1-3]. Они обладают наглядностью и компактным пространством признаков, т.к. по 2...3 компонентам спектра можно сделать заключение о балансе активности симпатического и парасимпатического отделов вегетативной нервной системы [1]. Кроме того, применение спектрального анализа позволяет количественно оценить различные частотные составляющие ВСР, а корреляционный анализ характеризует взаимосвязь ритма сердца с каким-либо другим процессом (например, изменением объема воздуха в легких) [4]. При этом возможно представление дискретных отсчётов наблюдаемых диагностических сигналов, одновременно поступающих от M датчиков, в виде единого векторного процесса для его дальнейшей обработки. Тогда результаты наблюдений за состоянием здоровья, поступающие по M различным информационным каналам, могут быть представлены в виде реализации \mathbf{X} дискретного векторного M -мерного случайного процесса:

$$\mathbf{X}=[\mathbf{x}_0; \mathbf{x}_1; \dots; \mathbf{x}_t; \dots; \mathbf{x}_{T-1}], \quad (1)$$

где $\mathbf{x}_t=[x_{0,t}; x_{1,t}; \dots; x_{m,t}; \dots; x_{M-1,t}]^T$ — M -мерный векторный t -й временной отсчёт процесса наблюдения; $x_{m,t}$ — значение t -го наблюдения в m -м канале; T — количество векторных отсчётов \mathbf{x}_t в реализации \mathbf{X} ; $m=0, 1, \dots, M-1$; $t=0, 1, \dots, T-1$.

Примерами реализаций \mathbf{X} таких процессов могут служить изменения объема воздуха в легких, сигналы капиллярного кровотока и ритма сердца, поступающие от $M \geq 2$ датчиков. Ниже рассмотрены результаты наблюдения за пациентом в случае, когда данные поступают по трём каналам $M=3$: ритма сердца, объема воздуха в легких и капиллярного кровотока. Как показали эксперименты, для выявления периодических изменений в короткой (5...10 с) реализации \mathbf{X} анализируемого процесса эффективно применение векторных авторегрессионных моделей (VAR) [5, 6].

Цель работы — построение векторных авторегрессионных моделей для задач параметрического спектрального экспресс-анализа процессов регуляции физиологических функций человека по коротким наблюдениям, одновременно поступающим от нескольких диагностических датчиков.

Постановка задачи. Математическое описание векторного процесса \mathbf{X} линейной авторегрессионной моделью предполагает, что текущий M -компонентный отсчёт \mathbf{x}_t может быть выражен через аддитивную взвешенную сумму p предыдущих отсчётов этого процесса [5]:

$$\mathbf{x}_t = \sum_{k=1}^p \mathbf{A}_k \mathbf{x}_{t-k} + \boldsymbol{\varepsilon}_t, \quad (2)$$

где $\boldsymbol{\varepsilon}_t$ — M -мерный вектор-столбец t -го векторного отсчёта $\boldsymbol{\varepsilon}_t=[\varepsilon_{0,t}; \varepsilon_{1,t}; \dots; \varepsilon_{m,t}; \dots; \varepsilon_{M-1,t}]^T$ реализации $\boldsymbol{\varepsilon}=[\boldsymbol{\varepsilon}_0; \boldsymbol{\varepsilon}_1; \dots; \boldsymbol{\varepsilon}_t; \dots; \boldsymbol{\varepsilon}_{T-p-1}]$ векторного процесса ошибки линейного предсказания; \mathbf{A}_k — $(M \times M)$ -мерная матрица k -го коэффициента линейного предсказания. Предполагается, что векторный процесс $\boldsymbol{\varepsilon}$ ошибки представляет собой некоррелированный шум [7].

Для нахождения неизвестных p $(M \times M)$ -мерных коэффициентов \mathbf{A}_k удобно представить их как единую $(M \times pM)$ -мерную матрицу $\mathbf{A}=[\mathbf{A}_1; \mathbf{A}_2; \dots; \mathbf{A}_k; \dots; \mathbf{A}_p]$. Тогда уравнение (2) можно представить в следующем виде [8]:

$$\mathbf{x}_t = \mathbf{A} \tilde{\mathbf{x}}_{t-1} + \boldsymbol{\varepsilon}_t, \quad (3)$$

где $\tilde{\mathbf{x}}_{t-1}$ — pM -мерный вектор-столбец преды-

дущих $t-p$ значений реализации \mathbf{X} , которые сгруппированы последовательно. Структура вектора $\tilde{\mathbf{x}}_{t-1}$ показана ниже:

$$\tilde{\mathbf{x}}_{t-1}^T = [\mathbf{x}_{t-1}^T; \mathbf{x}_{t-2}^T; \dots; \mathbf{x}_{t-k}^T; \dots; \mathbf{x}_{t-p}^T].$$

Для нахождения всей группы \mathbf{A} из p матричных коэффициентов \mathbf{A}_k векторной авторегрессии помножим обе части уравнения (3) справа на вектор $\tilde{\mathbf{x}}_{t-1}^{T*}$:

$$\mathbf{x}_t \tilde{\mathbf{x}}_{t-1}^{T*} = \mathbf{A} \tilde{\mathbf{x}}_{t-1} \tilde{\mathbf{x}}_{t-1}^{T*} + \boldsymbol{\varepsilon}_t \tilde{\mathbf{x}}_{t-1}^{T*}, \quad (4)$$

где $*$ — знак комплексного сопряжения. Усредняя (4) по t , получаем [8]:

$$\mathbf{A} = \mathbf{k}\mathbf{K}^{-1}, \quad (5)$$

где \mathbf{k} , \mathbf{K} — $(M \times pM)$ -мерная и $(pM \times pM)$ -мерная автокорреляционные матрицы описываемого векторного процесса (соответственно). Матрица \mathbf{K} имеет структуру теплицевой и эрмитовой блочной ленточной матрицы:

$$\mathbf{K} = \begin{bmatrix} \mathbf{K}_0 & \mathbf{K}_1 & \dots & \mathbf{K}_{p-1} \\ \mathbf{K}_1^{T*} & \mathbf{K}_0 & \dots & \vdots \\ \vdots & \dots & \ddots & \mathbf{K}_1 \\ \mathbf{K}_{p-1}^{T*} & \dots & \mathbf{K}_1^{T*} & \mathbf{K}_0 \end{bmatrix}, \quad (6)$$

где подматрицы \mathbf{K}_k , $k=0, 1, \dots, p-1$ представляют собой корреляционные матрицы k -го порядка векторного M -компонентного процесса. Подматрица \mathbf{K}_k может быть представлена как $(M \times M)$ -мерная матрица следующего вида:

$$\mathbf{K}_k = \begin{bmatrix} K_{0,0k} & \dots & K_{0,mk} & \dots & K_{0,M-1k} \\ \vdots & & \vdots & & \vdots \\ K_{m,0k} & \dots & K_{m,mk} & \dots & K_{m,M-1k} \\ \vdots & & \vdots & & \vdots \\ K_{M-1,0k} & \dots & K_{M-1,mk} & \dots & K_{M-1,M-1k} \end{bmatrix},$$

где $K_{j,mk}$ — коэффициент корреляции k -го порядка между j -м и m -м процессами, входящими в качестве компонент в рассматриваемый векторный процесс, $j, m=0, 1, \dots, M-1$. При $j=m$ величина $K_{m,mk}$ представляет собой автокорреляционный коэффициент k -го порядка m -го процесса. Аналогично структуре \mathbf{K} в (6) $(M \times pM)$ -мерная автокорреляционная матрица \mathbf{k} имеет вид:

$$\mathbf{k} = [\mathbf{K}_1 \mid \mathbf{K}_2 \mid \dots \mid \mathbf{K}_p]. \quad (7)$$

Матрицу \mathbf{P} мощностей возбуждающего шума [7] с мерностью $(M \times M)$ можно получить, решая линейную систему уравнений вида [8]:

$$\mathbf{P} = [\mathbf{I}; -\mathbf{A}] \begin{bmatrix} \mathbf{K}_0 \\ -\mathbf{k}^{T*} \end{bmatrix}. \quad (8)$$

где \mathbf{I} — $(M \times M)$ -мерная единичная матрица.

Недостатком изложенного подхода является отсутствие возможности учёта при нахождении параметров \mathbf{A} старших подматриц \mathbf{K}_k автокорреляции при $k > p$. Нарастивание лага p , т.е. порядка векторной модели, не только придаёт ей излишнюю громоздкость, но и в практических приложениях, связанных с необходимостью спектрального анализа коротких выборок экспериментальных процессов, вызывает появление ложных спектральных пиков [7]. Они возникают из-за неточных оценок подматрицы \mathbf{K}_k автокорреляции старших порядков k , что обусловлено ограниченностью длины T выборки \mathbf{X} исследуемого процесса. Вместе с тем низкие порядки p модели часто не дают возможность получить удовлетворительные спектральные оценки из-за неадекватности модели. Это обусловлено тем, что реальная связность исследуемого процесса может оказаться выше, чем p , а ещё значительные по величине коэффициенты обратной к \mathbf{K} в (8) матрицы \mathbf{K}^{-1} не участвуют в определении параметров \mathbf{A} модели.

Для преодоления данного противоречия предлагается использовать переопределённую векторную авторегрессионную модель небольшого порядка $p=1 \dots 10$, учитывающую при нахождении параметров \mathbf{A} дополнительные старшие подматрицы \mathbf{K}_k автокорреляции при $(p+P) \geq k > p$, где P — глубина переопределённости модели.

Аналитическое решение. Одновременный учёт не одной $\boldsymbol{\varepsilon}_t$, а серии из $(P+1)$ ошибок $\boldsymbol{\varepsilon}_t, \boldsymbol{\varepsilon}_{t+1}, \dots, \boldsymbol{\varepsilon}_{t+P}$ векторного линейного предсказания приводит к модификации выражения (3):

$$\begin{aligned} & [\mathbf{x}_t; \mathbf{x}_{t+1}; \dots; \mathbf{x}_{t+P}] = \\ & = \mathbf{A}_p [\tilde{\mathbf{x}}_{t-1}; \tilde{\mathbf{x}}_t; \dots; \tilde{\mathbf{x}}_{t+P-1}] + [\boldsymbol{\varepsilon}_t; \boldsymbol{\varepsilon}_{t+1}; \dots; \boldsymbol{\varepsilon}_{t+P}], \end{aligned} \quad (9)$$

где \mathbf{A}_p — матрица коэффициентов переопределённой векторной авторегрессионной модели. Для решения системы (9) из $(P+1)$ матричных линейных уравнений с глубиной переопределённости P помножим обе части (9) на матрицу $[\tilde{\mathbf{x}}_{t-1}; \tilde{\mathbf{x}}_t; \dots; \tilde{\mathbf{x}}_{t+P-1}]^{T*}$, обобщив полученное ранее выражение (4):

$$\begin{aligned} & [\mathbf{x}_t; \mathbf{x}_{t+1}; \dots; \mathbf{x}_{t+P}] [\tilde{\mathbf{x}}_{t-1}; \tilde{\mathbf{x}}_t; \dots; \tilde{\mathbf{x}}_{t+P-1}]^{T*} = \\ & = \mathbf{A}_p [\tilde{\mathbf{x}}_{t-1}; \tilde{\mathbf{x}}_t; \dots; \tilde{\mathbf{x}}_{t+P-1}] [\tilde{\mathbf{x}}_{t-1}; \tilde{\mathbf{x}}_t; \dots; \tilde{\mathbf{x}}_{t+P-1}]^{T*} + \\ & + [\boldsymbol{\varepsilon}_t; \boldsymbol{\varepsilon}_{t+1}; \dots; \boldsymbol{\varepsilon}_{t+P}] [\tilde{\mathbf{x}}_{t-1}; \tilde{\mathbf{x}}_t; \dots; \tilde{\mathbf{x}}_{t+P-1}]^{T*}. \end{aligned} \quad (10)$$

В развёрнутом виде формула (10) приобретает структуру матричного уравнения для нахождения матрицы \mathbf{A}_p коэффициентов векторной линейной переопределённой авторегрессии

с глубиной переопределённости $P > 1$:

$$\begin{aligned}
 & [\mathbf{x}_t; \mathbf{x}_{t+1}; \dots; \mathbf{x}_{t+P}] \begin{bmatrix} \mathbf{x}_{t-1}^{T*} & \mathbf{x}_{t-2}^{T*} & \dots & \mathbf{x}_{t-p}^{T*} \\ \mathbf{x}_t^{T*} & \mathbf{x}_{t-1}^{T*} & \dots & \mathbf{x}_{t-p+1}^{T*} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_{t+P-1}^{T*} & \mathbf{x}_{t+P-2}^{T*} & \dots & \mathbf{x}_{t+P-p}^{T*} \end{bmatrix} = \\
 & = \mathbf{A}_P \begin{bmatrix} \mathbf{x}_{t-1} & \mathbf{x}_t & \dots & \mathbf{x}_{t+P-1} \\ \mathbf{x}_{t-2} & \mathbf{x}_{t-1} & \dots & \mathbf{x}_{t+P-2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_{t-p} & \mathbf{x}_{t-p+1} & \dots & \mathbf{x}_{t+P-p} \end{bmatrix} \times \\
 & \times \begin{bmatrix} \mathbf{x}_{t-1}^{T*} & \mathbf{x}_{t-2}^{T*} & \dots & \mathbf{x}_{t-p}^{T*} \\ \mathbf{x}_t^{T*} & \mathbf{x}_{t-1}^{T*} & \dots & \mathbf{x}_{t-p+1}^{T*} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_{t+P-1}^{T*} & \mathbf{x}_{t+P-2}^{T*} & \dots & \mathbf{x}_{t+P-p}^{T*} \end{bmatrix} + \\
 & + [\boldsymbol{\varepsilon}_t; \boldsymbol{\varepsilon}_{t+1}; \dots; \boldsymbol{\varepsilon}_{t+P}] \times \\
 & \times \begin{bmatrix} \mathbf{x}_{t-1}^{T*} & \mathbf{x}_{t-2}^{T*} & \dots & \mathbf{x}_{t-p}^{T*} \\ \mathbf{x}_t^{T*} & \mathbf{x}_{t-1}^{T*} & \dots & \mathbf{x}_{t-p+1}^{T*} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_{t+P-1}^{T*} & \mathbf{x}_{t+P-2}^{T*} & \dots & \mathbf{x}_{t+P-p}^{T*} \end{bmatrix}. \tag{11}
 \end{aligned}$$

После усреднения по t выражение (11) приобретает вид, аналогичный (5):

$$\mathbf{k}_P = \mathbf{A}_P \mathbf{K}_P, \tag{12}$$

где

$$\mathbf{k}_P = \begin{bmatrix} \mathbf{x}_t \\ \mathbf{x}_{t+1} \\ \vdots \\ \mathbf{x}_{t+P} \end{bmatrix}^T \begin{bmatrix} \mathbf{x}_{t-1}^{T*} & \mathbf{x}_{t-2}^{T*} & \dots & \mathbf{x}_{t-p}^{T*} \\ \mathbf{x}_t^{T*} & \mathbf{x}_{t-1}^{T*} & \dots & \mathbf{x}_{t-p+1}^{T*} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_{t+P-1}^{T*} & \mathbf{x}_{t+P-2}^{T*} & \dots & \mathbf{x}_{t+P-p}^{T*} \end{bmatrix}$$

представляет собой $(M \times pM)$ -мерную автокорреляционную матрицу, а \mathbf{K}_P — квадратную $(pM \times pM)$ -мерную автокорреляционную матрицу анализируемого векторного процесса.

При получении (12) принят во внимание факт отсутствия корреляции между векторными M -мерными отсчётами $\boldsymbol{\varepsilon}t$ процесса ошибки линейного предсказания и \mathbf{x}_k анализируемого процесса [6] при любых $t, k=0, 1, \dots, T-1$:

$$\begin{bmatrix} \boldsymbol{\varepsilon}_{0,t} \\ \boldsymbol{\varepsilon}_{1,t} \\ \vdots \\ \boldsymbol{\varepsilon}_{M-1,t} \end{bmatrix} \begin{bmatrix} x_{0,k} \\ x_{1,k} \\ \vdots \\ x_{M-1,k} \end{bmatrix}^{T*} = \mathbf{0},$$

где $\mathbf{0}$ — нулевая $(M \times M)$ -мерная матрица.

Отметим, что для регуляризации задачи нахождения по (12) \mathbf{A}_P путём обращения мат-

рицы \mathbf{K}_P целесообразно добавить к элементам её главной диагонали действительный параметр $\lambda > 0$ регуляризации:

$$\mathbf{A}_P = \mathbf{k}_P (\mathbf{K}_P + \lambda \mathbf{I})^{-1}, \tag{13}$$

где \mathbf{I} — $(pM \times pM)$ -мерная единичная матрица.

Параметр λ может быть подобран экспериментально. Так, в рассмотренном ниже примере при $p=1 \dots 10, P=1 \dots 10$ и нормированных к единице дисперсиях K_{m,m_0} всех M исследуемых процессов для обеспечения удовлетворительной точности обращения матрицы \mathbf{K}_P достаточно было выполнить условие $10^{-8} < \lambda < 10^{-5}$. При этом анализировался максимальный модуль элементов $E_{m,k}, m=0, 1, \dots, M-1, k=0, 1, \dots, pM-1$ $(M \times pM)$ -мерной матрицы \mathbf{E} ошибок вычислений, полученной из решения проверочного уравнения:

$$\mathbf{E} = \mathbf{A}_P \mathbf{K}_P - \mathbf{k}_P.$$

Эксперименты показали, что величина максимального модуля ошибки оставалась в диапазоне $0 < |E_{m,k}| < 10^{-4}$ при $M=3; p=1, 2, \dots, 5; P=0, 1, \dots, 10$. Поскольку $|E_{m,k}| \ll K_{m,m_0}$, то можно сделать вывод о том, что введение переопределённости в систему матричных уравнений вида (10), описывающую модель медицинских диагностических сигналов, не приводит к существенному влиянию плохой обусловленности \mathbf{K}_P на решение (12) относительно \mathbf{A}_P . На практике сохраняется возможность получения по (13) удовлетворительной по точности оценки матрицы \mathbf{A}_P коэффициентов векторной авторегрессии.

Квадратную $(M \times M)$ -мерную матрицу \mathbf{P}_P мощностей возбуждающего шума для переопределённой модели можно получить по выражению, аналогичному (8), заменив в нём матрицу \mathbf{A} на найденные по (12) значения \mathbf{A}_P :

$$\mathbf{P}_P = [\mathbf{I}; -\mathbf{A}_P] \begin{bmatrix} \mathbf{K}_0 \\ -\mathbf{k}^{T*} \end{bmatrix}. \tag{14}$$

Отметим, что для моделирования диагностических сигналов, которые подобны приведённым в рассматриваемом ниже примере, можно использовать вместо величины \mathbf{P}_P , найденной из (14), полученную по (8) матрицу \mathbf{P} дисперсий возбуждающего шума, т.к. $\mathbf{P}_P \approx \mathbf{P}$ [8].

Экспериментальные исследования. Проанализируем предлагаемую методику спектрального анализа на примере описания экспериментальной последовательности \mathbf{X} , состоящей из трёх $(M=3)$ компонент: $x_{0,t}$ капиллярного кровотока, $x_{1,t}$ ритма сердца и $x_{2,t}$ дыхания.

Наблюдения продолжительностью 5 с представляли собой модельные последовательности X , сформированные на основе экспериментальных сигналов $x_{0,t}$ от фотоплетизмографа, $x_{1,t}$ кардиографа и $x_{2,t}$ пневмотахометра (контактного датчика дыхания). Частота дискретизации экспериментальных сигналов составляла 100 Гц.

На рисунке 1 изображен фрагмент анализируемой реализации X . Верхняя кривая $x_{0,t}$ представляет собой фрагмент фотоплетизмограммы, характеризующей интенсивность $x_{0,t}$ капиллярного кровотока; линия, расположенная в центре, отображает результат $x_{1,t}$ работы кардиографа; внизу показаны изменения $x_{2,t}$ объёма грудной клетки. Все кривые отображены в относительных единицах нормированной интенсивности.

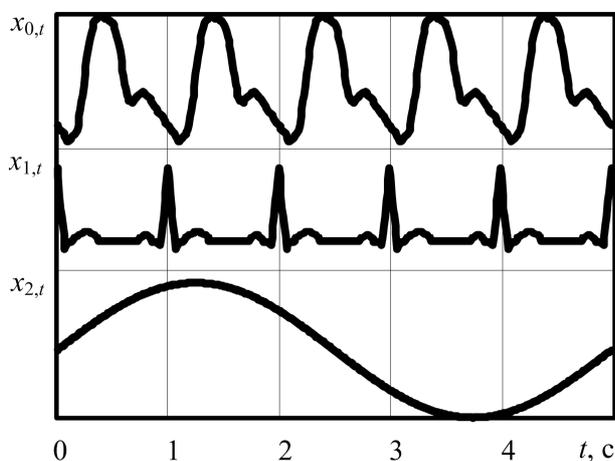


Рисунок 1

С целью подавления шумов, возникающих в процессе записи диагностических сигналов, применена их низкочастотная фильтрация равновесным фильтром скользящего-среднего двадцатого порядка. Подавление высокочастотных компонент не сказывается на качестве выделения доминантных частот F , находящихся ниже 2 Гц [1].

В рассматриваемом примере анализа спектрального состава для экспресс-диагностики состояния процессов регуляции основных физиологических функций человека достаточно выделения одной-двух доминантных компонент [1]. Поэтому целесообразно ограничиться низким порядком $p=2$ векторной авторегрессионной модели.

Результаты векторного спектрального анализа реализации X показаны на рисунке 2, где представлены оценки нормированных спектральных плотностей мощности $S_m(F)$ фотоплетизмограммы $S_0(F)$, кардиограммы $S_1(F)$ и пневмотахометра $S_2(F)$, изображённых точечной, пунктирной и штрихпунктирной линиями (соответственно). Истинные доминантные частоты F_m , заданные при формировании модельной последовательности X , составляли $F_0=F_1=1$ Гц для фотоплетизмограммы $x_{0,t}$ и кардиограммы $x_{1,t}$, а для пневмотахометра была принята истинная доминантная частота $F_2=0,2$ Гц (см. рисунок 1).

Из рисунка 2 видно, что векторная модель низкого ($p=2$) порядка не даёт возможности произвести удовлетворительную оценку доминантных частот процессов, характеризующих сердечно-сосудистую деятельность организма человека. Различия между истинными значениями частот F_0 , F_1 и их оценками составляют более 0,2 Гц, т.е. относительные отклонения превышают 20 %.

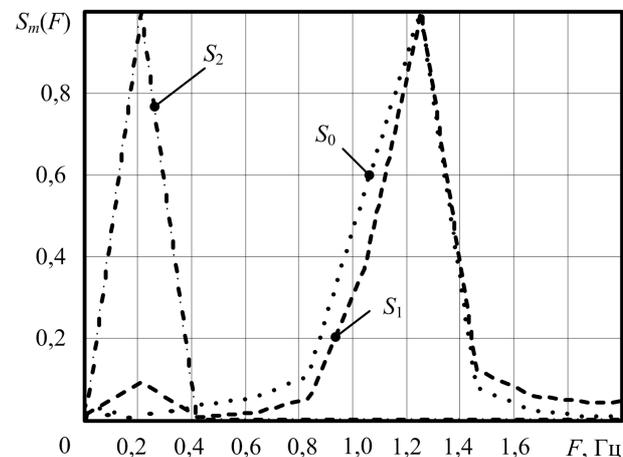


Рисунок 2

Попытка увеличить порядок p векторной модели при сохранении того же объёма статистического материала (время наблюдения 5 с) не только делает модель более громоздкой, но и не приводит к значительному повышению точности измерения частот F_0 , F_1 . Так, например, при $p=10$ относительные отклонения составляют величину 8,6 %.

Вместе с тем спектральные оценки, полученные с помощью переопределённой векторной модели низкого порядка $p=2$, имеют более высокую точность.

На рисунке 3 представлен фрагмент нормированных спектральных плотностей мощности $S_m(F)$, полученных с помощью векторной авторегрессионной модели второго порядка ($p=2$) при глубине переопределённости $P=5$. Фрагмент представляет собой диапазон частот от 0,5 до 1,5 Гц, содержащий истинную доминантную частоту $F_0=F_1=1$ Гц.

Сопоставление рисунков 2 и 3 даёт возможность качественно оценить эффективность предлагаемого метода, а для её количественного описания вводится критерий относительного отклонения ΔF_m оцененных доминирующих частот.

Сопоставление рисунков 2 и 3 даёт возможность качественно оценить эффективность предлагаемого метода, а для её количественного описания вводится критерий относительного отклонения ΔF_m оцененных доминирующих частот.

нантных частот \tilde{F}_m от истинных F_m , выраженного в процентах.

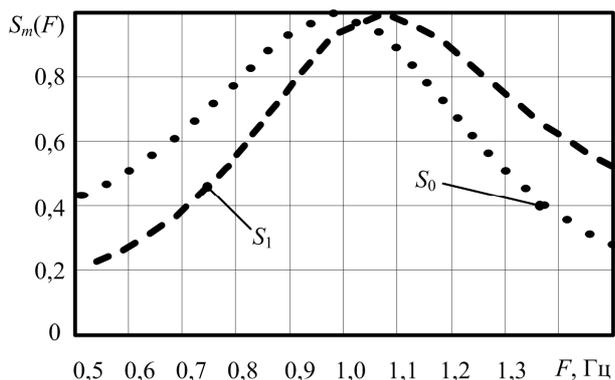


Рисунок 3

Анализ эффективности. Проведём оценку результативности предложенной методики векторного спектрального анализа по критерию:

$$\Delta F_m = \frac{|\tilde{F}_m - F_m|}{F_m} 100 \% \quad (15)$$

Зависимости ΔF_m от глубины P переопределённости для рассмотренного выше примера приведены на рисунке 4. На нём относительные отклонения ΔF_0 , ΔF_1 изображены точечной и пунктирной линиями (соответственно), а сплошной чертой — логарифмический тренд $L(P)$ функции $\overline{\Delta F}(P) = [\Delta F_0(P) + \Delta F_1(P)]/2$ усреднённого относительного отклонения оценки частоты пульса. Логарифмический тренд может быть рассчитан по формуле:

$$L(P) = 20 e^{-P/2.3} \% \quad (16)$$

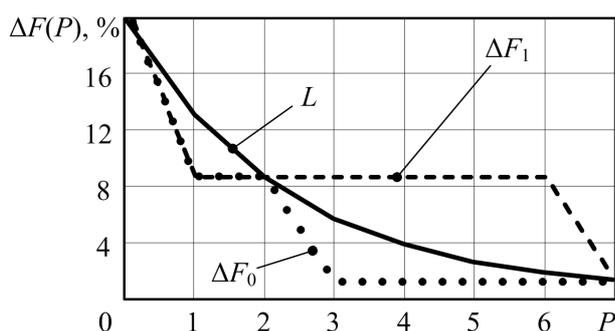


Рисунок 4

Из рисунка 4 видно, что относительные отклонения ΔF существенно (до 1,5...10 раз) уменьшаются при наращивании глубины P переопределённости векторной модели анализируемого порядка $p=2$. Тенденция по улучшению адекватности рассматриваемого параметричес-

кого метода векторного спектрального анализа сохраняется при увеличении порядка p . Так, для приведённого примера при $p=4$ усреднённое относительное отклонение $\overline{\Delta F}$ составляет 14 %, а при введении в модель того же порядка $p=4$ переопределённости глубиной $P=6$ величина $\overline{\Delta F}$ падает до 9 %.

Выводы. Таким образом, возможно по короткой (5...10 с) выборке установить характер периодических процессов, происходящих в сердечнососудистой системе человека. При этом имеется возможность повышения точности параметрических векторных спектральных оценок по критерию (15) путём введения в векторную авторегрессионную (VAR) модель заданного порядка p переопределённости достаточной глубины P . Общая тенденция уменьшения относительной ошибки $\overline{\Delta F}$ при наращивании P носит логарифмический характер (16). Выигрыши $\overline{\Delta F}(0)/\overline{\Delta F}(P)$ составляют несколько (1,5...10) раз и достигаются за счёт учёта $(P-p)$ старших подматриц \mathbf{K}_k автокорреляции анализируемого векторного процесса при нахождении матрицы \mathbf{A}_p параметров векторной авторегрессионной модели.

Библиографический список

1. Баевский Р.М., Кириллов О.И., Клецкин С.М. Математический анализ измерений сердечного ритма при стрессе.— М.: Наука, 1984.— 221 с.
2. Морозов А.А. Анализ и преобразование биомедицинских сигналов. Часть 1: учебное пособие.— М.: Изд-во МГТУ им. Н.Э.Баумана, 2000.— 24 с.
3. Рябыкина Г.В., Соболев А.В. Вариабельность ритма сердца.— М.: «Стар'Ко», 1998.— 200 с.
4. Гринченко В.Т., Грудницкий А.Г. Модель взаимодействия сердечно-сосудистой и респираторной систем // Акустичний вісник.— 2006.— Т. 9.— № 3.— С. 16-26.— (Киев, Украина).
5. Juselius K. The cointegrated VAR model. Methodology and applications.— New York: Oxford University Press Inc., 2006.— 440 p.
6. Миронов М.А., Башаев А.В., Полосин С.А. Оптимальная оценка параметров модели авторегрессии векторных гауссовских процессов по экспериментальным данным // Радиотехника.— 2002.— № 7.— С. 6-11.
7. Марпл-мл. С.Л. Цифровой спектральный анализ и его приложения: пер. с англ.— М.: Мир, 1990.— 584 с.
8. Андреев В.Г. Векторный регрессионный спектральный анализ отражений от вращающегося объекта // Вестник РГРТУ.— № 3.— 2010.— С. 43-48.

УДК 621.395

М.В. Илюшин, О.О. Басов, А.В. Радаев, А.В. Степанов

ОБОСНОВАНИЕ ОБЪЕКТИВНОГО КРИТЕРИЯ КАЧЕСТВА ЗВУЧАНИЯ СИНТЕЗИРОВАННОГО ШИРОКОПОЛОСНОГО РЕЧЕВОГО СИГНАЛА

Обоснована необходимость объективного измерения качества звучания синтезированного речевого сигнала (РС) по показателям естественности звучания речи и узнаваемости говорящего. На основе результатов экспериментальных исследований широкополосного РС (ШРС) обоснован критерий, обеспечивающий лучшую корреляцию с субъективными измерениями. Показано, что переход к кодированию ШРС обеспечивает повышение естественности звучания речи и узнаваемости говорящего.

Ключевые слова: объективные методы измерения качества звучания синтезированного речевого сигнала, широкополосный речевой сигнал, коэффициент корреляции.

Введение. Одной из важнейших задач в области кодирования РС является получение надежных оценок качества звучания синтезированного на приеме РС, позволяющих сопоставить алгоритмы кодирования [1].

К основным показателям качества звучания синтезированного РС, отражающим главные требования абонентов, относятся:

- уровень громкости;
- разборчивость;
- естественность звучания;
- узнаваемость говорящего;
- отсутствие помех [2, 3].

Существующие алгоритмы кодирования РС обеспечивают приемлемые оценки показателей разборчивости, громкости принимаемых речевых сообщений и отсутствие помех. Вместе с тем анализ результатов исследований [3, 4, 5] подтверждает тенденцию роста требований абонентов к качеству звучания синтезированного РС с точки зрения обеспечения естественности звучания речи и узнаваемости говорящего. В качестве одного из возможных направлений повышения качества РС по указанным показателям может рассматриваться переход от кодирования узкополосного РС (УРС) к передаче ШРС с диапазоном частот от 50 до 7000 Гц [6].

Цель работы – определить метод оценивания качества звучания синтезированного ШРС по показателям естественности звучания речи и узнаваемости говорящего, обеспечивающий лучшую корреляцию с субъективными оценками.

Анализ методов оценивания качества звучания синтезированного речевого сигнала.

К настоящему времени разработано большое количество методов оценивания качества звучания синтезированного УРС, однако проблема оценивания качества звучания синтезированного ШРС остается на сегодняшний день неразрешенной [3, 7].

Для оценивания качества ШРС по показателям естественности звучания речи и узнаваемости говорящего необходимо, чтобы метод оценивания отвечал следующим требованиям:

- позволял выразить качество звучания количественной мерой;
- обладал инвариантностью к типу оцениваемой системы кодирования;
- допускал измерение в процессе речевой коммуникации;
- позволял автоматизировать процесс измерения;
- учитывал свойства слухового восприятия;
- избавлял от необходимости использования экспертов.

В настоящее время выделяют физические, субъективно-статистические и объективные методы оценивания качества звучания синтезированного речевого сигнала.

Физические методы основаны на результатах измерения физических параметров исследуемого тракта, например коэффициента передачи, частотной характеристики, амплитудной характеристики. Невозможность сведения результатов измерения данных параметров к единому коли-

чественному показателю привела в свое время к необходимости разработки новых методов, получивших название субъективно-статистических [8, 9].

Данные методы предполагают привлечение экспертов, которые на основании своих слуховых ощущений выносят суждение о качестве речевого сигнала, прошедшего через испытуемый тракт. Исторически первым показателем оценивания качества звучания синтезированного РС в рамках субъективно-статистического метода была громкость. Главным недостатком данного показателя является невозможность учета других показателей качества, так как более громкая речь оказывалась иногда менее разборчивой, чем тихая [10]. В связи с этим был введен показатель разборчивости, мерой оценки качества звучания РС в котором является процент правильно принятых элементов речи из числа переданных. В зависимости от того, какие элементы речи передаются, выделяют звуковую, слоговую, словесную и фразовую разборчивости.

Метод артикуляционных испытаний, обеспечивая устойчивые количественные оценки, сопровождается следующими недостатками:

- невозможностью автоматизации процесса измерения;
- большой трудоемкостью;
- необходимостью большого объема речевого материала.

Некоторых из указанных недостатков удается избежать при использовании метода измерения эквивалента затухания по разборчивости, основанного на сопоставлении разборчивости речи на испытуемом тракте с разборчивостью речи на специальном эталонном тракте. Данный метод обеспечивает незначительные величины погрешностей измерений, однако не способен решить проблемы, связанные с невозможностью получения обобщенной характеристики качества передачи и автоматизации процесса оценивания.

Метод выбора для измерения разборчивости заключается в передаче связанных, хотя и лишенных смысла, фраз из пяти слов, первые два из которых являются вводными, а остальные три – испытательными, и сопоставлении услышанных слов со сходными по звучанию словами, представленными в специальных таблицах.

В связи с непрерывным ростом требований абонентов к качеству звучания синтезированного РС с точки зрения возможности передачи системами телефонии не только смысловой, но и эстетической информации, оценка качества, произведенная по любому отдельно взятому показателю, являлась приближенной и не

учитывала абонентскую оценку. Для устранения данного недостатка было предложено определять интегральное качество, оценка показателя которого выражалась в баллах и характеризовала общее впечатление слушателя от принимаемой речи.

Одним из методов оценки интегрального качества был рекомендован метод заданных категорий. Измерения по методу заданных категорий строятся на том, что экспертам предлагается высказать суждение о качестве звучания того или иного образца РС, подлежащего субъективной оценке, используя определенный критерий качества и соответствующую ему шкалу оценок. Обычно применяют критерии напряженности внимания и заметности искажений. Однако критерий напряженности внимания требует предоставления экспертам полностью неизвестного текста, что затрудняет его применение при моделировании. Кроме того, критериям заметности искажений и напряженности внимания присущ ряд общих недостатков, приводящих к увеличению численности бригады экспертов [11].

Разновидностью вышеназванного метода является метод мнений. В процессе измерения качества испытуемый тракт поочередно предоставляется не менее чем двенадцати парам обычных абонентов, которые ведут служебные или личные переговоры в течение трех-пяти минут. Для облегчения процесса общения абонентам рекомендуется использовать специальные таблицы. В качестве результатов испытаний могут быть получены следующие оценки качества:

- градации интегрального качества речи;
- градации качества речи по узнаваемости;
- градации качества речи по смысловой разборчивости [12, 13].

Достоинством данного метода является максимальное приближение условий испытаний к условиям реального использования тракта. К недостаткам метода относится недостаточная четкость границ оценок, обусловленная трудностями, возникающими при определении абонентами степени напряжения своего внимания и нахождении по заданной шкале той категории, которая соответствует его ощущениям.

Оценивать интегральное качество звучания синтезированного РС возможно также с помощью метода парных сравнений. Оценку качества речи испытуемого тракта формируют путем сравнения с эталонным трактом. Возможны два варианта оценки качества звучания РС: либо по критерию «лучше – хуже», либо в баллах по пятибалльной шкале. К достоинствам данного метода следует отнести способность

абонентов субъективно оценить интегральное качество передачи РС.

Стандартом оценки качества передачи речевого сигнала является оценка группой слушателей, получаемая прослушиванием в субъективных тестах, которые проводятся в соответствии с рекомендациями сектора стандартизации в области телефонии Международного союза электросвязи (МСЭ – Т) [14]. Результаты этих тестов – усредненные значения средней экспертной оценки. Шкала средней субъективной оценки MOS (Mean Opinion Score) определяет пять классов качества РС.

Кроме того, МСЭ – Т рекомендует к использованию при оценке разборчивости речи диагностический тест на рифмы DRT (Diagnostic Rhyme Test), модифицированный тест на рифмы MRT (Modified Rhyme Test), а в качестве единицы искажения – единицу ошибки квантования QDU (Quantization Distortion Unit) [10]. Однако оценки названных субъективных тестов для большинства кодеров, работающих на скоростях выше 4 кбит/с, не позволяют выделить отличия одного кодера от другого по их характеристикам.

В Российской Федерации методы измерений и нормы качества передачи речи регламентированы ГОСТ Р50840–95 [12] и Р51061–97 [13]. Данные документы устанавливают пять классов качества звучания речи на основе измерения разборчивости речи артикуляционным методом, что обусловлено высокой коррелированностью оценки качества звучания речи и степени разборчивости [15]. В [13] определен метод парных сравнений для оценивания качества передачи речевого сигнала по показателям естественности его звучания и узнаваемости говорящего.

Субъективные тесты являются дорогостоящими, время- и трудоемкими, исключают возможность автоматизации процесса измерения и требуют специальной подготовки квалифицированной группы экспертов. Поэтому исследователи непрерывно ведут поиск объективных критериев и методов оценивания качества звучания восстановленного РС, хорошо коррелирующихся с результатами субъективных тестов.

Объективные методы используют некоторые формализованные параметры, позволяющие судить о степени различия между исходным и синтезированным РС. Объективные измерения эффективны с точки зрения цены и повторяемости результатов, позволяют исследователю получить быструю и надежную оценку качества речи непосредственно в процессе разработки системы кодирования речи. Оценка меры иска-

жения РС при этом может проводиться как во временной, так и в частотной областях [16].

Наиболее общим и простым измерением искажений во временной области является среднеквадратичная ошибка (СКО), определяемая как

$$\text{СКО} = \frac{1}{N} \sum_{i=1}^N (x(i) - y(i))^2, \quad (1)$$

где x – вектор входных значений; y – вектор оценок.

В выражении (1) допускается, что искажения, вносимые каждым элементом вектора x , имеют равный вес. В общем случае для отражения вклада отдельных элементов в искажение вводятся неравные веса в виде взвешивающей матрицы. Указанный метод позволяет лишь сравнить форму огибающих исходного и синтезированного РС.

Поэтому для количественной оценки качества звучания синтезированного РС во временной области чаще используют критерий отношения сигнал-шум (ОСШ):

$$\text{ОСШ} = 10 \lg \frac{\sum_{i=1}^N x^2(i)}{\sum_{i=1}^N (x(i) - y(i))^2}, \quad (2)$$

где $x(i)$ и $y(i)$ – i -е отсчеты исходного и синтезированного РС; N – общее количество отсчетов в исследуемом речевом фрагменте.

ОСШ учитывает общие мощности сигнала и шума на всей длительности испытательного сигнала. При исследовании некоторых речевых кодеков большое значение имеют кратковременные отношения сигнал-шум, вычисленные на коротких сегментах РС. Таким образом учитывается сегментный характер слухового восприятия элементов речи. Критерий сегментного отношения сигнал-шум (ОСШ_{сег}) имеет вид:

$$\text{ОСШ}_{\text{сег}} = \frac{10}{M} \sum_{m=0}^{M-1} \lg \sum_{i=N_m}^{N_m+N-1} \left(\frac{x^2(i)}{(x(i) - y(i))^2} \right), \quad (3)$$

где $x(i)$ и $y(i)$ – вычисленные на m -м сегменте i -е отсчеты исходного и синтезированного речевых сигналов; N – длина сегмента и M – число сегментов в речевом фрагменте.

Учитывая результаты исследований [3, 8], можно выделить ОСШ_{сег} как более точную меру искажений по сравнению с СКО и ОСШ. Однако приведенные критерии объективного метода оценивания в основном отражают степень зашумленности РС и показывают слабую корреляцию с результатами субъективных тестов [7, 8]. К тому же, если качество кодеров формы речевой волны может быть оценено по степени соответ-

ствия формы огибающей восстановленного РС исходному с помощью названных критериев, то для алгоритмов низкоскоростного параметрического сжатия точное восстановление формы сигнала является сложной задачей, и поэтому методы оценивания качества звучания синтезированного речевого сигнала во временной области мало применимы [17]. В дополнение можно отметить, что в случае применения низкоскоростных систем кодирования РС на основе линейного предсказания слуховой аппарат человека более чувствителен к возникающим при этом частотным искажениям, нежели к амплитудным и фазовым [3, 7, 17].

Для того чтобы оценка качества звучания РС отражала критерии слухового восприятия, принципы ее формирования должны быть основаны на анализе спектрально-корреляционных характеристик РС. Критерий оценки качества звучания синтезированного РС в частотной области, предложенный в [8], учитывает сегментно-спектральный характер слухового восприятия речевого сигнала и имеет вид

$$L = \frac{1}{N_c} \sum_{i=1}^{N_c} \lambda(i)l(i), \quad (4)$$

где $\lambda(i)$ – мера спектральной неравномерности РС на i -м сегменте анализа; $l(i)$ – мера субполосных искажений РС на i -м сегменте анализа; N_c – число сегментов в речевом фрагменте. К недостаткам метода следует отнести значительные вычислительные затраты и требуемый большой объем речевого материала.

Перспективными с точки зрения оперативной оценки качества звучания синтезированного РС в процессе разработки систем кодирования речевой информации являются меры, измеряющие различие между спектрами исходного $P(\omega)$ и синтезированного $\hat{P}(\omega)$ РС. Наиболее популярные меры искажений получили название расстояние Итакуры-Саито ISD (Itakura-Saito distance) и кепстральное расстояние CD (Spectral distance) [7, 17, 18]. Мера ISD имеет вид:

$$\text{ISD} = \frac{1}{\pi} \int_0^{\pi} \left(\frac{P(\omega)}{\hat{P}(\omega)} - \log \frac{P(\omega)}{\hat{P}(\omega)} - 1 \right) d\omega \quad (5)$$

Кепстральное расстояние CD определяет различие между спектральными огибающими исходного и синтезированного сигнала, основано на кепстральных коэффициентах, рассчитываемых по параметрам линейного предсказания, и имеет вид

$$\text{CD} = \frac{1}{M} \sum_{m=1}^M \sum_{n=1}^Q (C_n - C'_n)^2, \quad (6)$$

где C_n и C'_n – кепстры сегментов исходного и синтезированного речевых сигналов; M – число сегментов в речевом фрагменте; Q – число кепстральных коэффициентов.

Главным недостатком указанных методов является зависимость значений коэффициентов корреляции между оценками ISD, CD и MOS от типа и скорости передачи системы кодирования.

В [3, 10, 19] авторы предлагают использовать показатель разборчивости речи в качестве основного критерия объективной оценки качества передачи речевого сигнала.

Приведенные в [10] методы объективного оценивания разборчивости речи основаны на предположении, что результирующая разборчивость речи определяется суммой вкладов в отдельных частотных полосах, а величина вклада зависит от ОСШ в каждой полосе.

Первый метод, основанный на расчете индекса артикуляции AI (Articulation Index), подразумевает, что ширина полос, на которые разделяется весь частотный диапазон речевого сигнала, выбирается таким образом, чтобы вклад каждой полосы в результирующую разборчивость был одинаковым. Данный метод в работах [11, 19] получил название «метода формантной разборчивости». В этих же источниках представлены результаты исследований и на их основании построены взаимозависимости между различными видами разборчивости для русского языка.

Метод определения индекса разборчивости речи SII (Speech Intelligibility Index), предложенный американскими учеными как дальнейшее развитие метода AI, предполагает различное число и размер частотных полос, в которых определяется ОСШ для вычисления суммарного коэффициента SII:

- критические полосы (21 полоса);
- третьоктавные полосы (18 полос);
- равные по вкладу критические полосы (17 полос);
- октавные полосы (6 полос).

Данный метод обеспечивает несколько лучшую корреляцию с оценками, полученными в субъективно-статистических испытаниях по сравнению с методом AI, что обусловлено учетом эффектов слуховой перцепции аудиторной системой человека [20].

Перспективными являются психоакустические методы и критерии оценивания качества звучания синтезированного РС, основанные на моделях слуха человека. Данные объективные методы направлены на имитацию с помощью технических и программных средств алгоритма слуховой перцепции человека. Большинство

психоакустических методов для оценки качества звучания РС основаны на сравнении исходного (неискаженного) речевого сигнала и искаженного выходного сигнала системы кодирования. Мерой оценки качества является разница между «внутренними представлениями» обоих сигналов, формируемыми с помощью функций от параметров восприятия [7, 9, 10].

Первый объективный критерий оценивания качества звучания РС [16], основанный на учете особенностей восприятия речи человеком, получил название искажение спектра барков BSD (Bark Spectral Distortion) и определялся соотношением

$$\text{BSD} = \sum_{i=1}^N (S_x^{(k)}(i) - S_y^{(k)}(i))^2, \quad (7)$$

где $S_x^{(k)}$ и $S_y^{(k)}$ – спектры барков k -го сегмента исходного и синтезированного РС в i -й критической полосе; N – общее количество критических полос.

Критерий BSD обеспечивает лучшую корреляцию с субъективно-статистическими испытаниями по сравнению с рассмотренными выше объективными методами. Этот факт объясняется тем, что аудиторная система человека работает в неравномерной частотной шкале, получившей название шкала барков [21]. Мера BSD позволяет учитывать эффекты слухового восприятия в критических полосах.

С целью отражения такого важного свойства слуха, как маскирование, критерий BSD был дополнен и представлен в виде модифицированного искажения спектра барков MBSD (Modified Bark Spectral Distortion):

$$\text{MBSD} = \frac{1}{N} \sum_{n=1}^N \sum_{i=1}^K M(n, i) D(n, i), \quad (8)$$

где $M(n, i)$ и $D(n, i)$ – i -е параметры искажений уровня ощущения и разницы уровней громкости n -го сегмента анализа; N – число сегментов в речевом фрагменте; K – общее количество критических полос [16].

В рамках МСЭ – Т были разработаны рекомендации Р.861 [22], регламентирующие вопросы объективного анализа речевых кодеков с учетом свойств слуха. В рекомендациях был утвержден алгоритм PSQM (Perceptual Speech Quality Measure) для измерения качества узкополосных (300-3400 Гц) кодеров речи. С учетом его недостатки, обусловленных невозможностью анализа сигналов с переменной задержкой и кратковременных искажений, в основе рекомендаций Р.862 [23] был принят алгоритм PESQ (Perceptual Evaluation of Speech Quality), исполь-

зующий модель аудиторной системы человека и ориентированный на анализ РС в сетях пакетной коммутации. Основными недостатками PESQ являются сложность реализации модели аудиторной системы человека и отсутствие корреляции с результатами MOS для русского языка.

Проанализировав многообразие методов и критериев оценивания качества звучания РС, можно сделать следующие выводы.

1. Состоятельность оценки качества речи в значительной степени зависит от выбора критерия оценки.

2. Превалирование интересов абонентов среди факторов, определяющих тенденции развития речевых служб в системе связи, делает более актуальными методы субъективной оценки, характерной особенностью которых является участие в них бригад дикторов и аудиторов с определением последними качества звучания речи. При этом качество речи определяется величиной, характеризующей субъективную оценку звучания речи на выходе исследуемого тракта.

3. Субъективные тесты являются дорогостоящими, время- и трудоемкими, исключают возможность автоматизации процесса измерения и требуют специальной подготовки квалифицированной группы экспертов, поэтому исследователи непрерывно ведут поиск объективных критериев и методов оценивания качества звучания синтезированного РС, хорошо коррелирующихся с результатами субъективных тестов.

4. Внедрение новых технологий обработки речи и возросшие требования абонентов к качеству звучания синтезированного РС привели к расширению перечня показателей качества речи: наравне с традиционными (громкость и разборчивость) все более активное применение находят естественность звучания голоса и узнаваемость говорящего.

5. Для оценивания качества звучания РС необходимы методы, позволяющие получить обобщенную интегральную оценку качества речи с учетом ее громкости, разборчивости, естественности звучания и узнаваемости говорящего.

6. Психоакустические методы позволяют получить устойчивые оценки качества передачи речевого сигнала, но характеризуются сложностью реализации модели аудиторной системы человека.

7. Корректность критериев объективного и психоакустического методов оценки качества звучания синтезированного РС характеризуется корреляцией оценок, вычисленных с их использованием, с оценками, полученными субъективными методами.

Таким образом, работоспособность метода и

критерия объективного оценивания качества звучания ШРС, ориентированных на восприятие аудиторной системой человека, будет определяться их способностью быстро и надежно предоставлять оценку, максимально предсказывающую величину общепринятой субъективной оценки качества.

Экспериментальные исследования. Для обоснования критерия качества звучания синтезированного ШРС были проведены экспериментальные исследования, включающие три этапа.

На первом этапе были получены средние субъективные оценки качества звучания синтезированного РС по показателям естественности его звучания и узнаваемости говорящего для пяти низкоскоростных кодеков (таблица 1).

Таблица 1

Стандарт	Скорость кодирования, кбит/с	Оценка (ГОСТ 51061-97)	Оценка MOS	Оценка MBSD
G.723.1	5,3	4,11	3,66	$0,219 \times 10^{-3}$
G.723.1	6,3	4,18	3,91	$0,258 \times 10^{-3}$
G.726	16	3,62	2,73	$0,289 \times 10^{-3}$
G.728	16	4,34	3,75	$0,162 \times 10^{-3}$
G.729a	8	4,22	3,98	$0,213 \times 10^{-3}$

Измерения качества речи проводились с помощью метода парных сравнений с контрольным трактом согласно Р51061-97 [13].

На втором этапе с помощью компьютерного моделирования были получены средние объективные оценки качества звучания синтезированного РС для указанных низкоскоростных кодеков и набора сбалансированных речевых фраз, произносимых различными дикторами. Определялись следующие меры искажений: ОСШ_{сер}, SSD, ISD, CD и MBSD.

На третьем этапе вычислялись коэффициенты корреляции между объективными и субъективными оценками для выбранных мер по формуле [8]

$$r = \frac{\sum_{i=1}^n (X(i) - \bar{X})(Y(i) - \bar{Y})}{\sqrt{\sum_{i=1}^n (X(i) - \bar{X})^2 \cdot \sum_{i=1}^n (Y(i) - \bar{Y})^2}}, \quad (9)$$

где $X(i)$ и $Y(i)$ – объективный и субъективный показатели качества передачи i -го кодека; \bar{X} и \bar{Y} – средние значения объективного и субъективного показателей качества на множестве объектов исследования; n – число кодеков.

В таблице 2 представлены значения коэффициентов корреляции между объективными оценками и оценками, полученными с помощью метода парных сравнений с контрольным

трактом [13]. Коэффициент корреляции между субъективными оценками, полученными согласно [13], и субъективными оценками MOS составил $r = 0,94$, что обусловлено высокой коррелированностью оценки качества звучания речи по показателям естественности и узнаваемости говорящего и оценками MOS.

Таблица 2

Критерий объективного оценивания качества звучания синтезированного речевого сигнала	Коэффициент корреляции, r
ОСШ _{сер}	0,74
SSD	0,68
ISD	0,61
CD	0,63
MBSD	0,83

Учитывая полученные результаты (таблицы 1, 2), можно выделить критерий MBSD как меру искажений, позволяющую получить объективные оценки качества звучания синтезированного РС по показателям естественности звучания речи и узнаваемости говорящего.

Как следует из таблицы 3, переход от кодирования узкополосного РС к ШРС обеспечивает повышение качества синтезированного РС по показателям естественности его звучания и узнаваемости говорящего.

Таблица 3

Стандарт	Скорость кодирования, кбит/с	Оценка MOS	Оценка MBSD
G.728	16	3,75	$0,162 \times 10^{-3}$
G.722.2	15,85	4,0	$0,048 \times 10^{-3}$

Прослушивание в субъективных тестах сигналов, синтезированных указанными кодерами, подтвердило преимущество кодера ШРС перед кодером УРС по показателям естественности звучания речи и узнаваемости говорящего.

Выводы. Таким образом, в работе обоснован критерий MBSD, позволяющий объективно оценить качество звучания синтезированного ШРС по показателям естественности звучания речи и узнаваемости говорящего.

Существующие алгоритмы кодирования ШРС обеспечивают разборчивость речи примерно 92 %. При такой и более высокой разборчивости речи переговоры ведутся без переопросов. Понимание речи не требует от абонента какого-либо напряжения.

Обоснованный критерий обеспечивает наилучшую, по сравнению с известными подходами к оцениванию качества речи, корреляцию с субъективно-статистическими методами. Критерий MBSD позволяет получить интегральную

оценку качества звучания синтезированного ШРС, т.е. учесть все аспекты снижения качества – от уменьшения разборчивости до ухудшения естественности речи и узнаваемости абонента.

Библиографический список

1. *Быков, С.В.* Цифровая телефония / В.И. Журавлев, И.А. Шалимов учеб. пособие для вузов. – М.: Радио и связь, 2003. – 144 с.
2. *Велян, Г.В.* Передача речи по сетям электро-связи. – М.: Радио и связь, 1985.
3. *Соболев, В.Н.* Информационные технологии в синтетической телефонии. монография – М.: ИРИАС, 2007. – 360 с.
4. *Беллами, Д.* Цифровая телефония : пер. с англ. / Под ред. А.Н. Берлина, Ю.Н. Чернышова. – М.: Эко – Трендз, 2004. – 640 с. : ил.
5. *Шалимов, И.А.* Аналитический обзор систем кодирования речи // Телекоммуникации. 2006. №2. – С. 7–15.
6. *Рыболовлев, А.А.* Объективные основы повышения естественности (натуральности) синтезированной речи при расширении полосы частот речевого сигнала до диапазона 50–7000 Гц / М.В. Илюшин. Доклады 11-й международной конференции «Цифровая обработка сигналов и ее применения», Москва, 2009. – С. 234–237.
7. *Шалимов, И.А.* Практикум по цифровой телефонии: учеб. пособие / Академия ФСБ России, 2008. – 344 с. 111 ил., 139 табл.
8. *Ситняковский, И.В.* Цифровая сельская связь / И.В. Ситняковский, В.И. Мейкшан, под ред. М.Д. Венедиктова. – М. : Радио и связь, 1994. – 248 с.
9. *Попов, О.Б.* Цифровая обработка сигналов в трактах звукового вещания / С.Г. Рихтер: учеб. пособие для вузов. – М.: Горячая линия – Телеком, 2007. – 341 с.
10. *Шелухин, О.И.* Цифровая обработка и передача речи / Н. Ф. Лукьянцев; под ред. О.И. Шелухина. – М. : Радио и связь, 2000. – 456 с.: ил.
11. *Сапожков, М.А.* Вокодерная связь / В.Г. Михайлов. – М. : Радио и связь, 1983. – 248 с.
12. ГОСТ Р 50840-95. Передача речи по трактам связи. Методы оценки качества, разборчивости и узнаваемости. – М.: ИПК Издательство стандартов, 1996. – 230 с.
13. ГОСТ Р 51061-97. Системы низкоскоростной передачи речи по цифровым каналам. Параметры качества речи и методы измерений. – М.: Госстандарт России, 1997. – 230 с.
14. ITU-T Recommendation P.800. Method for the subjective assessment of transmission quality, 1996.
15. *Михайлов, В. Г.* Измерение параметров речи / Л. В. Златоустова. – М.: Радио и связь, 1987. – 168 с.
16. *W. Yang.* Enhanced Modified Bark Spectral Distortion (EMBSD): An Objective Speech Quality Measure Based On Audible Distortion And Cognition Model / A Dissertation of the Requirement for the Degree Doctor of Philosophy – May, 1999.
17. *Поляков, А.Н.* Об одном из способов решения задачи определения оптимальных управляющих параметров системы низкоскоростной компрессии речевой информации // Телекоммуникации. 2008. №3. – С. 15–18.
18. *H. Ozer et al.* / Digital Signal Processing. 16 (2006) P. 389–401.
19. *Калинцев, Ю.К.* Разборчивость речи в цифровых вокодерах. – М.: Радио и связь, 1991. – 220 с.
20. *Алдошина, И.А.* Основы психоакустики (подборка статей с сайта <http://www.625-net.ru>).
21. *Радзишевский, А.Ю.* Основы аналогового и цифрового звука. - М. : Издательский дом «Вильямс», 2006. – 288 с. : ил.
22. ITU-T Recommendation P.861. Objective Quality measurement of telephone-band (300-3400 Hz) speech codecs, 1996.
23. ITU-T Recommendation P.862. PESQ an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs, 2001.