

ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ
Государственное образовательное учреждение
высшего профессионального образования
УЛЬЯНОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

ОСНОВЫ ТЕОРИИ ОБРАБОТКИ ИЗОБРАЖЕНИЙ

Сборник лабораторных работ
для студентов специальности “Прикладная математика”

Составитель В. Р. Крашенинников

Ульяновск 2005

ББК 32.84 я 7
УДК 621.391 (076)
О–75

Рецензент заведующий кафедрой математической кибернетики и информатики
Ульяновского государственного университета
доктор техн. наук, профессор И. В. Семушин

Основы теории обработки изображений: сборник лабораторных работ/
О–75 Составитель В. Р. Крашенинников.–Ульяновск: УлГТУ, 2004.–35 с.

Сборник лабораторных работ составлен в соответствии с программой курса «Основы теории обработки изображений» и предназначен для студентов дневной формы обучения специальности «Прикладная математика» специализации «Обработка изображений» экономико-математического факультета. Сборник охватывает методы описания и имитации изображений, прогноз, обнаружение объектов. Большое внимание уделено адаптивным алгоритмам. Сборник можно рассматривать как приложение к учебному пособию «Основы теории обработки изображений» В. Р. Крашенинникова.

Сборник подготовлен на кафедре САПР.

Одобрено секцией методических пособий
научно-методического совета университета

ББК 32.84 я 7
УДК 621.391 (076)

© В. Р. Крашенинников, составление, 2005
© Оформление. УлГТУ, 2005

СОДЕРЖАНИЕ

Список сокращений	3
Введение	4
1. Порядок выполнения лабораторных работ	4
2. Техника безопасности при выполнении работ	5
3. Генератор стандартных гауссовских случайных величин	5
4. Лабораторная работа № 1 «Имитация изображений и оценка их параметров»	5
4. Лабораторная работа № 2 «Оценка изображений»	12
5. Лабораторная работа № 3 «Обнаружение сигналов на фоне изображений»	22
6. Лабораторная работа № 4 «Адаптивное обнаружение сигналов на фоне изображений»	28
Библиографический список	35

Список сокращений

- И – изображение
- КФ – ковариационная функция
- НКФ – нормированная ковариационная функция,
то есть корреляционная функция
- ОП – отношение правдоподобия
- ПГ – псевдоградиент, псевдоградиентный
- ПК – персональный компьютер
- ПРВ – плотность распределения вероятностей
- ПСТ – поле случайных точек
- СВ – случайная величина
- СКО – среднеквадратическое отклонение
- СП – случайное поле
- ФП – функция правдоподобия
- ФР – функция распределения (вероятностей)

ВВЕДЕНИЕ

Решение многих проблем науки и практики приводит к необходимости извлечения полезной информации из различного рода многомерных данных, которые, по аналогии с оптическими изображениями, называются *многомерными изображениями*, или просто *изображениями* (И). Такие задачи возникают в очень многих областях знаний: в медицине, радио-, тепло- и гидролокации, исследовании Космоса и Земли, телевидении и т. д. Например, диагностика различных заболеваний по И внутренних органов человека, обнаружение лесных пожаров, поиск перспективных для ловли рыбы акваторий, оценка экологического состояния регионов, навигационные задачи и т. д. При решении задач, связанных с речевыми и другими одномерными сигналами, эти сигналы бывает полезно преобразовывать в И.

Характерно, что эти задачи приходится решать при наличии различного рода мешающих факторов – помех, мешающих И, переменчивости условий наблюдения, динамики наблюдаемого объекта, взаимного перемещения приемника и объекта и т. п. Объем исходных данных обычно очень велик (глобальный мониторинг Земли, массовые медицинские обследования), они поступают с большой скоростью и требуют обработки в режиме реального времени. Оператор не в состоянии справиться с таким потоком информации.

Единственным выходом из такой ситуации является компьютерная обработка И. Для этого необходимо создание соответствующих математических методов описания и обработки И, а также программного обеспечения применительно к конкретным задачам.

Данный сборник содержит лабораторные работы, предусмотренные в курсе «Основы теории обработки изображений». В разделах «Теоретическая часть» к каждой работе дается краткая сводка теоретических сведений. Для более подробного изучения теории следует обратиться к учебному пособию «Основы теории обработки изображений» [1].

1. ПОРЯДОК ВЫПОЛНЕНИЯ ЛАБОРАТОРНЫХ РАБОТ

1. Ознакомиться с этими методическими указаниями. Для более подробного ознакомления с материалом следует обратиться к учебному пособию [1].
2. Получить у преподавателя задание на лабораторную работу.
3. Составить алгоритм имитации или обработки И.
4. Составить программу, реализующую алгоритм и включающую в себя визуализацию И. К языку программирования никаких требований не предъявляется, студент может использовать тот язык, которым он владеет, и который обеспечивает выполнение работы.
4. Отладить программу и получить требуемые результаты.

2. ТЕХНИКА БЕЗОПАСНОСТИ ПРИ ВЫПОЛНЕНИИ ЛАБОРАТОРНЫХ РАБОТ

Лабораторные работы выполняются на ПК, поэтому должны выполняться все правила техники безопасности обращения с электроприборами.

Запрещается:

1. Выполнять работы без преподавателя или лаборанта.
2. Прикасаться к оголенным проводам.
3. Производить перекоммутацию ПК.
4. Оставлять без присмотра включенные ПК.
4. Загромождать рабочее место одеждой и другими предметами, не относящимися к выполняемой работе.

Запрещается также изменять структуру директорий и удалять файлы, кроме собственных, без разрешения преподавателя.

3. ГЕНЕРАТОР СТАНДАРТНЫХ ГАУССОВСКИХ СЛУЧАЙНЫХ ВЕЛИЧИН

При выполнении лабораторных работ данного сборника требуется разыгрывать (генерировать) нормальные СВ со стандартным распределением вероятностей, то есть с нулевым средним значением и единичной дисперсией. Достаточно хорошее приближение требуемого распределения можно получить из датчика $RAND()$ равномерно распределенных на отрезке $[0;1]$ случайных чисел, который обычно входит в математическое обеспечение как стандартная функция. Стандартные нормальные СВ можно получать по формуле

$$\xi = \sum_{m=1}^{12} RAND() - 6,$$

то есть из суммы двенадцати случайных равномерных величин вычитается шесть. В силу центральной предельной теоремы получается СВ, близкая к стандартной нормальной.

4. ЛАБОРАТОРНАЯ РАБОТА № 1 «ИМИТАЦИЯ ИЗОБРАЖЕНИЙ И ОЦЕНКА ИХ ПАРАМЕТРОВ»

4.1. Цель работы

Изучение методов имитации И и оценки их параметров.

4.2. Теоретическая часть

4.2.1. Случайные поля

Наиболее распространенными в настоящее время являются информационные комплексы, включающие в себя пространственные системы датчиков и цифровую вычислительную технику. Поэтому мы будем рассматривать И с

дискретными пространственными и временными переменными. Не ограничивая общности, можно считать, что И заданы на многомерных прямоугольных сетках с единичным шагом. На рис. 4.1.а и 4.1.б изображены двумерная и трехмерная сетки. В общем случае И задано в узлах n -мерной сетки $\Omega = \{\bar{j} = (j_1, \dots, j_n) : j_m = \overline{1, M_m}, m = \overline{1, n}\}$. Отдельные элементы И называются **пикселями**. В данном сборнике лабораторных работ размерность И будет не более трех.

Если данные представляют собой временную последовательность И, то иногда удобно считать эту последовательность одним И, увеличив размерность сетки на единицу. Например, последовательность из плоских И (рис. 4.1.а) можно рассматривать как одно трехмерное И (рис. 4.1.б).

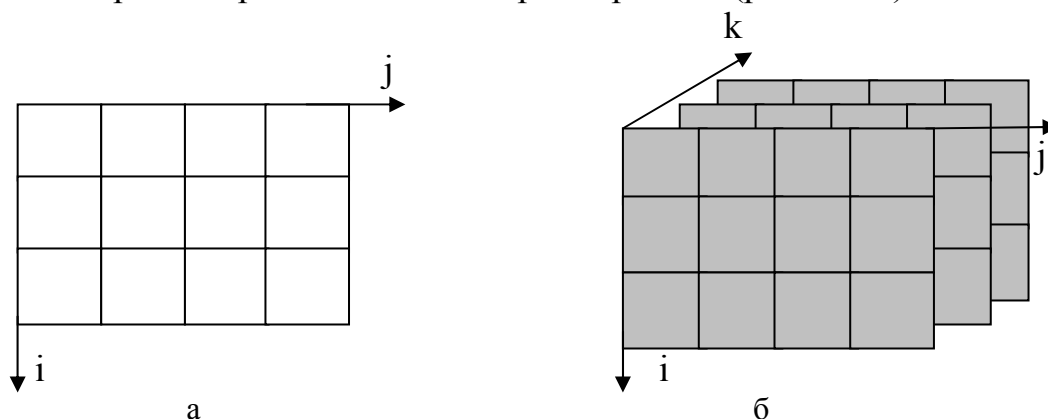


Рис. 4.1.

Если требуется временную переменную выделить особо, то будем ее записывать сверху: $X = \{x_j^k : \bar{j} \in \Omega, k \in K\}$. Это И задано на прямом произведении $\Omega \times K$ сеток Ω и K , где K – множество значений временного индекса. Сечение $x^k = \{x_j^k : \bar{j} \in \Omega\}$, то есть совокупность отсчетов И при фиксированном значении временного индекса k , называется **k -м кадром И**. Каждый кадр задан на сетке Ω . Например, на рис. 4.1.б изображено три двумерных кадра.

Таким образом, И можно рассматривать как функцию, определенную на многомерной сетке. Значение элементов И невозможно точно предсказать заранее, поэтому естественно рассматривать эти значения как СВ, применяя аппарат теории вероятностей и математической статистики. **Итак, приходим к основной модели И – системе СВ, заданных на многомерной сетке. Такие системы называются дискретными СП, или случайными функциями нескольких переменных.**

4.2.2. Авторегрессионные модели случайных последовательностей

Рассмотрим простейшую **линейную** авторегрессионную модель

$$x_1 = \sigma \xi_1, \quad x_i = \rho_1 x_{i-1} + \sigma \sqrt{1 - \rho_1^2} \xi_i, \quad i = 2, 3, \dots, M_1, \quad (4.1)$$

где ξ_i – независимые гауссовские стандартные СВ. Эта модель порождает **гауссовскую стационарную марковскую последовательность** с нулевым средним, дисперсией σ^2 и КФ $V_x(a) = M[x_i x_{i+a}] = \sigma^2 \rho_1^{|a|}$. Параметр ρ_1 равен коэффициенту корреляции между соседними элементами порождаемого процесса, а σ равно СКО процесса. При ρ_1 , близких к единице, процесс становится более гладким; при малых ρ_1 , напротив, значения процесса слабо зависимы между собой; при отрицательных ρ_1 корреляция между соседними значениями процесса отрицательна, поэтому он часто меняет знак.

Более сложные авторегрессионные модели

$$x_i = \sum_{j=1}^m \rho_j x_{i-j} + \beta \xi_i, \quad i = m+1, m+2, \dots, \quad (4.2)$$

с соответствующими условиями на ПРВ значений первых m членов x_1, x_2, \dots, x_m определяют **марковские последовательности m -го порядка**.

4.2.3. Линейные авторегрессионные модели изображений Хабиби

Рассмотрим модель Хабиби двумерного изображения размеров $M_1 \times M_2$:

$$\begin{aligned} x_{11} &= \sigma \xi_{11}, & x_{i1} &= \rho_1 x_{i-1,1} + \sigma \sqrt{(1-\rho_1^2)} \xi_{i1}, \quad i = 2, 3, \dots, M_1, \\ x_{1j} &= \rho_2 x_{1,j-1} + \sigma \sqrt{(1-\rho_2^2)} \xi_{1,j}, \quad j = 2, 3, \dots, M_2, \\ x_{ij} &= \rho_1 x_{i-1,j} + \rho_2 x_{i,j-1} - \rho_1 \rho_2 x_{i-1,j-1} + \sigma \sqrt{(1-\rho_1^2)(1-\rho_2^2)} \xi_{i,j}. \end{aligned} \quad (4.3)$$

С помощью первых двух уравнений формируется первая строка И, точно так же, как и в (4.1). Третье уравнение дает первый элемент очередной строки, остальные элементы которой формируются слева направо с помощью четвертого уравнения.

Существует и другой способ построения такого же И. Первая строка $x_{11}, x_{12}, \dots, x_{1M_1}$ формируется тем же способом с помощью первых двух уравнений модели (4.3). Пусть уже сформирована i -я строка $x_{i1}, x_{i2}, \dots, x_{iM_2}$. Построим точно так же с помощью первых двух уравнений модели (4.3) вспомогательную строку y_1, y_2, \dots, y_{M_2} . Тогда очередная $(i+1)$ -я строка И определяется по формуле

$$x_{i+1,j} = \rho_2 x_{ij} + \sqrt{(1-\rho_2^2)} y_j, \quad j = 1, 2, \dots, M_2. \quad (4.4)$$

Порождаемое поле имеет **факторизуемую экспоненциальную** КФ:

$$V_x(a, b) = M[x_{ij} x_{i+a, j+b}] = \sigma^2 \rho_1^{|a|} \rho_2^{|b|}. \quad (4.5)$$

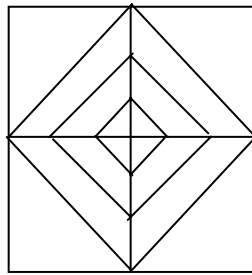
Аналогичным образом можно строить И большей размерности. Например, трехмерное изображение $X = \{x_{ijk}\}$ размеров $M_1 \times M_2 \times M_3$ можно построить следующим образом. Формируем сначала первый двумерный «слой» $X_1 = \{x_{ij1}\}$ размеров $M_1 \times M_2$ ранее описанным способом. Пусть уже построен k -й слой $X_k = \{x_{ijk}\}$. Сформируем вспомогательный двумерный слой $Y = \{y_{ij}\}$ тех же размеров и тем же способом. Тогда очередной $(k+1)$ -й слой $X_{k+1} = \{x_{ij,k+1}\}$ определяется по формуле

$$x_{ij,k+1} = \rho_3 x_{ijk} + \sqrt{(1-\rho_3^2)} y_{ij}, \quad i=1,2,\dots,M_1, \quad j=1,2,\dots,M_2. \quad (4.6)$$

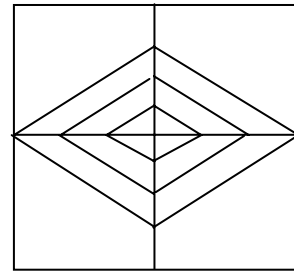
Порождаемое И имеет **множительную экспоненциальную** КФ:

$$V_x(a,b,c) = M \left[x_{ijk} x_{i+a, j+b, k+c} \right] = \sigma^2 \rho_1^{|a|} \rho_2^{|b|} \rho_3^{|c|}. \quad (4.7)$$

Это трехмерное И можно рассматривать и как последовательность плоских И (кадров), где k – номер кадра (рис. 4.1).



$$\rho_1 = \rho_2 = 0,9$$



$$\rho_1 = 0,8, \rho_2 = 0,9$$

Рис. 4.2.

Факторизуемость КФ (4.5) и (4.7) приводит к анизотропии И – элементы И, одинаково коррелированные с элементом x_{ab} , расположены на ромбе с центром в (a,b) , а не на окружности (рис. 4.2). Поэтому корреляция убывает медленнее вдоль координатных осей, чем по диагональным направлениям. В результате плоским И свойственны полосы вдоль осей, и И напоминают клетчатую ткань. Для получения И с другими КФ нужно применить авторегрессионные модели – изменить значения коэффициентов и, возможно, увеличить количество слагаемых в правых частях уравнений авторегрессии, то есть повысить ее порядок.

4.2.4. Волновые модели изображений

В волновой модели И определяется равенством

$$x_{\bar{j}}^t = \sum_{\{k:\tau_k \leq t\}} f((\bar{j}, t), (\bar{u}_k, \tau_k), \bar{\omega}_k), \quad (4.8)$$

где $(n+1)$ -мерная область определения $\{\bar{j}, t\}$ поля может быть сеточной или непрерывной; $\{\bar{u}_k, \tau_k\}$ – дискретное **поле случайных точек** (ПСТ) в $(n+1)$ -мерном непрерывном пространстве; t и τ_k интерпретируются как время; $\bar{\omega}_k$ – случайный вектор параметров функции f .

Это И можно представить как результат воздействия случайных возмущений или волн $f(\bar{j}, t, \bar{u}_k, \tau_k, \bar{\omega}_k)$, возникающих в случайных местах \bar{u}_k в случайные моменты времени τ_k и изменяющихся по заданному закону во времени и пространстве. Выбор функции f , параметров ПСТ и $\bar{\omega}$ позволяет получить широкий класс И.

Рассмотрим частный случай волновой модели, для которой корреляционные задачи анализа и синтеза легко решаются. Пусть

$$f(\bar{j}, t, \bar{u}_k, \tau_k, \bar{\omega}_k) = c \exp(-2(\rho_k / R_k)^2) \exp(-\mu |t - \tau_k|) \xi_k, \quad (4.9)$$

где ПСТ – пуассоновское с постоянной плотностью λ ; $\rho_k = |\bar{j} - \bar{u}_k|$ – расстояние между \bar{j} и \bar{u}_k ; $\{R_{kj}\}$ – система независимых неотрицательных одинаково распределенных СВ с ПРВ $\omega(\alpha)$; $\{\xi_k\}$ – система независимых одинаково распределенных СВ. В этом случае волны неподвижны, независимы между собой, имеют сферические сечения по пространству и экспоненциально затухают со временем; система $\{\xi_k\}$ определяет интенсивность волн, а $\{R_{kj}\}$ – их пространственный масштаб.

Порождаемое И X стационарно, однородно, имеет нулевое среднее и изотропную по пространству КФ

$$V(\rho, t) = \frac{c^2 \pi^{n/2} \lambda}{2^{n+1}} e^{-\mu t} \int_0^\infty \alpha^n \exp(-\rho^2 / \alpha^2) \omega(\alpha) d\alpha. \quad (4.10)$$

При $\rho = t = 0$ из (4.10) находим дисперсию И

$$\sigma_n^2 = \frac{c^2 \pi^{n/2} \lambda}{2^{n+1} \mu} M[R^n], \quad (4.11)$$

пропорциональную плотности λ ПСТ, эффективному интервалу $1/\mu$ затухания волн и среднему значению n -й степени пространственного масштаба R .

Имитация дискретного И на n -мерной сетке $\{\bar{j}\}$ с шагом квантования Δt по времени на основе модели рассмотренного частного вида может быть осуществлена с помощью следующего алгоритма. В начальный момент $t_0 = 0$ значения И во всех узлах равны нулю. В каждый последующий момент $t_m = m\Delta t$ на непрерывном пространстве или на сетке, несколько перекрывающей $\{\bar{j}\}$, формируется пуассоновское ПСТ с плотностью $\lambda\Delta t$. В каждой

сформированной точке ПСТ \bar{u}_k разыгрываются СВ ξ_k и R_k , после чего производится преобразование

$$x_{\bar{j}}^{t_m} = x_{\bar{j}}^{t_{m-1}} \exp(-\mu \cdot \Delta t) + c \sum_k \exp(-2(\rho_k / R_k)^2) \xi_k \quad (4.12)$$

всех значений И на сетке $\{\bar{j}\}$. При таком моделировании в (4.12) можно учитывать только достаточно большие по сравнению с уровнем квантования слабые. Достоинством такого алгоритма является его рекуррентность, что позволяет легко реализовать имитацию И на ЭВМ.

Пуассоновское ПСТ можно сформировать следующим способом. За время Δt на площади кадра ΔS в среднем должно возникать $a = \lambda \Delta S \Delta t$ точек. Если $a \ll 1$, то можно считать, что с вероятностью a возникает только одна точка. Плотность λ обычно мала, поэтому в качестве площадок можно брать отдельные пиксели (единичной площади). Таким образом, в каждом пикселе возникает одна точка из ПСТ с вероятностью $a = \lambda \Delta t$. Для имитации этого возникновения нужно разыграть случайное число φ , равномерно распределенное на отрезке $[0;1]$; точка появляется, если оказалось $\varphi < a$, и не появляется в противном случае.

Рассмотрим теперь решение корреляционных задач анализа и синтеза. Из (4.10) следует, что построенное поле имеет экспоненциальную нормированную КФ (НКФ) $e^{-\mu t}$ по времени и НКФ

$$r(\rho) = \frac{1}{M[R^n]} \int_0^\infty \alpha^n \exp(-\rho^2 / \alpha^2) \omega(\alpha) d\alpha \quad (4.13)$$

по пространству. Таким образом, при решении задачи анализа, когда ПРВ $\omega(\alpha)$ задана, искомая НКФ может быть найдена аналитически или численным интегрированием. При решении задачи синтеза, когда НКФ $r(\rho)$ задана, необходимо решить интегральное уравнение (4.13) относительно неизвестной ПРВ $\omega(\alpha)$.

Поскольку найти аналитическое решение задачи синтеза удастся не всегда, рассмотрим метод ее приближенного решения. Из (4.13) следует, что при вырожденном распределении ($R = \alpha = const$) получаем КФ $\exp(-\rho^2 / \alpha^2)$. Пусть теперь задана произвольная невозрастающая НКФ $r(\rho)$. Аппроксимируем ее с достаточной точностью суммой гауссоид с положительными коэффициентами: $r(\rho) \approx h(\rho) = \sum_i q_i \exp(-\rho^2 / \alpha_i^2)$, где $\sum_i q_i = 1$, так как $r(0) = 1$.

Тогда при дискретном распределении $P(R = \alpha_i) = k^{-1} q_i / \alpha_i^n$, где $k = \sum q_i / \alpha_i^n$, порождаемое поле будет иметь НКФ, в точности равную $h(\rho)$. Таким образом, построенная модель позволяет приближенно решать задачу синтеза с помощью вариации только распределения вероятностей масштаба R . Решение, очевидно, будет точным, если $r(\rho) = h(\rho)$.

Пример 1. Пусть требуется сформировать последовательность плоских И, имеющих НКФ $r(\rho) = 0,3 \exp(-\rho^2 / 2^2) + 0,7 \exp(-\rho^2 / 3^2)$ по пространству. Тогда $n=2$, $k=0,3/2^2 + 0,7/3^2 = 0,153$, $P(R=2) = 0,491$, $P(R=3) = 0,509$. Таким образом, R равняется 2 или 3 с вероятностями 0,491 и 0,509. Если коэффициент корреляции на единичном интервале времени должен быть равен, например, 0,99, то из уравнения $\exp(-\mu \cdot 1) = 0,99$ находим $\mu = -\ln 0,99 = 0,01$. Параметры c и λ найдем из (4.11) при заданной дисперсии σ^2 , учитывая, что $M[R^n] = M[R^2] = 2^2 \cdot 0,491 + 3^2 \cdot 0,509 = 6,545$, имеем $c^2 \lambda = 2^3 \mu / M[R^2] \pi \sigma^2 = 0,0039 \sigma^2$. Взяв $c = 1$, получаем $\lambda = 0,0039 \sigma^2$.

4.2.4. Оценка параметров изображений

При обработке И или построении их моделей часто возникает задача оценки их параметров. Пусть имеется, например, плоское однородное И $X = \{x_{ij}\}$. Требуется оценить его основные характеристики – математическое ожидание m_x , дисперсию σ_x и КФ $V_x(i, j) = M[(x_{ab} - m_x)(x_{a+i, b+j} - m_x)]$. Обычно И состоит из большого числа пикселей, поэтому достаточно хорошие оценки этих характеристик можно получить в виде выборочных средних:

$$\hat{m}_x = \frac{1}{N} \sum_{i,j} x_{ij}, \quad \hat{\sigma}_x^2 = \frac{1}{N} \sum_{i,j} (x_{ij} - \hat{m}_x)^2 = \frac{1}{N} \sum_{i,j} x_{ij}^2 - \hat{m}_x^2, \quad (4.14)$$

$$\hat{V}_x(a, b) = \frac{1}{N} \sum_{i,j} (x_{ij} - \hat{m}_x)(x_{i+a, j+b} - \hat{m}_x) = \frac{1}{N} \sum_{i,j} x_{ij} x_{i+a, j+b} - \hat{m}_x^2, \quad (4.15)$$

где N – количество элементов в соответствующем усреднении.

Отметим, что в случае неоднородности И оценки (4.14) и (4.15) следует применять только к участкам, где И относительно однородны. Иначе будут получаться неверные, чрезмерно усредненные характеристики.

4.3. Содержание отчета

1. Титульный лист.
2. Цель лабораторной работы.
3. Задание на работу.
4. Краткое изложение теоретической части.
5. Алгоритм имитации И.
6. Распечатка программы имитации.
7. Распечатка нескольких И.
8. Таблица оценок параметров имитированных И (для КФ плоского И – дополнительно график изолиний, как на рис. 4.2).
9. Выводы по работе.

4.4. Контрольные вопросы

1. Основная модель И в виде сеточного СП.
2. Гауссовские СП. Однородные и стационарные СП.
3. Корреляционная функция И.
4. Линейные авторегрессионные модели СП.
5. Авторегрессионная модель Хабиби многомерных И, их КФ.
6. Волновая модель И, задачи ее анализа и синтеза.
7. Оценка основных характеристик И.

4.4. Задания на лабораторную работу

1. Имитация последовательности кадров И заданных размеров с заданными характеристиками, используя заданную модель. Составленная программа должна обеспечивать визуализацию имитированных И.
2. Оценка среднего значения, дисперсии и КФ имитированных И.

5. ЛАБОРАТОРНАЯ РАБОТА № 2 «ОЦЕНКА ИЗОБРАЖЕНИЙ»

5.1. Цель работы

Изучение некоторых методов оценки элементов изображений по совокупности наблюдений.

5.2. Теоретическая часть

5.2.1. Оценка элементов изображения как статистическая задача оценки параметров

При обработке И часто требуется оценить значения его элементов, имея ту или иную информацию. Пусть требуется построить оценку \hat{x} элементов $\bar{x} = (x_1, \dots, x_m)^T$ изображения X по имеющимся наблюдениям $\bar{z} = (z_1, \dots, z_n)^T$. Вектор \bar{x} может состоять и из одного элемента x . Если вектор наблюдений \bar{z} не содержит элементы \bar{x} , то оценка \hat{x} называется **прогнозом**. В этом случае \bar{z} обычно содержит элементы И, соседние с \bar{x} . Такая задача может возникнуть, например, если \bar{x} по каким-то причинам не наблюдаемы. Если же \bar{z} содержит зашумленные (искаженные) наблюдения элементов \bar{x} , то говорят о задаче **оценки** или **фильтрации** \bar{x} . В этой лабораторной работе исследуется задача прогноза, но теоретическая часть может быть отнесена к любому виду оценки.

Рассматриваемая задача оценки элементов И является одной из основных задач математической статистики. Оптимальной оценкой (оптимальным

решающим правилом) считается оценка, минимизирующая средние потери. Чаще всего применяется квадратичная функция потерь. При построении оценки следует учитывать всю имеющуюся информацию: модель И и возможных его искажений.

Отметим несколько методов построения оценок. **Метод максимума апостериорной ПРВ** – оптимальной оценкой $\hat{\bar{x}}$ является точка максимума апостериорной ПРВ $w(\bar{x} | \bar{z})$ по \bar{x} :

$$w(\hat{\bar{x}}, \bar{z}) = \max_{\bar{x}} w(\bar{x} | \bar{z}). \quad (5.1)$$

Метод максимального правдоподобия – в качестве оценки берется точка максимума ФП $P(\bar{z} | \bar{x}) = w(\bar{z} | \bar{x})$:

$$P(\bar{z} | \hat{\bar{x}}) = \max_{\bar{x}} P(\bar{z} | \bar{x}). \quad (5.2)$$

Метод наименьших квадратов – оценка должна минимизировать средний квадрат ошибки оценки: $M[(\hat{x}_i - x_i)^2] = \min$.

Метод моментов - в качестве оценки берется такой вектор $\hat{\bar{x}}$, при котором теоретические значения первых моментов наблюдений \bar{z} равны их выборочным значениям.

5.2.2. Оценка гауссовских параметров по гауссовским наблюдениям

Во многих приложениях И имеют гауссовские распределения, поэтому рассмотрим этот случай подробнее.

Совместная ПРВ m гауссовских оцениваемых случайных величин x_1, \dots, x_m , составляющих гауссовский вектор $\bar{x} = (x_1, \dots, x_m)^T$, имеет вид

$$w(x_1, \dots, x_m) = w(\bar{x}) = \frac{1}{(2\pi)^{m/2} \det^{1/2}(V_{xx})} \exp\left(-\frac{1}{2}(\bar{x} - \bar{m})^T V_{xx}^{-1}(\bar{x} - \bar{m})\right), \quad (5.3)$$

где $\bar{m} = M[\bar{x}] = (M[x_1], M[x_2], \dots, M[x_m])^T = (m_1, \dots, m_m)^T$ – вектор математических ожиданий (математическое ожидание вектора \bar{x}); $V_{xx} = M[(\bar{x} - \bar{m})(\bar{x} - \bar{m})^T] = \{M[(x_i - m_i)(x_j - m_j)^T]; i = \overline{1, m}, j = \overline{1, m}\}$ – матрица ковариаций (автоковариаций). Отметим, что V_{xx} – симметричная матрица: $V_{xx}^T = V_{xx}$. В частности, если $\bar{m} = 0$, т. е. $M[x_i] = 0$, то

$$w(\bar{x}) = \frac{1}{(2\pi)^{m/2} \det^{1/2}(V_{xx})} \exp\left(-\frac{1}{2}\bar{x}^T V_{xx}^{-1}\bar{x}\right), \quad (5.4)$$

где $V_{xx} = M[\bar{x} \bar{x}^T]$.

Аналогичную совместную ПРВ имеет и гауссовский вектор наблюдений $\bar{z} = (z_1, \dots, z_n)^T$ с ковариационной матрицей V_{zz} .

Рассмотрим задачу оценки гауссовского вектора $\bar{x} = (x_1, \dots, x_m)^T$, когда имеется гауссовский вектор наблюдений $\bar{z} = (z_1, \dots, z_n)^T$. Пусть известны все средние значения и все ковариации величин $x_i, i = \overline{1, m}$, и $z_i, i = \overline{1, n}$. Без потери общности можно считать, что $M[\bar{x}] = 0$ и $M[\bar{z}] = 0$, так как мы можем центрировать все величины, вычитая из них их математические ожидания. Если оптимальность оценки понимать в смысле минимума средних квадратов ошибок, то есть при квадратичной функции потерь $M[(\hat{x}_i - x_i)^2]$, то оптимальной оценкой будет

$$\hat{\bar{x}} = V_{xz} V_{zz}^{-1} \bar{z}, \quad (5.5)$$

где $V_{xz} = M[\bar{x} \bar{z}^T] = \{M[x_i z_j], i = \overline{1, m}, j = \overline{1, n}\}$ – матрица ковариаций x_i и z_j (кросс-ковариаций). Отметим, что $V_{xz} = V_{zx}^T$. Отсюда следует очень важный факт: **оптимальная оценка гауссовских параметров по гауссовским наблюдениям линейна (есть линейная функция наблюдений \bar{z}).**

Матрица ковариаций ошибок оценки (5.5) есть

$$M[(\hat{\bar{x}} - \bar{x})(\hat{\bar{x}} - \bar{x})^T] = V_{xx} - V_{xz} V_{zz}^{-1} V_{zx} = T. \quad (5.6)$$

При этом дисперсии ошибок равны диагональным элементам матрицы T . Матрица ковариаций ошибок оценок и наблюдений равна нулю:

$$M[(\hat{\bar{x}} - \bar{x}) \bar{z}^T] = 0. \quad (5.7)$$

Это очень важное обстоятельство: **ошибки оптимальных оценок не коррелированы с наблюдениями.**

Отметим, что оценка (5.5) есть **оптимальная линейная оценка \bar{x} для любых центрированных векторов \bar{x} и \bar{z}** , то есть это оптимальная оценка среди оценок вида $F \bar{z}$, где F – любая матрица. Но эта оценка не обязательно оптимальна для негауссовских векторов.

Рассмотрим теперь важный случай оценки одного параметра, то есть пусть $\bar{x} = x$ и $\hat{x} = \bar{\alpha}^T \bar{z}$, где $\bar{\alpha}$ – весовой вектор оценки. Тогда (5.7) принимает вид $M[(\bar{\alpha}^T \bar{z} - x) \bar{z}^T] = 0$, где

$$\bar{\alpha}^T = V_{xz} V_{zz}^{-1}, \quad \bar{\alpha} = V_{zz}^{-1} V_{zx}. \quad (5.8)$$

Из (5.8) следует, что $\bar{\alpha}$ является решением системы линейных уравнений

$V_{zz} \bar{\alpha} = V_{zx}$ или в развернутом виде

$$\begin{pmatrix} V_{z_1 z_1} & V_{z_1 z_2} & \dots & V_{z_1 z_n} \\ \dots & \dots & \dots & \dots \\ V_{z_n z_1} & V_{z_n z_2} & \dots & V_{z_n z_n} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \dots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} V_{x z_1} \\ \dots \\ V_{x z_n} \end{pmatrix}. \quad (5.9)$$

При этом средний квадрат ошибки оценки определяется из (5.6):

$$M[(\hat{x} - x)^2] = V_{xx} - V_{xz} V_{zz}^{-1} V_{zx} = \sigma^2 - V_{xz} V_{zz}^{-1} V_{zx}. \quad (5.10)$$

Этот средний квадрат ошибки можно представить в другом виде. Рассмотрим определитель полной ковариационной матрицы V (наблюдений и оцениваемого параметра):

$$|V| = \begin{vmatrix} V_{xx} & V_{xz} \\ V_{zx} & V_{zz} \end{vmatrix}. \quad (5.11)$$

Тогда

$$M[(\hat{x} - x)^2] = |V|/|V_{zz}|, \quad (5.12)$$

то есть средний квадрат ошибки оценки равен отношению определителя полной ковариационной матрицы к определителю ковариационной матрицы наблюдений.

Пример 1. Рассмотрим оценку гауссовского сигнала x с параметрами $(0, \sigma_x^2)$ по зашумленному наблюдению $z = x + \theta$, где θ – гауссовский шум с параметрами $(0, \sigma_\theta^2)$, независимый от x . В нашем случае, используя (5.8), получаем

$$V_{zz} = M[z^2] = M[(x + \theta)^2] = \sigma_x^2 + \sigma_\theta^2, \quad V_{zx} = M[xz] = M[x(x + \theta)] = \sigma_x^2, \\ \alpha = V_{zx}V_{zz}^{-1} = \sigma_x^2(\sigma_x^2 + \sigma_\theta^2)^{-1} = \sigma_x^2/(\sigma_x^2 + \sigma_\theta^2).$$

Итак, оптимальной оценкой является

$$\hat{x} = z\sigma_x^2/(\sigma_x^2 + \sigma_\theta^2). \quad (5.13)$$

Дисперсию ошибки этой оценки получим из (5.10):

$$M[(\hat{x} - x)^2] = \sigma_x^2 - V_{zx}V_{zz}^{-1}V_{zx} = \sigma_x^2 - \sigma_x^2 \frac{1}{\sigma_x^2 + \sigma_\theta^2} \sigma_x^2 = \sigma_x^2 \left(1 - \frac{\sigma_x^2}{\sigma_x^2 + \sigma_\theta^2}\right) = \frac{\sigma_x^2 \sigma_\theta^2}{\sigma_x^2 + \sigma_\theta^2},$$

или, по-другому, из (5.12):

$$M[(\hat{x} - x)^2] = \frac{\begin{vmatrix} \sigma_x^2 & \sigma_x^2 \\ \sigma_x^2 & \sigma_x^2 + \sigma_\theta^2 \end{vmatrix}}{\sigma_x^2 + \sigma_\theta^2} = \frac{\sigma_x^2 \sigma_\theta^2}{\sigma_x^2 + \sigma_\theta^2}.$$

Пример 2. Рассмотрим прогноз элемента x_i последовательности, определяемой моделью (1.1) с $\rho_1 = \rho$, по наблюдениям четырех его ближайших элементов

$\bar{z} = (x_{i-2}, x_{i-1}, x_{i+1}, x_{i+2})^T$. Из симметрии можно сделать вывод, что веса наблюдений

$x_{i\pm 1}$, (а также $x_{i\pm 2}$) одинаковы, следовательно, оценка имеет вид

$$\hat{x}_i = \alpha_1 x_{i-1} + \alpha_1 x_{i+1} + \alpha_2 x_{i-2} + \alpha_2 x_{i+2} = \alpha_1 (x_{i-1} + x_{i+1}) + \alpha_2 (x_{i-2} + x_{i+2}) = \alpha_1 z_1 + \alpha_2 z_2.$$

Поэтому для сокращения размерности представим наблюдения в виде

$\bar{z} = (x_{i-1} + x_{i+1}, x_{i-2} + x_{i+2})^T = (z_1, z_2)^T$. Найдем элементы матриц V_{zz} и V_{zx} из (5.8):

$$V_{z_1, z_1} = M[z_1 z_1] = M[(x_{i-1} + x_{i+1})^2] = M[x_{i-1}^2 + 2x_{i-1}x_{i+1} + x_{i+1}^2] = \sigma^2 + 2\sigma^2\rho^2 + \sigma^2 = 2(1 + \rho^2)\sigma^2,$$

$$V_{z_1, z_2} = M[z_1 z_2] = M[(x_{i-1} + x_{i+1})(x_{i-2} + x_{i+2})] = M[x_{i-1}x_{i-2} + x_{i-1}x_{i+2} + x_{i+1}x_{i-2} + x_{i+1}x_{i+2}] =$$

$$= 2\rho_1(1 + \rho^2)\sigma^2, \quad V_{x_i, z_1} = M[x_i z_1] = M[x_i(x_{i-1} + x_{i+1})] = M[x_i x_{i-1} + x_i x_{i+1}] = 2\rho\sigma^2,$$

$$V_{x_i, z_2} = M[x_i z_2] = M[x_i(x_{i-2} + x_{i+2})] = M[x_i x_{i-2} + x_i x_{i+2}] = 2\rho^2\sigma^2, \quad V_{z_2, z_2} = 2(1 + \rho^4)\sigma^2.$$

Таким образом,

$$\bar{\alpha}^T = V_{zz} V_{zz}^{-1} = (2\rho \sigma^2, 2\rho^2 \sigma^2) \begin{pmatrix} 2(1+\rho^2)\sigma^2 & 2\rho(1+\rho^2)\sigma^2 \\ 2\rho(1+\rho^2)\sigma^2 & 2(1+\rho^4)\sigma^2 \end{pmatrix}^{-1} = \left(\frac{\rho}{1+\rho^2}, 0 \right). \quad \text{Итак, опти-}$$

мальный прогноз есть

$$\hat{x}_i = \bar{\alpha}^T \bar{z} = \left(\frac{\rho}{1+\rho^2}, 0 \right) \begin{pmatrix} x_{i-1} + x_{i+1} \\ x_{i-2} + x_{i+2} \end{pmatrix} = \frac{\rho}{1+\rho^2} (x_{i-1} + x_{i+1}), \quad (5.14)$$

а дисперсия его ошибки $M[(\hat{x}_i - x_i)^2] = (1 - \rho^2)/(1 + \rho^2)$. Отметим, что в эту оценку не вошли наблюдения $x_{i\pm 2}$ (их веса оказались нулевыми). Это является следствием марковости рассматриваемой последовательности: условное распределение вероятностей элемента x_i при известных значениях элементов $x_{i\pm 1}$ не зависит от более удаленных элементов. Поэтому прогноз (5.14) не изменится, если в \bar{z} добавить к $x_{i\pm 1}$ любые другие элементы (конечно, кроме самого x_i – тогда задача прогноза бессмысленна).

Пример 3. Рассмотрите самостоятельно обобщение примера 2, когда наблюдения элементов x_i зашумлены: $z_i = x_i + \theta_i$, где $\{\theta_i\}$ – независимые между собой и от x_i гауссовские помехи с нулевым средним и дисперсией σ_θ^2 . Убедитесь в том, что веса наблюдений $z_{i\pm 2}$ не будут нулевыми, так как последовательность z_i не марковская.

Пример 4. Рассмотрим теперь прогноз элемента x_{ij} И, определяемого моделью (4.3), по наблюдениям его восьми ближайших соседей $\bar{z} = (x_{i\pm 1, j}, x_{i, j\pm 1}, x_{i\pm 1, j\pm 1})^T$, см. рис. 5.1.

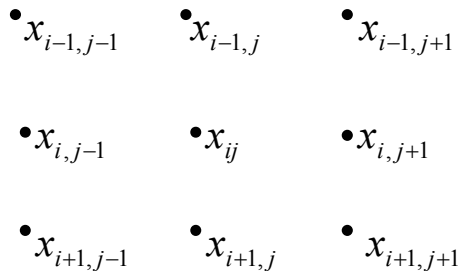


Рис. 5.1.

Как и в примере 2, в силу марковости, добавление к уже имеющимся другим наблюдений прогноза не улучшает. Из-за симметрии веса у четырех наблюдений $x_{i\pm 1, j\pm 1}$ (расположенных по диагоналям), у двух наблюдений $x_{i\pm 1, j}$ и у двух наблюдений $x_{i, j\pm 1}$ будут одинаковыми. Предположим для простоты, что $\rho_1 = \rho_2 = \rho$. Тогда веса всех четырех наблюдений $x_{i\pm 1, j}$ и $x_{i, j\pm 1}$ будут одинаковыми. Следовательно, наблюдения можно представить в виде $\bar{z} = (z_1, z_2)^T$, где $z_1 = x_{i, j-1} + x_{i, j+1} + x_{i-1, j} + x_{i+1, j}$ и $z_2 = x_{i-1, j-1} + x_{i-1, j+1} + x_{i+1, j-1} + x_{i+1, j+1}$. Покажите, что $V_{z_1, z_1} = 4(1 + 3\rho^2)\sigma^2$, $V_{z_2, z_2} = 4(1 + \rho^2)^2 \sigma^2$, $V_{z_1, z_2} = 8\rho(1 + \rho^2)\sigma^2$, $V_{x_{ij}, z_1} = 4\rho\sigma^2$, $V_{x_{ij}, z_2} = 4\rho^2\sigma^2$, найдите оптимальный прогноз $\hat{x}_{i, j} = \alpha_1 z_1 + \alpha_2 z_2$ и дисперсию его ошибки.

5.2.3. Адаптивная оценка изображений

В предыдущих выкладках предполагалось, что модели И и возможных помех заданы вместе со своими параметрами. На практике это далеко не всегда так, поэтому в реальных ситуациях имеется априорная неопределенность в описании данных. Для преодоления этой неопределенности приходится применять адаптивные алгоритмы обработки, в частности, оценки И. При этом имеются различные подходы к построению адаптивных алгоритмов.

В *идентификационных алгоритмах* сначала по всем имеющимся данным оцениваются все недостающие неизвестные характеристики $\bar{\gamma}$ обрабатываемых И. Затем полученные оценки $\hat{\gamma}$ используются как точные при определении параметров $\bar{\alpha}$ алгоритма обработки. При этом, конечно, в оценках И возникают дополнительные ошибки из-за неточности определения характеристик И.

Пример 4. Предположим, что в примерах 2-4 неизвестны параметры σ^2 и ρ . Тогда, применяя результаты лабораторной работы 1, эти параметры можно оценить. Затем, используя оценки этих параметров как точные значения, воспользоваться уже полученными алгоритмами.

В алгоритмах *без идентификации* параметры $\bar{\alpha}$ алгоритма обработки И подбираются так, чтобы результат обработки был как можно лучше. Пусть оптимальность понимается в смысле минимума критерия (функционала) качества $J(\bar{\alpha})$. Тогда нужно найти точку минимума $\bar{\alpha}^*$ этого функционала, что можно сделать и без промежуточных оценок каких-либо характеристик γ исходных данных. При этом $\bar{\alpha}$ могут подбираться итерационно в процессе текущей обработки по наблюдениям за текущими значениями $J(\bar{\alpha})$.

Для реализации таких алгоритмов необходима оценка текущего значения $J(\bar{\alpha})$, т. е. критерий должен быть наблюдаемым, что является ограничением на применение этого подхода. Иногда выход может быть найден с помощью замены $J(\bar{\alpha})$ на наблюдаемый критерий $J_1(\bar{\alpha})$ с той же точкой минимума.

Весьма перспективными для реализации этого подхода являются *псевдоградиентные* адаптивные алгоритмы.

5.2.4. Псевдоградиентные алгоритмы

Пусть структура процедуры обработки И определена, а критерий качества решения задачи сформулирован в терминах минимизации функционала $J(\bar{\alpha})$, где $\bar{\alpha}$ – параметры алгоритма обработки. Ввиду априорной неопределенности описания данных нет возможности заранее определить оптимальные параметры $\bar{\alpha}^*$. Поэтому необходима некоторая процедура адаптации, составляющая вместе с процедурой обработки адаптивный алгоритм, в кото-

ром параметры $\bar{\alpha}$ определяются на основании конкретной реализации (наблюдения) Z объекта обработки.

Существует ряд численных методов поиска экстремумов. Наиболее распространенными являются различные модификации градиентного алгоритма

$$\bar{\alpha}_n = \bar{\alpha}_{n-1} - \mu_n \nabla J(\bar{\alpha}_{n-1}), \quad (5.15)$$

где $\bar{\alpha}_n$ – следующее за $\bar{\alpha}_{n-1}$ приближение к точке минимума; μ_n – положительная числовая последовательность, определяющая длину шагов; $\nabla J(\bar{\alpha})$ – градиент функции $J(\bar{\alpha})$. Каждый шаг в (5.15) делается в направлении скорейшего убывания $J(\bar{\alpha})$.

Применению этих методов в обработке И препятствует необходимость многократных и громоздких вычислений $\nabla J(\bar{\alpha}_{n-1}, Z)$, каждое из которых обычно включает в себя всю процедуру обработки Z при параметрах $\bar{\alpha}_{n-1}$. Значительно сократить объем вычислений можно, если вместо $\nabla J(\bar{\alpha}_{n-1}, Z)$ взять усечение $\nabla Q(\bar{\alpha}_{n-1}) = \nabla J(\bar{\alpha}_{n-1}, Z_n)$, то есть вычислять градиент не по всей реализации Z , а только по некоторой ее части Z_n , например, в скользящем окне на И. Но тогда в (5.15) вместо точного значения градиента будет использоваться его значение со случайной ошибкой $\bar{\delta}_n$. Последовательность $\bar{\alpha}_n$ становится случайной, поэтому случаен и сам факт ее сходимости к $\bar{\alpha}^*$. Эти случайные ошибки, вообще говоря, не являются серьезным препятствием для сходимости $\bar{\alpha}_n \rightarrow \bar{\alpha}^*$.

В 1973 г. Я. З. Цыпкиным и Б. Т. Поляком было введено понятие **псевдоградиента** (ПГ), на основе которого разработан единый подход к анализу и синтезу алгоритмов стохастической минимизации функционалов. Класс ПГ алгоритмов очень широк и включает в себя все (или почти все) алгоритмы адаптации и обучения. Эти алгоритмы основаны на процедуре

$$\bar{\alpha}_n = \bar{\alpha}_{n-1} - \mu_n \bar{\beta}_n, \quad (5.16)$$

где $\bar{\beta}_n$ – случайное (в частности, детерминированное) направление, вообще говоря, зависящее от предыдущих значений $\bar{\alpha}_i$ и от номера шага n . Направление $\bar{\beta}_n$ называется **псевдоградиентом** функционала $J(\bar{\alpha})$ в точке $\bar{\alpha}_{n-1}$, если выполнено **условие псевдоградиентности**

$$[J(\bar{\alpha}_{n-1})]^T M[\bar{\beta}_n] \geq 0, \quad (5.17)$$

где левая часть есть скалярное произведение, поэтому **ПГ в среднем составляет острый угол с точным значением градиента**.

Алгоритм (5.17) называется **псевдоградиентным**, если $\bar{\beta}_n$ является ПГ на каждом шаге. В этом случае шаги в (5.17) будут производиться в среднем в сторону уменьшения $J(\bar{\alpha})$, поэтому выполнение относительно слабых ус-

ловий оказывается достаточным для сходимости с вероятностью единица при любом начальном приближении $\bar{\alpha}_0$.

Главным из этих условий является строгое неравенство в (5.17) при $\bar{\alpha}_{n-1} \neq \bar{\alpha}^*$, единственность точки минимума, а также $\sum \mu_n = \infty$ (что обеспечивает возможность дойти из любой точки до $\bar{\alpha}^*$) и $\sum \mu_n^2 < \infty$ (что обеспечивает возможность асимптотического уменьшения дисперсии колебаний последовательности $\bar{\alpha}_n$). Последним двум условиям удовлетворяют, например, последовательности вида $\mu_n = 1/(a + bn)$.

До сих пор предполагалось, что задачей является нахождение точки минимума $\bar{\alpha}^*$ функционала $J(\bar{\alpha}, Z)$, единой для всей реализации Z . Такая точка $\bar{\alpha}^*$ существует, но обработка будет оптимальной, если данные Z **однородны**. Для сходимости $\bar{\alpha}_n \rightarrow \bar{\alpha}^*$ при этом требуется сходимость $\mu_n \rightarrow 0$. Если же, начиная с некоторого момента, ограничить μ_n снизу (например, взять постоянные $\mu_n = \mu$), то дисперсии ошибок оценок $\bar{\alpha}_n$ параметров $\bar{\alpha}^*$ перестанут уменьшаться, а сами $\bar{\alpha}_n$ будут колебаться около $\bar{\alpha}^*$. Таким образом, если обработку однородных данных производить одновременно с оценкой $\bar{\alpha}^*$ (и при $\mu_n = \mu$), то по достижении установившегося режима будет осуществляться **квазиоптимальная** обработка.

Если произойдет скачкообразное изменение характеристик Z или переход к обработке других данных, то могут измениться и значения требуемых оптимальных параметров $\bar{\alpha}_n^*$. Если обработка будет просто продолжена, то непосредственно после этого скачка возможно значительное ухудшение качества обработки, после чего постепенно снова будут достигнуты квазиоптимальные результаты.

При плавном изменении характеристик наблюдений Z (точнее, при плавном изменении оптимальных значений параметров $\bar{\alpha}^*$), соизмеримом со скоростью переходного процесса процедуры (5.17), появляется возможность применения ПГ алгоритмов к обработке неоднородных И без их сегментации на участки относительно однородной структуры.

5.2.4. Псевдоградиентные алгоритмы прогноза изображений

Будем считать прогноз \hat{x}_i элемента x_i оптимальным, если достигается минимум среднего квадрата ошибки $M[(\hat{x}_i - x_i)^2] = M[\Delta_i^2]$ прогноза. В этом случае оптимальным прогнозом будет условное математическое ожидание $M[x_i | \bar{z}_i]$, где \bar{z}_i – вектор наблюдений, по которому строится прогноз. Пусть определен вид функции прогноза $\hat{x}_i = f(\bar{\alpha}_i, \bar{z}_i)$, где $\bar{\alpha}_i$ – параметры функции

прогноза, зависящие от характеристик И. Например, в случае гауссовских изображений с нулевым средним оптимальный прогноз линеен: $\hat{x}_i = \bar{\alpha}_i^T \bar{z}_i$.

Если вид функции прогноза определен, то задача сводится к его оптимизации, то есть к нахождению оптимальных значений $\bar{\alpha}_i^*$ параметров $\bar{\alpha}_i$, при которых достигается минимум квадрата ошибки:

$$J(\bar{\alpha}_i) = M[(f(\bar{\alpha}_i, \bar{z}_i) - x_i)^2]. \quad (5.18)$$

Построим ПГ алгоритм, минимизирующий этот функционал. Возьмем алгоритм (5.16) с ПГ вектором

$$\bar{\beta}_n = \nabla(f(\bar{\alpha}_n, z_n) - x_n)^2 = 2\Delta_n \frac{\partial f(\bar{\alpha}_n, z_n)}{\partial \bar{\alpha}_n} \quad \text{или} \quad \bar{\beta}_n = \text{sign} \left(\Delta_n \frac{\partial f(\bar{\alpha}_n, z_n)}{\partial \bar{\alpha}_n} \right). \quad (5.19)$$

Нахождение градиента в (5.19) затруднений не вызывает, так как функция прогноза f задана.

Например, при линейном прогнозе $\partial f(\bar{\alpha}_n, \bar{z}_n) / \partial \bar{\alpha}_n = \partial(\bar{\alpha}_n^T \bar{z}_n) / \partial \bar{\alpha}_n = \bar{z}_n$. Отсюда получаем ПГ алгоритм настройки весовых коэффициентов прогноза:

$$\bar{\alpha}_{n+1} = \bar{\alpha}_n - \mu_n \text{sign}(\Delta_n \bar{z}_n), \quad (5.20)$$

где $\Delta_n = \hat{x}_n - x_n = \bar{\alpha}_n^T \bar{z}_n - x_n$ – ошибка прогноза. Процедура (5.20) выполняется в порядке просмотра элементов И одновременно с построением самого прогноза. В качестве начального приближения $\bar{\alpha}_0$ вектора весовых коэффициентов можно взять любое, например, (0.5; 0.5).

Замечание. В построенном алгоритме негласно предполагалось, что элементы x_i наблюдаемы, что было необходимо для вычисления текущих значений градиента функционала (5.18). Однако у нас обычно имеются зашумленные наблюдения $z_i = x_i + \theta_i$, оптимальный прогноз которых при независимости и центрированности шумов θ_i совпадает с оптимальным прогнозом элементов x_i истинного И. Поэтому функционал (5.18) можно заменить вспомогательным функционалом

$$J_1(\bar{\alpha}_i) = M[(f(\bar{\alpha}_i, z_i) - z_i)^2], \quad (5.21)$$

то есть в (5.18)-(5.20) нужно только заменить x на z .

Пример 5. Применим описанный ПГ алгоритм для прогноза изображений из примера 4. Прогноз очередного элемента И имеет вид $\hat{x}_n = \bar{\alpha}_n^T \bar{z}_n = \alpha_{n1} z_{n1} + \alpha_{n2} z_{n2}$ и $\Delta_n = \alpha_{n1} z_{n1} + \alpha_{n2} z_{n2} - z_n$. Из (5.20) получаем

$$\bar{\alpha}_{n+1}^T = (\alpha_{n+1,1}, \alpha_{n+1,2}) = (\alpha_{n,1}, \alpha_{n,2}) - \mu_n \text{sign}(\Delta_n z_{n,1}, \Delta_n z_{n,2}), \quad (5.22)$$

то есть весовыми коэффициентами для прогноза следующего элемента x_{n+1} будут $\alpha_{n+1,1} = \alpha_{n,1} - \mu_n \text{sign}(\Delta_n z_{n,1})$ и $\alpha_{n+1,2} = \alpha_{n,2} - \mu_n \text{sign}(\Delta_n z_{n,2})$.

На рис. 5.2 приведен результат применения этого алгоритма. На рис. 5.2.а показано исходное И, на рис.5.2.б – его оптимальный прогноз, на рис. 5.2.в – его ошибки. Рис. 5.2.г соответствует адаптивному ПГ прогнозу, рис. 5.2.д – его ошибкам. Визуально рисунки

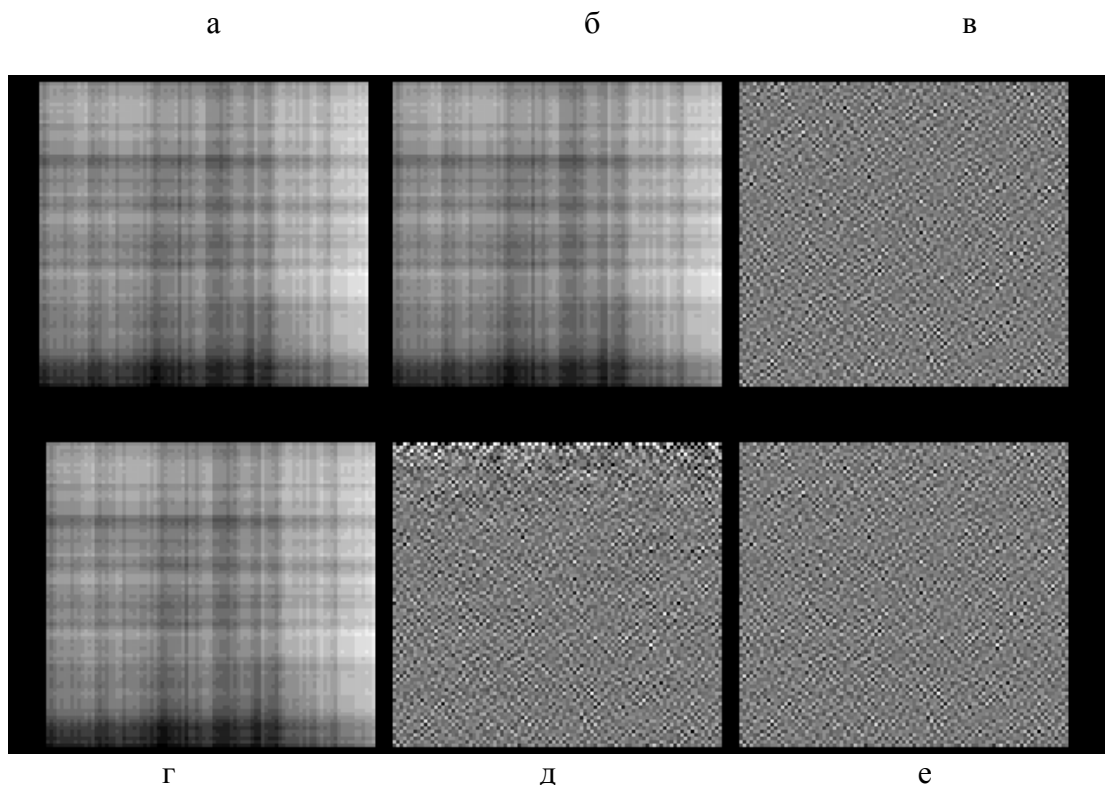


Рис. 5.2.

5.2.б и 5.2.г одинаковы, так как ПГ прогноз близок к оптимальному. Тем не менее ошибки прогноза (увеличенные для визуализации) на рис. 5.2.в и 5.2.д различаются между собой. В нескольких первых (верхних) строках ПГ прогноза (рис. 5.2.д) ошибки относительно большие, так как идет процесс подстройки параметров прогноза. Этот процесс быстро устанавливается. На рис. 5.2.е показаны ошибки прогноза при втором проходе И, который начат со значений параметров прогноза, установившихся к концу первого прохода. Анализ рис. 5.2.в и 5.2.е позволяет сделать вывод об отсутствии какой-либо разницы между ними, что говорит о том, что прогноз практически сошелся к оптимальному.

5.3. Содержание отчета

1. Титульный лист.
2. Цель лабораторной работы.
3. Задание на работу.
4. Краткое изложение теоретической части.
5. Описание алгоритмов прогноза.
6. Распечатка программ прогноза.
7. Распечатка спрогнозированных И.
8. Сводка оценок качества прогноза (СКО ошибки).
9. Выводы по работе.

5.4. Контрольные вопросы

1. Постановка задачи оценки И.
2. Прогноз и фильтрация И.
3. Оценка И как статистическая задача.
4. Прогноз гауссовских И, линейность оптимального прогноза.
5. Идентификационный адаптивный прогноз.
6. Безыдентификационный адаптивный прогноз.
7. Псевдоградиентная адаптация.
8. Псевдоградиентный адаптивный прогноз.

5.4. Задания на лабораторную работу

1. Построить оптимальный прогноз авторегрессионной последовательности с известными параметрами. Вычислить выборочное СКО ошибки и сравнить его с теоретическим значением.
2. Построить адаптивный идентификационный прогноз той же последовательности. Сравнить конечные значения весовых коэффициентов с их оптимальными значениями, полученными в п.1 этого задания. Вычислить выборочное СКО ошибки и сравнить его с СКО, полученным в п.1.
3. Построить ПГ прогноз той же последовательности. Вычислить выборочное СКО ошибки и сравнить его с СКО, полученными в п.1 и п.2.
4. Выполнить задания в пп.1-3 для И, заданного моделью Хабиби. Программы должны обеспечивать визуализацию результатов (см. рис. 5.2).

6. ЛАБОРАТОРНАЯ РАБОТА № 3 «ОБНАРУЖЕНИЕ СИГНАЛОВ НА ФОНЕ ИЗОБРАЖЕНИЙ»

6.1. Цель работы

Изучение неадаптивных методов обнаружения сигналов на фоне мешающих изображений.

6.2. Теоретическая часть

6.2.1. Постановка задачи обнаружения

Рассмотрим задачу обнаружения сигналов на фоне мешающего И, то есть случайного сеточного поля помех $Y = \{y_j^i : \bar{j} \in \Omega, i = I\}$. Наблюдения $Z = \{z_j^i\}$ заданы на той же сетке и могут содержать (гипотеза H_1) или не содержать (гипотеза H_0) сигнал S . По имеющимся наблюдениям Z требуется

выбрать одну из этих гипотез. Поставленная задача является частным случаем двухальтернативных статистических решений.

Если известны совместные условные ПРВ $w_z(Z|H_0)$ и $w_z(Z|H_1)$ наблюдений Z соответственно при отсутствии и наличии сигнала S , то, как это следует из общих принципов статистических решений, оптимальное решающее правило основано на отношении правдоподобия (ОП):

$$\Lambda(Z) = \frac{w_z(Z|H_1)}{w_z(Z|H_0)} \begin{cases} \geq \Lambda_0 \Rightarrow H_1 \\ < \Lambda_0 \Rightarrow H_0 \end{cases}, \quad (6.1)$$

где $\Lambda(Z)$ – ОП и Λ_0 – порог, значение которого зависит от априорной информации и понятия оптимальности обнаружения.

При обработке И характерна малость размеров апертуры сигнала относительно размеров сетки, поэтому будем предполагать, что действие сигнала проявляется только в некоторой ограниченной области $G \subset \Omega \times I$. Тогда при наличии сигнала изменяются наблюдения Z_G только в этой области, а в наблюдениях $Z_D = Z \setminus Z_G$ в дополнительной области $D = (\Omega \times I) \setminus G$ наличие сигнала никак не сказывается (рис. 6.1). Используя это обстоятельство и предполагая аддитивное взаимодействие $Z_G = Y_G + S_G$ сигнала и помех, можно привести ОП к виду

$$\Lambda(Z) = \frac{w_Y(Z_G - S_G|Z_D)}{w_Y(Z_G|Z_D)}, \quad (6.2)$$

содержащему условные ПРВ наблюдений в области G при заданных значениях наблюдений вне этой области.

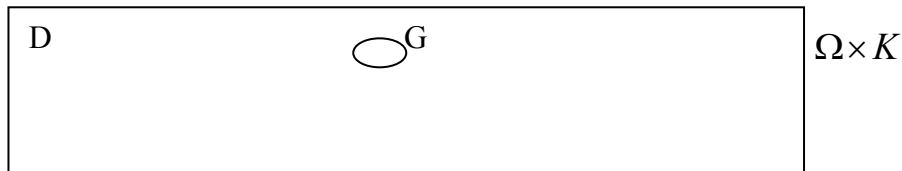


Рис. 6.1.

Отметим, что положение обнаруживаемого сигнала обычно неизвестно, поэтому обнаружение нужно производить для каждого его возможного положения, то есть для всевозможных расположений области G на I .

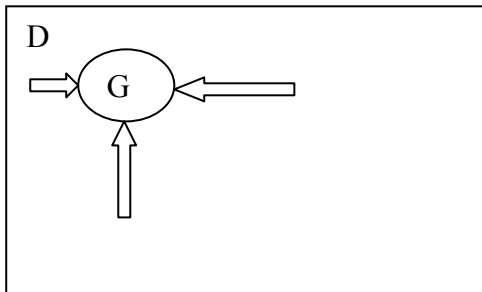
6.2.2. Обнаружение при гауссовской аппроксимации распределений помех

Предположим, что ПРВ $w_Y(Y)$ помех может быть аппроксимировано гауссовским распределением с нулевыми средними. Если, кроме того, сигнал аддитивно взаимодействует с помехами, то с учетом (6.2) и (5.3) решающее правило (6.1) приводится к виду

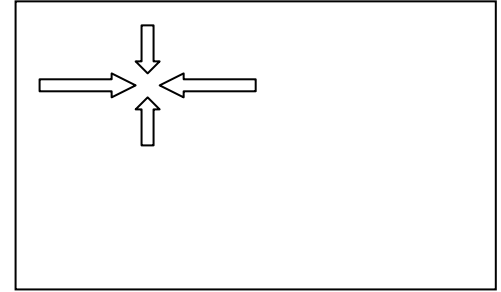
$$\lambda(Z) = S_G^T V_G^{-1} \Delta_G = S_G^T V_G^{-1} (Z_G - \hat{Z}_G) \begin{cases} \geq \lambda_0 \Rightarrow H_1 \\ < \lambda_0 \Rightarrow H_0 \end{cases}, \quad (6.3)$$

где $\lambda_0 = \ln \Lambda_0 + \frac{1}{2} S_G^T V_G^{-1} S_G^T$ – новый порог, соответствующий статистике $\lambda(Z)$; $V_G = M[\Delta_G \times \Delta_G]$ – ковариационная матрица ошибок $\Delta_G = Z_G - \hat{Z}_G$ прогноза $\hat{Z}_G = \hat{Y}_G = M[Y_G | Z_D]$ значений Y_G в области G по наблюдениям Z_D вне G .

Вычитание прогноза \hat{Z}_G из Z_G можно рассматривать как компенсацию помех в области G при известных значениях помех Z_D вне этой области, а разности $\Delta_G = Z_G - \hat{Z}_G$ – как остатки этой компенсации. Назовем этот тип прогноза **прогнозом в область**, так как он строится для всех наблюдений в G по всем наблюдениям вне этой области (рис. 6.2.а).



а



б

Рис. 6.2.

Таким образом, алгоритм обнаружения (6.3) включает в себя компенсацию помех в G путем вычитания из наблюдений Z_G их прогноза \hat{Z}_G по наблюдениям Z_D . После компенсации осуществляется весовое суммирование остатков компенсации $\Delta_G = Z_G - \hat{Z}_G$ с весами $S_G^T V_G^{-1}$ и, наконец, сравнение полученной статистики $\lambda(Z)$ с порогом λ_0 .

Статистику $\lambda(Z)$ в (6.3) можно представить в более удобной для использования форме:

$$\lambda(Z) = S_G^T L_Y^{-1} \Delta^* = S_G^T L_Y^{-1} (Z - Z^*), \quad (6.4)$$

где Z^* – совокупность прогнозов; L_Y – диагональная матрица дисперсий ошибок $\Delta^* = Z - Z^*$ этих прогнозов. При этом прогноз $z_j^* = M[z_j | (Z \setminus z_j)]$ каждого наблюдения z_j по остальным наблюдениям из Z строится при предположении об отсутствии сигнала, то есть теми же действиями, которые выполняются при прогнозе мешающих изображений Y . Будем называть этот тип прогноза **прогнозом в точку** (рис. 6.2.б).

Таким образом, статистика (6.4) включает в себя вычитание $\Delta^* = Z - Z^*$ из наблюдений Z их прогнозов в точку Z^* , деление получен-

ных разностей их собственными дисперсиями и суммирование полученных значений с весами, равными отсчетам сигнала S_G . Полученная статистика сравнивается с порогом λ_0 .

Отметим, что $L_Y^{-1}\Delta_Z^*$ вычисляется один раз. При обнаружении сигнала в разных положениях G меняется только область весового суммирования в (6.4). Если же использовать (6.3), то компенсацию помех приходится вычислять для каждого возможного положения области заново.

Несмотря на равенство статистик (6.3) и (6.4), между ними есть принципиальное различие. В (6.3) прогноз и компенсация выполняются по наблюдениям Z_D , которые сигнала не содержат, поэтому при наличии сигнала в G он будет искажен только ошибками прогноза. Если эти ошибки малы, то остатки компенсации Δ_G будут близки к S_G (визуально можно увидеть сигнал с небольшими искажениями). В (6.4) при построении прогноза в точку используются все остальные наблюдения, в том числе и содержащие сигнал. Поэтому в остатках Δ_G^* этой компенсации каждый отсчет сигнала будет искажен не только ошибками прогноза мешающего И, но и остальными отсчетами сигнала. Даже при малых ошибках прогноза визуально будет наблюдаться очень искаженный сигнал.

Более предпочтительной является статистика (6.4) в силу следующих обстоятельств. Поскольку матрица L_Y диагональная, в (6.4) требуется только нормирование остатков компенсации их собственными дисперсиями. Построение прогноза в точку выполняется более унифицированно: этот прогноз не зависит от формы области G . Поэтому после выполнения преобразования $L_Y^{-1}\Delta_Z^*$ уже несложно осуществить обнаружение сигналов любых форм и размеров. Значительно облегчается синтез квазиоптимальных алгоритмов, в которых при компенсации используется прогноз только части наблюдений. Дисперсии ошибок прогнозов σ_ε^2 в точку для однородных И приблизительно равны между собой (по крайней мере, на некотором удалении от границ И), поэтому (6.4) может быть аппроксимировано выражением

$$\lambda(Z) = S_G^T \Delta_Z^* / \sigma_\varepsilon^2 \quad (6.5)$$

или даже заменено статистикой

$$\lambda(Z) = S_G^T \Delta_Z^* \quad (6.6)$$

с заменой порога λ_0 на $\lambda_1 = \sigma_\varepsilon^2 \lambda_0$.

6.2.3. Характеристики обнаружения сигналов

Важнейшими характеристиками обнаружения являются вероятность ложной тревоги

$$P_F = P(\lambda(Z) \geq \lambda_0 | H_0) \quad (6.7)$$

и вероятность правильного обнаружения

$$P_D = P(\lambda(Z) \geq \lambda_0 | H_1). \quad (6.8)$$

В задачах обнаружения сигналов обычно используется критерий Неймана-Пирсона: задается вероятность ложной тревоги P_F , по которой из уравнения (6.7) находится порог λ_0 . Затем из (6.8) находится вероятность правильного обнаружения P_D .

Например, при гауссовском распределении помех Y и их аддитивном взаимодействии с сигналом статистика $\lambda(Z)$ имеет нормальное распределение. При гипотезе H_0 сигнал в наблюдениях отсутствует, среднее значение статистики (6.2) $m_0 = 0$ и дисперсия $\sigma_0^2 = S_G^T V_G^{-1} S_G$, поэтому

$$P_F = 1 - \Phi(\lambda_0 / \sigma_0), \quad (6.9)$$

где $\Phi(*)$ – функция Лапласа (полная). При гипотезе H_1 сигнал в наблюдениях имеется, среднее значение статистики становится равным $m_1 = S_G^T V_G^{-1} S_G$, а дисперсия сохраняется. Таким образом,

$$\sigma_0^2 = \sigma_1^2 = m_1 = S_G^T V_G^{-1} S_G. \quad (6.10)$$

Из (6.8) и (6.10) получаем

$$P_D = 1 - \Phi((\lambda_0 - m_1) / \sigma_0) = 1 - \Phi(\lambda_0 / \sigma_0 - \sqrt{m_1}). \quad (6.11)$$

Пример 1. Рассмотрим обнаружение сигнала $S_G^T = c \cdot (1, 2, -3)$ на фоне случайного процесса, заданного авторегрессионной моделью (4.1) с параметрами $\rho_1 = \rho$ и $\sigma = 1$, где c – постоянный множитель. Из примера 2 лабораторной работы 2, следует, что прогнозы в каждую точку зависят только от двух ближайших соседей и дисперсии ошибок одинаковы, поэтому будем использовать статистику в форме (6.5). Возьмем область $G = (i - 1, i, i + 1)$. Тогда $z_i^* = a(z_{i-1} + z_{i+1})$, где $a = \rho / (1 + \rho^2)$,

$$\begin{aligned} \lambda(Z) &= [c(z_{i-1} - a(z_{i-2} + z_i)) + 2c(z_i - a(z_{i-1} + z_{i+1})) - 3c(z_{i+1} - a(z_i + z_{i+2}))] / \sigma_\varepsilon^2 = \\ &= c[-az_{i-2} + (1 - 2a)z_{i-1} + (2 + 2a)z_i - (3 + 2a)z_{i+1} + 3az_{i+2}] / \sigma_\varepsilon^2, \end{aligned}$$

где $\sigma_\varepsilon^2 = (1 - \rho^2) / (1 + \rho^2)$ – дисперсия ошибки прогноза в точку. При наличии сигнала имеем $M[z_{i-1}] = c$, $M[z_i] = 2c$, $M[z_{i+1}] = -3c$, поэтому

$m_1 = c[(1 - 2a)c + (2 + 2a)2c + (3 + 2a)3c] / \sigma_\varepsilon^2 = (14 + 8a)c^2 / \sigma_\varepsilon^2$. Такое же значение, согласно (6.10), имеют дисперсии $\sigma_0^2 = \sigma_1^2$. Пусть, например, $\rho = 0.5$, тогда

$a = 0.4$, $\sigma_\varepsilon^2 = 0.6$, $\sigma_0^2 = \sigma_1^2 = m_1 = 28.67c^2$. Пусть задана вероятность ложной тревоги $P_F = 0.001$, тогда из (6.9) имеем $0.001 = 1 - \Phi(\lambda_0 / 5.35c)$, $\Phi(\lambda_0 / 5.35c) = 0.999$,

$\lambda_0 / 5.35c = 3.1$, $\lambda_0 = 16.6c$. Подставляя эти значения в (6.10), получаем вероятность правильного обнаружения $P_D(c) = 1 - \Phi(3.1 - 5.35c)$, график которой показан на рис.6.3. Как уже отмечалось ранее, можно использовать статистику (6.6), то есть исключить в статистике деление на σ_ε^2 , но тогда нужно использовать порог

$\sigma_\varepsilon^2 \lambda_0 = 0.6 \cdot 16.6c = 10c$. Характеристики обнаружения при этом не изменятся.

Отметим при этом, что $P_D(0) = 1 - \Phi(3.1) = P_F$, поэтому график начинается не из начала координат. Получается, что при $c = 0$, когда сигнал нулевой, то есть его фактически нет, он тем не менее обнаруживается. Этот странный факт связан с тем, что имеется веро-

ятность ложной тревоги P_F , которая как раз и равна вероятности принятия решения о наличии сигнала, когда его на самом деле нет (это то же самое, что есть нулевой сигнал). Поэтому, когда имеется нулевой сигнал (например, если при передаче настоящего сигнала отказал передатчик), мы его и «обнаруживаем» с вероятностью P_F .

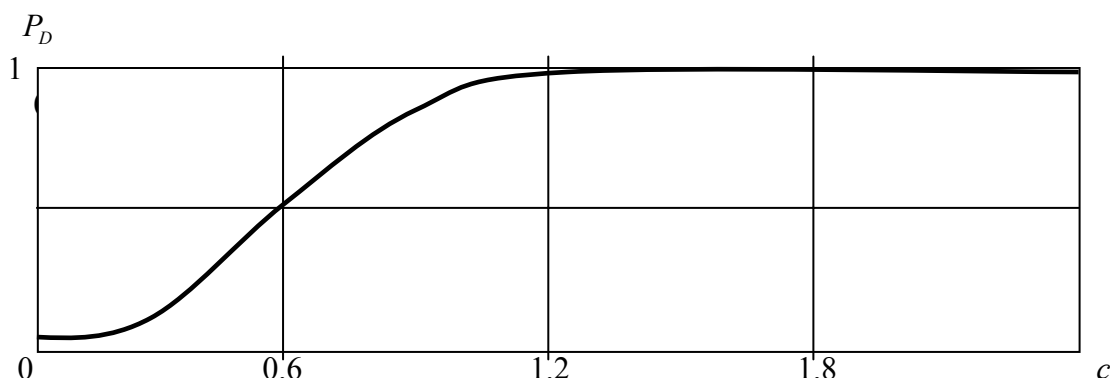


Рис. 6.3.

Пример 2. Рассмотрим обнаружение точечного сигнала $S = c$ на изображении, заданном моделью Хабиби (4.3) с параметрами $\rho_1 = \rho_2 = \rho$ и $\sigma = 1$. В данном примере алгоритмы, основанные на прогнозе в область и в точку, совпадают. Используя результаты примера 4 лабораторной работы 2, имеем $\lambda(Z) = c(z_{i,j} - az_1 - bz_2) / \sigma_\varepsilon^2$. Сигнал $S = c$ в выражении этой статистики присутствует только в отсчете $z_{i,j}$, поэтому $\sigma_0^2 = \sigma_1^2 = m_1 = c^2 / \sigma_\varepsilon^2$. Дальнейшие выкладки производятся как в примере 1.

6.3. Содержание отчета

1. Титульный лист.
2. Цель лабораторной работы.
3. Задание на работу.
4. Краткое изложение теоретической части.
5. Описание алгоритмов обнаружения.
6. Распечатка программ обнаружения.
7. Распечатка картины обнаружения (исходное И, остатки компенсации, статистика и ее сечение порогом, см. рис. 7.1).
8. Сводка характеристик обнаружения (расчетные и выборочные).
9. Выводы по работе.

6.4. Контрольные вопросы

1. Постановка задачи обнаружения сигналов и ее статистическая формулировка.
2. Оптимальное правило обнаружения, отношение правдоподобия.
3. Алгоритм обнаружения при гауссовских помехах.
4. Алгоритмы обнаружения, основанные на прогнозе в область и в точку.

5. Характеристики обнаружения – вероятность ложной тревоги и правильного обнаружения.
6. Определение порога по критерию Неймана-Пирсона.

6.4. Задания на лабораторную работу

1. Произвести обнаружение протяженного сигнала заданной формы на фоне авторегрессионного процесса. Определить порог. Найти расчетные и выборочные характеристики обнаружения.
2. Произвести обнаружение сигнала на фоне И, заданного моделью Хабиби. Определить порог. Найти расчетные и выборочные характеристики обнаружения.

7. ЛАБОРАТОРНАЯ РАБОТА № 4 «АДАПТИВНОЕ ОБНАРУЖЕНИЕ СИГНАЛОВ НА ФОНЕ ИЗОБРАЖЕНИЙ»

7.1. Цель работы

Изучение адаптивных методов обнаружения сигналов на фоне мешающих изображений.

7.2. Теоретическая часть

7.2.1. Адаптация в обнаружении сигналов

В предыдущей лабораторной работе предполагалось, что все необходимые распределения И и вид сигнала заданы вместе со своими параметрами. На практике это далеко не всегда так, поэтому в реальных ситуациях имеется априорная неопределенность в описании данных.

Рассмотрим задачу обнаружения сигналов на фоне мешающего И, разобранную в п. 6.2.2. Оптимальный алгоритм обнаружения (6.3) включает в себя компенсацию помех путем вычитания из наблюдений их прогнозов, взвешенное суммирование остатков компенсации и сравнение полученной статистики $\lambda(Z)$ с порогом λ_0 . При этом статистика может вычисляться в формах (6.3)-(6.6), основываясь на прогнозе в область или в точку. При неопределенности описания исходных данных мы не можем определить параметры этого алгоритма – коэффициенты прогноза и порог.

Таким образом, при неопределенности описания исходных данных мы не можем применить имеющийся алгоритм обнаружения в чистом виде. Для преодоления этой неопределенности приходится применять адаптивные алгоритмы обнаружения.

Оставим собственно алгоритм обнаружения без изменения и добавим к нему процедуры определения параметров алгоритма. При этом, как

и в задачах оценки И, можно применять идентификационную и безыдентификационную адаптацию. В этой лабораторной работе будут применяться ПГ адаптивные алгоритмы обнаружения.

Прежде всего нужно выполнить прогноз и компенсацию мешающих изображений. Адаптивные ПГ алгоритмы прогноза уже были рассмотрены в лабораторной работе 3. Поэтому перейдем к адаптивному определению порога.

7.2.2. Адаптивное определение порога

Рассмотрим задачу определения порога λ в решающем правиле обнаружения

$$\Lambda(Z) \leq \lambda, \quad (7.1)$$

где $\Lambda(Z) = \Lambda$ – статистика правила. При использовании критерия Неймана-Пирсона порог определяется по заданной вероятности ложной тревоги (принятия решения о наличии сигнала, когда его нет) $p = P_F$ из соотношения

$$P(\Lambda \geq \lambda | H_0) = p \quad (7.2)$$

или

$$F(\lambda) = P(\Lambda < \lambda | H_0) = 1 - p = q, \quad (7.3)$$

где $F(\lambda)$ – условная ФР статистики Λ при отсутствии сигнала. Итак, порог λ является *q-квантилью* СВ Λ при отсутствии обнаруживаемого сигнала.

Если $F(x)$ известна, то q -квантиль можно определить из уравнения (7.3). Мы же рассматриваем именно случай неизвестности этой функции распределения. Кроме того, на неоднородных данных она может изменяться. Отсюда возникает задача оценки q -квантили (в обнаружении – стабилизации порога, точнее – стабилизации вероятности ложной тревоги на заданном уровне).

Пусть имеется последовательность наблюдаемых СВ $\Lambda_1, \Lambda_2, \dots$ (вычисленных значений статистики Λ). Требуется последовательно определять оценку q -квантили λ_{qn} очередной СВ Λ_n .

Можно, например, построить следующую оценку. Если имеется n выборочных значений $\Lambda_1, \dots, \Lambda_n$ статистики Λ , то величина $\hat{\lambda}_n$, отсекающая qn наименьших из наблюдаемых $\Lambda_1, \dots, \Lambda_n$, является оценкой q -квантили по относительной частоте события $\Lambda < \hat{\lambda}_n$. Но для такой сортировки нужно хранить все наблюдения. Если мешающие И неоднородны, то все «обнаружения» могут оказаться в области с большой дисперсией помех. Менее громоздкое и лучшее по характеристикам решение можно получить, применяя ПГ алгоритмы.

Одноконтурный алгоритм

Применим для оценки q -квантили ПГ алгоритм

$$\hat{\lambda}_{n+1} = \hat{\lambda}_n + \mu_n \begin{cases} q, & \text{если } \Lambda_n \geq \hat{\lambda}_n, \\ -p, & \text{если } \Lambda_n < \hat{\lambda}_n, \end{cases} \quad (7.4)$$

где μ_n – положительная числовая последовательность. В (7.4) следующая оценка $\hat{\lambda}_{n+1}$ увеличивается на $\mu_n q$, если Λ_n достигает порог $\hat{\lambda}_n$, в противном случае оценка уменьшается на $\mu_n p$.

Проанализируем действие алгоритма (7.4). Для этого найдем математическое ожидание приращения его оценок:

$$\begin{aligned} M[h_n] &= M[\hat{\lambda}_{n+1} - \hat{\lambda}_n] = \mu_n [qP(\Lambda_n \geq \hat{\lambda}_n) - pP(\Lambda_n < \hat{\lambda}_n)] = \\ &= \mu_n [q(1 - F_\Lambda(\hat{\lambda}_n)) - pF_\Lambda(\hat{\lambda}_n)] = \mu_n (q - F_\Lambda(\hat{\lambda}_n)). \end{aligned}$$

Если $\hat{\lambda}_n < \lambda_q$, то $F(\hat{\lambda}_n) < F(\lambda_q) = q$ и $M[h_n] > 0$, поэтому оценка будет в среднем повышаться. Если $\lambda_q < \hat{\lambda}_n$, то $q = F(\lambda_q) < F(\hat{\lambda}_n)$ и $M[h_n] < 0$, поэтому оценка будет в среднем понижаться. Если $\hat{\lambda}_n = \lambda_q$, то $M[h_n] = 0$. Следовательно, у оценок $\hat{\lambda}_n$ имеется тенденция двигаться к точной квантили λ_q . Если при этом $\mu_n \rightarrow 0$ (а также $\sum \mu_n = \infty$ и $\sum \mu_n^2 < \infty$), то $\hat{\lambda}_n \rightarrow \lambda_q$ с вероятностью 1.

Этот алгоритм очень прост в реализации, не требует задания распределения $F(x)$ и запоминания или сортировки наблюдаемых значений Λ_i .

В случае неоднородных данных ФР $F_{\Lambda_n}(x)$ величин Λ_n могут быть различны. Если при этом текущая квантиль λ_{qn} изменяется медленно, то алгоритм (7.4) может достаточно точно оценивать изменяющуюся квантиль, для этого нужно взять μ_n ограниченными снизу (например, $\mu_n = \text{const}$). Таким образом, этот алгоритм применим и для оценки переменных квантилей, что очень важно в задачах обнаружения сигналов на неоднородных И.

Однако этот алгоритм имеет и существенный недостаток. В задачах обнаружения характерны малые задаваемые значения ложной тревоги p (0.001 и меньше), а $q \approx 1$, т. е. нужно оценивать квантиль очень высокого порядка. Алгоритм (7.4) имеет очень большую асимметрию: шаги «вверх» значительно больше, чем «вниз». Увеличение порога происходит и в случае обнаружения сигнала, хотя увеличивать порог в этом случае не следует. Если И существенно неоднородно и квантиль меняется очень быстро, то алгоритм не успевает ее отслеживать, особенно при уменьшении точной квантили. В результате в среднем порог оказывается «заданным».

Значительно лучше алгоритм (7.4) оценивает квантили, близкие к 1/2, когда отмеченная асимметрия невелика. Воспользуемся этим для улучшения оценок квантилей.

Двухконтурный алгоритм

Предположим, что статистика Λ_n представима в виде $\Lambda_n = U_n K_n$, где распределения величин K_n постоянны или меняются очень медленно, а распределения величин U_n могут меняться относительно быстро. Такое представление применимо во многих приложениях. Например, когда тип распределения статистики Λ почти не меняется, а интенсивность помех, влияющая на U_n , изменяется достаточно быстро.

Будем искать оценку квантили λ_{qn} величины Λ_n в виде $\hat{\lambda}_n = \hat{U}_n \hat{K}_n$, где \hat{U}_n – оценка r -квантили величины Λ_n и $1/2 < r \ll q$, используя то, что такие квантили хорошо оцениваются алгоритмом (7.4). Тогда \hat{K}_n будет коэффициентом пропорциональности между λ_{rn} и λ_{qn} . Используем для нахождения оценок \hat{U}_n и \hat{K}_n две процедуры вида (7.4):

$$\begin{aligned} \hat{U}_{n+1} &= \hat{U}_n + \mu_n \begin{cases} r, & \text{если } \Lambda_n \geq \hat{U}_n, \\ -s, & \text{если } \Lambda_n < \hat{U}_n, \end{cases} \\ \hat{K}_{n+1} &= \hat{K}_n + \nu_n \begin{cases} q, & \text{если } \Lambda_n < \hat{U}_n \hat{K}_n, \\ -p, & \text{если } \Lambda_n \geq \hat{U}_n \hat{K}_n, \end{cases} \end{aligned} \quad (7.5)$$

при этом выбирается $\nu_n \ll \mu_n$, так как распределение величин K_n меняется значительно медленнее, чем распределение величин U_n .

Как показали испытания, этот алгоритм хорошо отслеживает порог даже при быстром его изменении. Например, при обнаружении сигналов на существенно неоднородных И ложные тревоги распределялись достаточно равномерно по полю кадра, а их количество укладывалось в допустимые пределы.

7.2.3. Неопределенность в задании обнаруживаемого сигнала

Возьмем статистику $\lambda(Z)$ в форме (6.6). Весовые коэффициенты остатков компенсации в $\lambda(Z)$ равны отсчетам обнаруживаемого сигнала S , который тоже может быть известен лишь приближенно. Будем считать, что по крайней мере форма $F^T = (s_1, \dots, s_k)$ сигнала задана, то есть

$$S_G^T = c F_G^T = c \cdot (s_1, \dots, s_k). \quad (7.6)$$

Таким образом, возможна еще и неопределенность общего коэффициента c в статистике:

$$\lambda(Z) = S_G^T \Delta^* = c F_G^T \Delta^*. \quad (7.7)$$

Пусть c – **неизвестный, но постоянный коэффициент**, то есть обнаруживаемые сигналы одинаковы. Предположим также, что этот коэффициент положителен (если известно, что он отрицателен, то можно сделать его положительным, сменив знак у формы). Решающее правило $c F_G^T \Delta^* \leq \lambda_0$ при статистике (7.7), очевидно, эквивалентно правилу $F_G^T \Delta^* \leq \lambda_0 / c$ с изменен-

ным порогом. Поэтому неопределенность коэффициента c можно перенести в порог, взяв статистику

$$\lambda(Z) = F_G^T \Delta^* . \quad (7.8)$$

Порог же, как показано в п. 7.2.2, можно определять независимо от конкретного вида статистики.

Совсем другая ситуация возникает, когда в (7.7) c – **неизвестный и переменный коэффициент**, то есть обнаруживаемые сигналы в силу каких-то причин могут быть разными. Это уже другая статистическая задача. Если нам требуется только обнаруживать сигналы, то это задача при сложной гипотезе H_1 . Если еще и требуется определить, какой именно сигнал обнаружен, то это задача многоальтернативного различения гипотез.

Рассмотрим случай, когда коэффициент c может принимать два значения: a (гипотеза H_{11}) и $-a$ (гипотеза H_{12}). Тогда для проверки гипотезы H_{11} против H_0 нужно использовать статистику $\lambda(Z) = F_G^T \Delta^*$, а для H_{12} против H_0 – статистику $\lambda(Z) = -F_G^T \Delta^*$. Порог в обоих случаях один и тот же. В первом приближении эти два варианта статистики можно представить одним

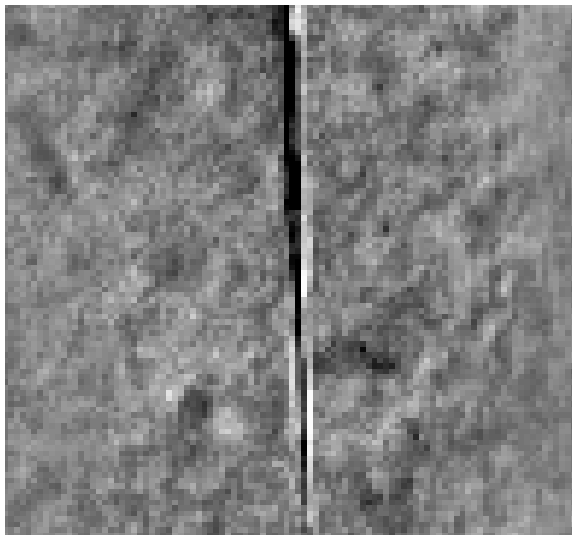
$$\lambda(Z) = |F_G^T \Delta^*| , \quad (7.9)$$

подбирая порог именно для этой статистики.

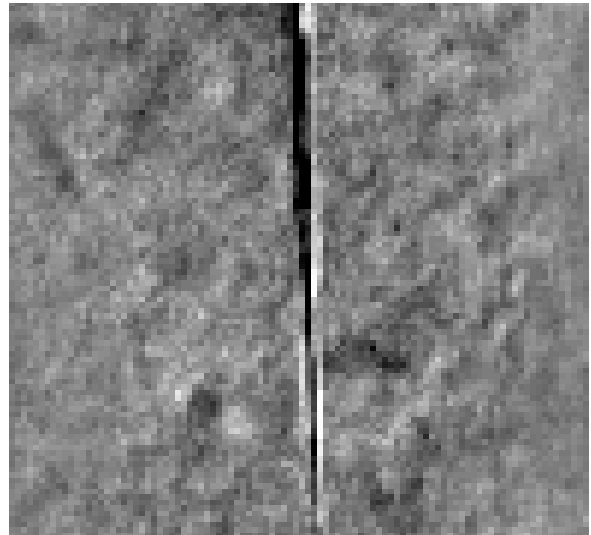
Пример 1. Имеются два кадра $Z^1 = \{z_{ij}^1\}$ и $Z^2 = \{z_{ij}^2\}$ центрированного гауссовского мешающего И. Аддитивные сигналы могут присутствовать только на втором кадре. Они представляют собой прямоугольники с одинаковыми отсчетами, равными 3 или -3. Другими словами, присутствие сигнала (объекта) в таком прямоугольнике повышает или понижает яркость на 3 единицы. Требуется построить алгоритм обнаружения. Применим адаптивный алгоритм со структурой, разобранный выше. Статистику решающего правила возьмем в форме (7.9).

Самой сложной составляющей обнаружения является компенсация мешающих И. Иногда при небольшой потере качества эту процедуру можно значительно упростить. Пусть И имеет сильную межкадровую корреляцию, значительно превышающую внутрикадровую. Такое положение имеет место, когда сюжет за время между регистрацией двух его кадров почти не изменяется и аппаратура работает достаточно стабильно. В этом случае достаточно точный прогноз элемента z_{ij}^2 можно построить по наблюдению одного элемента z_{ij}^1 с предыдущего кадра в виде $\hat{z}_{ij}^2 = \alpha z_{ij}^1$ (и это будет одновременно приближенный прогноз и в точку, и в область). Тогда ПГ адаптивная процедура компенсации мешающих И очень проста.

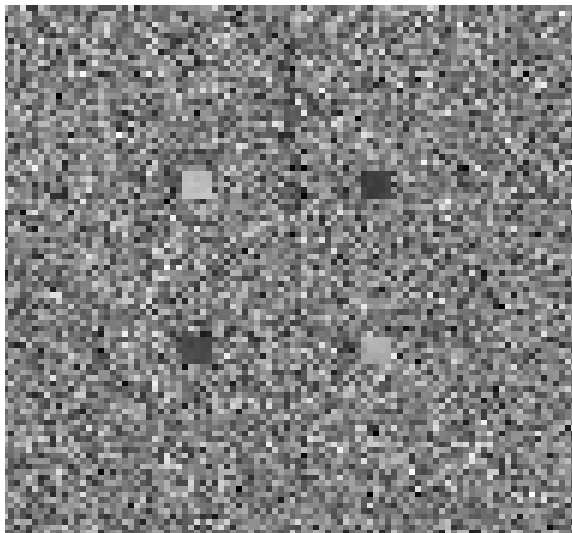
На рис. 7.1 показан пример применения этого алгоритма к И морского дна. Рис. 7.1.а – первый кадр, рис. 7.1.б – второй кадр на котором присутствуют четыре визуально незаметных прямоугольных объекта. На рис. 7.1.в – остатки компенсации, объекты уже просматриваются. На рис. 7.1.г – статистика (7.9), объекты заметны лучше. На рис. 7.1.д – сечение статистики порогом при вероятности ложной тревоги $P_f = 0.001$. На рис. 7.1.е – то же при $P_f = 0.0001$. Отметим, что в остатках компенсации на рис. 7.1.в просматривается действительная форма объектов, это следствие компенсации в область. Группы же ярких точек на рис. 7.1.д и 7.1.е имеют форму, отличную от прямоугольной формы объектов, что



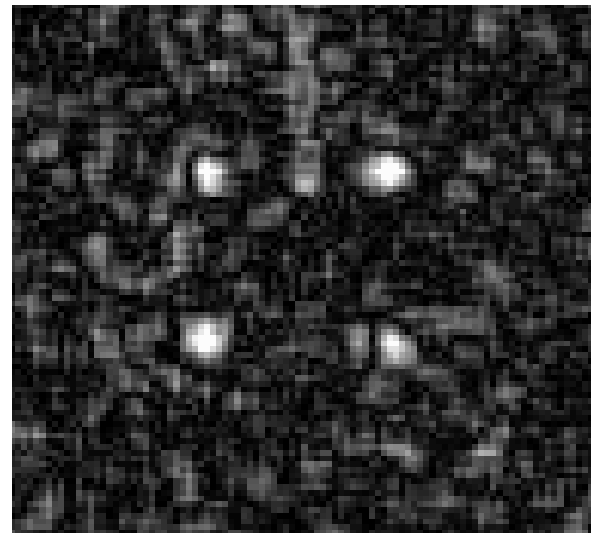
а



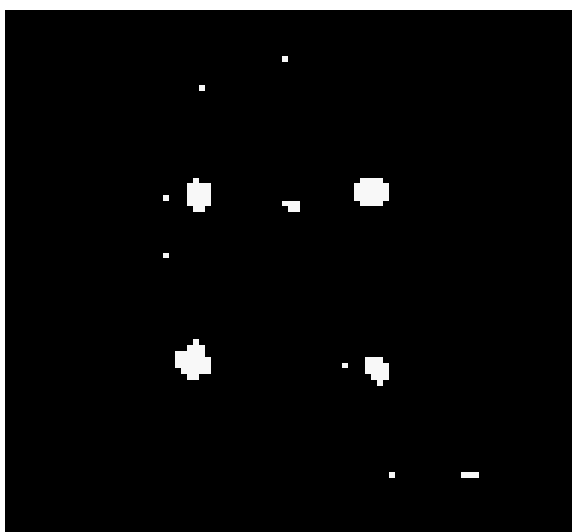
б



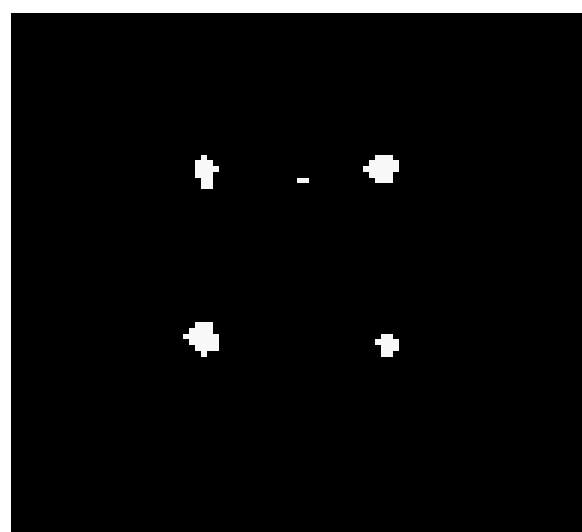
в



г



д



е

Рис. 7.1

объясняется суммированием по скользящему квадратному окну в (7.9). Обычно считается, что объект обнаружен, если имеется хотя бы одно превышение порога в области S_G расположения объекта.

7.3. Содержание отчета

1. Титульный лист.
2. Цель лабораторной работы.
3. Задание на работу.
4. Краткое изложение теоретической части.
5. Описание алгоритмов обнаружения.
6. Распечатка программ обнаружения.
7. Распечатка картины обнаружения (исходное И, остатки компенсации, статистика и ее сечение порогом).
8. Сводка характеристик обнаружения (зависимость количества обнаруженных сигналов от их яркости, то есть от коэффициента c в (4.6), и от вероятности ложной тревоги).
9. Выводы по работе.

7.4. Контрольные вопросы

1. В чем может состоять априорная неопределенность в задачах обнаружения сигналов?
2. Как преодолевается априорная неопределенность?
3. Адаптивная компенсация мешающих И.
4. Адаптивное определение порога.
Одноконтурный алгоритм и его недостатки.
Двухконтурный алгоритм.
5. Неопределенность описания обнаруживаемого сигнала и ее преодоление.

7.4. Задания на лабораторную работу

1. Произвести адаптивное обнаружение протяженного сигнала заданной формы на втором кадре заданной пары И. Порог определять адаптивно по заданной вероятности ложной тревоги.
2. Найти характеристики обнаружения (количество обнаруженных целей в зависимости от их яркости и вероятности ложной тревоги).

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Крашенинников, В. Р. Основы теории обработки изображений: учебное пособие. – Ульяновск: УлГТУ, 2003. – 152 с.
2. Васильев, К. К., Крашенинников, В. Р. Методы фильтрации многомерных случайных полей. – Саратов: СГУ, 1990. – 128 с.
3. Васильев, К. К., Крашенинников, В. Р. Представление и быстрая обработка многомерных изображений // Научно-технические технологии. – 2002. – Т. 3. – С. 4-24.
4. Методы компьютерной обработки изображений / под ред. В. А. Соффера. – М.: Физматлит, 2001. – 784 с.
5. Цифровая обработка изображений / под ред. А. А. Спектора. – Новосибирск: НГТУ, 2002. – 350 с.

Учебное издание

Крашенинников Виктор Ростиславович

Основы теории обработки изображений

Сборник лабораторных работ

Редактор Н. А. Евдокимова

Подписано в печать 20.10.2005. Формат 60x84/16. Бумага офсетная.

Печать трафаретная. Усл. печ. л. 2,09. Уч.-изд. л. 2.00.

Тираж 100 экз. Заказ .

Ульяновский государственный технический университет

432027, Ульяновск, Северный Венец, 32.

Типография УлГТУ, 432027, Ульяновск, Северный Венец, 32.