

## ИНФОРМАЦИОННАЯ ТЕХНОЛОГИЯ АНАЛИЗА КАЧЕСТВА СИСТЕМЫ ОБНАРУЖЕНИЯ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ

### 1. ВВЕДЕНИЕ

Как правило, качество системы обнаружения объектов на изображениях определяется по критериям, имеющим смысл вероятностей: ошибок обнаружения объектов, неправильной классификации изображения, недостижения системой цели своего функционирования и т.п. [1,2]. В интересных для практики ситуациях значения подобных показателей качества довольно малы, в связи с чем их экспериментальное измерение весьма затруднительно. Дополнительные сложности вызывает и высокая размерность воздействия на входе системы обнаружения, имеющей дело с изображениями. Ниже излагается метод проведения статистических экспериментов с системой обнаружения объектов на изображениях, позволяющий существенно сократить число экспериментов, необходимое для получения требуемой точности оценивания ее характеристик качества по сравнению с традиционно используемыми методами [3,4]. Отметим, что все изложенные результаты можно использовать и при оценке вероятностей редких событий, имеющих достаточно произвольный смысл.

### 2. ОЦЕНИВАНИЕ ХАРАКТЕРИСТИК КАЧЕСТВА СИСТЕМЫ ОБНАРУЖЕНИЯ ОБЪЕКТОВ С ПОМОЩЬЮ СТАТИСТИЧЕСКОГО МОДЕЛИРОВАНИЯ

Примем описание системы обнаружения в виде бинарной функции  $G(X)$  с аргументом  $X = (x_0, x_1, \dots, x_{L-1})$  -  $L$ -компонентным вектором случайного воздействия на ее входе. Эта функция принимает значение 0 или 1 при решении, соответственно, об отсутствии объекта во входном воздействии (гипотеза  $H_0$ ) или о его наличии (гипотеза  $H_1$ ). Качество системы характеризуется вероятностями ложного обнаружения -  $\alpha$  и пропуска объекта -  $\beta$  [2,4]:

$$\alpha = \int_{R^L} G(X) W(X/H_0) dX, \quad (1)$$

$$\beta = \int_{R^L} (1 - G(X)) W(X/H_1) dX, \quad (2)$$

где  $W(X/H_0)$  и  $W(X/H_1)$  - плотности распределения вероятностей входного вектора в случае справедливости гипотез  $H_0$  и  $H_1$  соответственно. Вероятности (1) и (2) по сути отличаются друг от друга лишь ситуацией на входе системы  $G$  и трактовкой ее отклика,

поэтому ниже для краткости ограничимся рассмотрением только одной из них, а именно,  $\alpha$ .

Стандартным методом оценивания неизвестной вероятности  $\alpha$  по результатам  $N$  независимых наблюдений над системой обнаружения объектов является метод прямого статистического моделирования [3,4,5,6], который заключается в вычислении оценки вида

$$\hat{\alpha} = \frac{1}{N} \sum_{n=0}^{N-1} G(X^{(n)}), \quad (3)$$

где  $X^{(n)}$  - реализация вектора  $X$  с плотностью вероятностей  $W(X/H_0)$  в  $n$ -ом эксперименте.

Оценка (3) является несмещенной, ее точность характеризуется шириной интервала  $[\alpha - d, \alpha + d]$ , в который она попадает с заданной доверительной вероятностью  $\gamma$ .

Выражение для величины  $d$  имеет вид [3,4,5]:

$$d = C_\gamma \sqrt{\frac{D_G}{N}}, \quad (4)$$

где  $D_G$  - дисперсия отклика системы  $G$ , равная здесь

$$D_G = \alpha(1 - \alpha); \quad (5)$$

$C_\gamma$  - коэффициент, определяемый по заданной  $\gamma$  с помощью функции Лапласа (обычно полагают  $C_\gamma = 2$  или  $C_\gamma = 3$ , что при  $N \gg 1$  соответствует доверительной вероятности  $\gamma \approx 0.95$  и  $\gamma \approx 0.997$ ). Из формул (4) и (5) следует выражение для относительной ошибки оценивания:

$$\delta = \frac{d}{\alpha} = C_\gamma \sqrt{\frac{1 - \alpha}{\alpha N}}. \quad (6)$$

Из формулы (6) видно, что при оценивании малых значений вероятности  $\alpha$  для получения высокой точности оценочная процедура (3) может потребовать очень большого числа экспериментов  $N$  с системой  $G$ . Улучшить точность оценки  $\hat{\alpha}$  при заданном  $N$  или, что эквивалентно, уменьшить число экспериментов при заданной точности позволяет применение методов понижения дисперсии при оценивании математического ожидания случайной величины. Одним из них является метод значимой выборки (МЗВ), который наиболее удобен в рассматриваемом случае [3,4,5,6].

Пусть все возможные входные векторы имеют плотности распределения вероятностей, различающиеся только значениями некоторого параметра  $\sigma$  (для случая оценивания вероятности  $\alpha$  это позволяет сделать переобозначение:  $W(X/H_0) = W(X;\sigma)$ ). Тогда суть МЗВ состоит в том, что входное воздействие с заданной плотностью  $W(X;\sigma)$  заменяется модифицированным, имеющим плотность  $W(X;\sigma_m)$ , а затем отклик системы  $G$  корректируется соответствующим весовым коэффициентом. Соотношение (1) можно переписать в другой форме:

$$\alpha = \int_{R^L} G(X) W(X;\sigma) dX = \int_{R^L} G(X) \frac{W(X;\sigma)}{W(X;\sigma_m)} W(X;\sigma_m) dX,$$

из которой следует, что при обработке  $N$  независимых наблюдений над системой наряду с оценкой (3) можно применять и оценку вида

$$\hat{\alpha} = \frac{1}{N} \sum_{n=0}^{N-1} G(X^{(n)}) \frac{W(X^{(n)}; \sigma)}{W(X^{(n)}; \sigma_m)}, \quad (7)$$

где  $X^{(n)}$  - реализация вектора  $X$  с модифицированной плотностью в  $n$ -ом эксперименте. Точность оценки (7) характеризуется показателем, аналогичным показателю (4):

$$d_m = C_\gamma \sqrt{\frac{D_Z}{N}}, \quad (8)$$

где  $D_Z$  - дисперсия случайной величины

$$Z = G(X) \frac{W(X; \sigma)}{W(X; \sigma_m)}.$$

Сопоставляя (4) и (8), видим, что при заданном  $N$  для повышения точности оценки (7) по сравнению с оценкой (3) необходимо выполнение неравенства

$$D_Z < D_G. \quad (9)$$

Это условие будет выполнено, если выбрать  $\sigma_m$  так, чтобы весовые коэффициенты  $W(X; \sigma)/W(X; \sigma_m)$ , используемые в оценке (7), были меньше единицы для тех  $X$ , при которых  $G(X)=1$ .

Если оценивать единственное значение  $\alpha$  для конкретного распределения вектора  $X$ , то можно найти оптимальное  $\sigma_m$ , минимизирующее дисперсию  $D_Z$ , и, соответственно, показатель (8) [7]. Однако МЗВ открывает и более широкую возможность снижения вычислительной сложности моделирования, позволяя оценивать сразу несколько значений и даже целую зависимость  $\alpha$  от варьируемого параметра распределения входного вектора на основе одних и тех же откликов системы  $G$  [8,9]. Действительно, если зафиксировать величину  $\sigma_m$ , при которой проводится моделирование, то после введения обозначений

$$k = \frac{\sigma}{\sigma_m}, \quad (10)$$

$$C(X, k) = \frac{W(X; k\sigma_m)}{W(X; \sigma_m)}, \quad (11)$$

из (7) получим оценку сразу всей функции  $\alpha(k)$ :

$$\hat{\alpha}(k) = \frac{1}{N} \sum_{n=0}^{N-1} G(X^{(n)}) C(X^{(n)}, k). \quad (12)$$

Ценность МЗВ состоит также и в том, что при выборе модифицированной плотности распределения входного вектора не нужна информация о процессах преобразования, осуществляемых системой. Поэтому к системе  $G$  можно отнести все

неизвестные и трудноформализуемые факторы, в том числе и формирование сложной модели наблюдения сигнала, действие шума и нелинейных искажений, процедуру принятия решения человеком-оператором. Единственным условием применимости МЗВ является наличие доступа к реализациям вектора  $X$  для их использования при расчете весовых коэффициентов в формулах (7) и (12). При проведении моделирования с применением синтезированных тестовых входных сигналов это условие всегда может быть выполнено.

Далее будем полагать, что входное воздействие  $X$  соответствует случайному шуму наблюдения изображения, на котором может находиться подлежащий обнаружению объект, а параметр  $\sigma$  имеет смысл СКО (среднеквадратическое отклонение) шума. В таком случае использование МЗВ при проведении экспериментов по обнаружению заключается в увеличении  $\sigma$  до некоторого значения  $\sigma_m$ , т.е. в переходе к "плохим" условиям наблюдения. Затем по формулам (11), (12) производится пересчет получаемых решений для "лучших" условий (для  $\sigma < \sigma_m$  или, по (10), для  $k < 1$ ). В результате получается оценка зависимости искомой вероятности от уровня шума. "Базовое" значение СКО  $\sigma_m$ , являющееся максимальным из рассматриваемых, следует здесь выбирать так, чтобы ему соответствовала настолько большая вероятность  $\alpha(1)$ , что ее можно было бы достаточно точно оценивать с помощью прямого статистического моделирования при малом  $N$ .

### 3. МОДИФИКАЦИЯ МЕТОДА ЗНАЧИМОЙ ВЫБОРКИ ДЛЯ МНОГОМЕРНОГО ВХОДНОГО ВОЗДЕЙСТВИЯ

В реальных задачах обнаружения объектов на изображениях входное воздействие является существенно многомерным. При этом на решение, принимаемое системой обнаружения, могут не влиять (или слабо влиять) многие компоненты вектора  $X$  (назовем такие компоненты "фиктивными"). Рост размерности  $X$  и появление в нем фиктивных компонент быстро снижает эффективность описанной выше схемы МЗВ. Причина этого ухудшения заключается в появлении рассогласования между весовыми коэффициентами в формулах (7), (12) и откликами системы: из-за дополнительных фиктивных компонент при  $G(X)=1$  весовые коэффициенты принимают значения как меньше, так и больше единицы, что в итоге приводит к нарушению условия (9).

Можно избежать негативных последствий такой ситуации, если произвести дополнительную рандомизацию входного воздействия путем его домножения на независимую неотрицательную случайную величину  $\xi$  с плотностью распределения вероятностей  $W_\xi(\xi)$ . Эта рандомизация позволяет свести многомерный случай к одномерному и использовать МЗВ в скалярном варианте, модифицируя плотность не вектора  $X$ , а величины  $\xi$  [8,9].

При поступлении произведения  $X\xi$  на вход системы вместо вероятности ложного обнаружения (1) имеем:

$$g = \int_{R^L} \int_0^\infty G(X\xi) W(X/H_0) W_\xi(\xi) dX d\xi. \quad (13)$$

Положим, что СКО компонент вектора  $X\xi$  близко к  $\sigma_m$ , т.е. реализуются плохие условия наблюдения, при которых вероятность (13) достаточно велика и удовлетворительно оценивается с помощью прямого статистического моделирования. Следуя схеме МЗВ, воспользуемся результатами этого же моделирования для оценивания меньших значений  $g$ , соответствующих лучшим условиям наблюдения. Для этого произведем модификацию входного воздействия, состоящую в умножении его СКО на положительный коэффициент  $k < 1$ . Такая модификация эквивалентна умножению на  $k$  только величины  $\xi$ . При этом вероятность (13), зависящая от параметра  $k$ , будет равна

$$g(k) = \int_{R^L} \int_0^\infty G(Xk\xi) W(X/H_0) W_\xi(\xi) dX d\xi = \int_{R^L} \int_0^\infty G(Xy) W(X/H_0) \frac{1}{k} W_\xi\left(\frac{y}{k}\right) dX dy, \quad (14)$$

где  $y = k\xi$  - новая переменная интегрирования. Из последнего выражения имеем два следствия. Во-первых, ему тождественна запись

$$g(k) = \int_{R^L} \int_0^\infty G(Xy) \frac{\frac{1}{k} W_\xi\left(\frac{y}{k}\right)}{W_\xi(y)} W(X/H_0) W_\xi(y) dX dy,$$

из которой следует, что в качестве оценки вероятности (14) можно использовать величину

$$\hat{g}(k) = \frac{1}{N} \sum_{n=0}^{N-1} G(X^{(n)}\xi_n) C(\xi_n, k), \quad (15)$$

где  $X^{(n)}$ ,  $\xi_n$  - реализации случайного вектора  $X$  и случайной величины  $\xi$  в  $n$ -ом эксперименте, а

$$C(\xi, k) = \frac{\frac{1}{k} W_\xi\left(\frac{\xi}{k}\right)}{W_\xi(\xi)} \quad (16)$$

- весовой коэффициент при усреднении откликов системы  $G$ . Во-вторых, выражение (14) может быть переписано в виде

$$g(k) = \int_0^\infty \alpha(y) \frac{1}{k} W_\xi\left(\frac{y}{k}\right) dy, \quad (17)$$

где

$$\alpha(y) = \int_{R^L} G(Xy) W(X/H_0) dX$$

- вероятность ложного обнаружения при модификации входного вектора  $X$  умножением на коэффициент  $y$ . Согласно п.2, значения функции  $\alpha(y)$  при  $0 < y < 1$  являются искомыми величинами, которые требуется определить по наблюдениям за системой  $G$ . Для оценивания этой зависимости нужно решить интегральное уравнение (17), заменив его точную левую часть на оценку (15). В общем случае для приближенного решения интегрального уравнения можно использовать итерационные методы [10] с естественными для  $\alpha(y)$  ограничениями на положительность и монотонное возрастание.

Однако, как будет показано ниже, существуют ситуации, когда можно обойтись без применения трудоемких итерационных процедур.

Для осуществления предлагаемой рандомизации входного воздействия и решения интегрального уравнения (17) необходимо выбрать конкретный вид плотности распределения вероятностей  $W_{\xi}(\xi)$ . При выборе этой плотности следует в первую очередь учитывать ее влияние на качество оценки (15): при единичных откликах системы  $G$  весовые коэффициенты (16) должны быть по возможности минимальными. Кроме того, желательно, чтобы выбранная плотность упрощала решение уравнения (17). Исходя из этих соображений, предлагается использовать для рандомизации случайную величину  $\xi$  с плотностью вида

$$W_{\xi}(\xi) = A \begin{cases} \xi^{a-1} & \text{при } 0 \leq \xi < 1, \\ \xi^{b-1} & \text{при } \xi \geq 1, \end{cases} \quad (18)$$

где  $a > 0$ ,  $b < 0$  - параметры распределения;

$$A = \frac{ab}{(b-a)}$$

- нормировочная константа.

Подставив (18) в уравнение (17), имеем:

$$g(k) = I_1(k) + I_2(k), \quad (19)$$

где

$$I_1(k) = Ak^{-a} \int_0^k \alpha(y) y^{a-1} dy, \quad I_2(k) = Ak^{-b} \int_k^{\infty} \alpha(y) y^{b-1} dy.$$

Для первых двух производных функции (19) имеем выражения:

$$g'(k) = -\frac{a}{k} I_1(k) - \frac{b}{k} I_2(k), \quad (20)$$

$$g''(k) = \frac{a(a+1)}{k^2} I_1(k) + \frac{b(b+1)}{k^2} I_2(k) + \frac{ab}{k^2} \alpha(k). \quad (21)$$

Формулы (19)-(21) могут рассматриваться как система трех линейных уравнений относительно  $I_1(k)$ ,  $I_2(k)$  и  $\alpha(k)$ . Из этой системы, в частности, находим:

$$\alpha(k) = L(g(k)) = g(k) + \frac{a+b+1}{ab} kg'(k) + \frac{1}{ab} k^2 g''(k). \quad (22)$$

Выражение (22) представляет собой решение интегрального уравнения (17), записанное в явной форме.

После проведения статистического моделирования для получения оценки  $\hat{\alpha}(k)$  следует подставить в (22) значения оценок  $\hat{g}(k)$ ,  $\hat{g}'(k)$  и  $\hat{g}''(k)$ . Оценка  $\hat{g}(k)$  здесь вычисляется по формуле (15) с весовыми коэффициентами, определяемыми из (16), (18):

$$C(\xi, k) = \begin{cases} k^{-a} & \text{при } 0 \leq \xi < k, \\ k^{-b} \xi^{b-a} & \text{при } k \leq \xi < 1, \\ k^{-b} & \text{при } \xi \geq 1. \end{cases} \quad (23)$$

Хорошие же оценки производных найти существенно труднее, поскольку весовые коэффициенты (23) и, следовательно, оценка  $\hat{g}(k)$  не являются дифференцируемыми по  $k$  функциями (хотя истинная функция  $g(k)$  - гладкая). В связи с этим перед численным дифференцированием следует аппроксимировать оценку  $\hat{g}(k)$  какой-либо достаточно гладкой функцией. Можно осуществить такое сглаживание с помощью кубических сплайнов, являющихся дважды непрерывно дифференцируемыми функциями в области своего определения [11]. Пусть  $g_s(k)$  - результат сплайн-аппроксимации функции  $\hat{g}(k)$ , а  $\hat{g}'(k)$  - оценка производной  $g'(k)$ , полученная численным дифференцированием функции  $g_s(k)$ . Для получения более гладкой зависимости функции  $\hat{g}'(k)$  от  $k$  ее также можно сгладить кубическим сплайном. Результат этого сглаживания обозначим  $g'_s(k)$ . Действуя аналогичным образом, определяем оценку второй производной -  $\hat{g}''(k)$  (дифференцированием  $g'_s(k)$ ) и ее сглаженный вариант -  $g''_s(k)$ . После этого получаем возможность вычислить и искомую оценку вероятности ложного обнаружения:

$$\hat{\alpha}(k) = L(g_s(k)) = g_s(k) + \frac{a+b+1}{ab} k g'_s(k) + \frac{1}{ab} k^2 g''_s(k). \quad (24)$$

Указанные два метода нахождения оценки вероятности ложного обнаружения  $\alpha(k)$  (из интегрального уравнения (17) и из уравнения (24)) являются весьма трудоемкими при реализации и требуют оптимизации большого количества параметров. Предпочтительнее в этом смысле подход, состоящий в следующем.

Если система обнаружения  $G$  является пороговой, а входные воздействия стандартными гауссовскими, то вероятность ложного обнаружения вычисляется аналитически:

$$\alpha_c(k) = 1 - \Phi(c/k),$$

где  $\Phi(x)$  - функция Лапласа,  $c$  - порог. В этом случае в соответствии с (17) легко рассчитывается и зависимость  $g_c(k)$ . Считая данные зависимости типичными, для каждой реализации оценки  $\hat{g}(k)$  и начального порогового значения  $c_0$  можно найти такое значение порога  $c$ , при котором  $g_c(k)$  будет наиболее близка к  $\hat{g}(k)$ . В случае среднеквадратического приближения для этого необходимо:

- для некоторого множества  $k_1, \dots, k_r$  значений аргумента  $k$  определить величины  $\Delta k_1, \dots, \Delta k_r$  таким образом, чтобы

$$g_{c_0}(k + \Delta k_i) = \hat{g}(k_i), \quad i = 1, \dots, r;$$

- вычислить масштабирующий коэффициент  $s$  по формуле:

$$s = \frac{\sum_{i=1}^r k_i (k_i + \Delta k_i)}{\sum_{i=1}^r (k_i + \Delta k_i)^2}; \quad (25)$$

- положить значение порога  $c = c_0/s$ .

Обозначим  $\Delta\hat{g}(k) = \hat{g}(k) - g_c(k)$  - погрешность приближения. Тогда из уравнения (22)

будем иметь:

$$\hat{\alpha}(k) = L(\hat{g}(k)) = L(g_c(k)) + L(\Delta\hat{g}(k)). \quad (26)$$

Поскольку  $L(g_c(k)) = \alpha_c(k)$ ,  $L(\Delta\hat{g}(k)) = \hat{\alpha}(k) - \alpha_c(k) = \Delta\hat{\alpha}(k)$ , то можно считать, что первое слагаемое в (26) представляет собой "опорное" решение, а второе - поправку к нему. В этом и состоит смысл предлагаемого метода оценивания зависимости  $\alpha(k)$ .

Для того, чтобы найти  $L(\Delta\hat{g}(k))$ , необходимо произвести сглаживание функции  $\Delta\hat{g}(k)$  ( $\Delta\hat{g}(k)$ , как и  $\hat{g}(k)$  дифференцируемой не является). Однако для этого можно не прибегать к кубической сплайн-аппроксимации. В силу малости значений  $\Delta\hat{g}(k)$  достаточную точность оценивания обеспечивает просто линейное сглаживание.

Нетрудно видеть, что изложенный подход к оцениванию  $\alpha(k)$  практически означает аппроксимацию системы обнаружения  $G$  неизвестной структуры пороговой системой с гауссовским воздействием на входе. Он не является таким общим, как два рассмотренных ранее, но для большинства систем обнаружения его использование вполне оправдано высокой вычислительной эффективностью и достигаемой точностью получаемых оценок.

#### 4. РЕЗУЛЬТАТЫ МОДЕЛИРОВАНИЯ

Сначала продемонстрируем эффективность МЗВ на примере оценивания вероятности ложного обнаружения в случае скалярной аддитивной модели наблюдения "объект плюс шум". Предположим, что при отсутствии объекта все возможные воздействия  $X$  на входе системы  $G$  являются гауссовскими случайными величинами с нулевыми математическими ожиданиями и могут отличаться лишь значениями СКО  $\sigma$ . Подставив выражение для гауссовской плотности вероятностей в формулу (11), получим конкретизацию выражения для весовых коэффициентов:

$$C(X, k) = \frac{1}{k} \exp\left[-\left(\frac{1}{k^2} - 1\right) \frac{X^2}{2\sigma_m^2}\right]. \quad (27)$$

"Базовое" значение СКО  $\sigma_m$ , соответствующее синтезируемым входным воздействиям и используемое в формуле (27), для определенности положим равным единице. В экспериментах по оцениванию зависимости  $\alpha(k)$  рассмотрим простейшую систему  $G$ , представляющую собой пороговое устройство:

$$G(X) = \begin{cases} 1 & \text{при } X \geq c, \\ 0 & \text{при } X < c, \end{cases} \quad (28)$$

где значение порога  $c$  установим так, чтобы при тестовом входном воздействии вероятность ложного обнаружения была равна 0.2:

$$\sigma_m = 1, c \approx 0.842, \alpha(1) = 0.2.$$

Для сформулированных условий статистическое моделирование заключалось в следующем. По  $N=100$  откликам системы осуществлялось вычисление оценок  $\hat{\alpha}(k)$  по



формуле (12). Параллельно, с учетом (27) и (28), рассчитывались истинные вероятности  $\alpha(k)$  для  $0 < k \leq 1$ . Для получения статистически устойчивых выводов было выполнено

$M=50$  серий экспериментов и вычислены:

- выборочная дисперсия оценки  $\alpha(k)$ :

$$D_{\hat{\alpha}}(k) = \frac{1}{M} \sum_{j=0}^{M-1} \hat{\alpha}_j^2(k) - \alpha^2(k), \quad (29)$$

где  $\hat{\alpha}_j(k)$  - оценка  $\alpha(k)$  в  $j$ -ой серии;

- размах ошибки оценивания, определенный с доверительной вероятностью  $\gamma \approx 0.95$  (экспериментальный аналог показателя (8) при  $C_\gamma = 2$ ):

$$\hat{d}_M(k) = 2\sqrt{D_{\hat{\alpha}}(k)}; \quad (30)$$

- относительная погрешность оценивания:

$$\hat{\delta}_M(k) = \frac{\hat{d}_M(k)}{\alpha(k)}. \quad (31)$$

Чтобы сравнить МЗВ с методом прямого статистического моделирования, для последнего по формуле (6) при  $C_\gamma = 2$  вычислялись теоретические относительные погрешности оценки вероятностей  $\alpha(k)$ .

Результаты экспериментов и расчетов представлены на рис.1, 2. На рис.1 изображены также графики двух конкретных оценок  $\hat{\alpha}_1(k)$  и  $\hat{\alpha}_2(k)$ , полученных по 100 экспериментам. Анализ результатов позволяет сделать вывод о том, что оценка  $\hat{\alpha}(k)$ , вычисленная по формулам (12), (25), при всех расчетных значениях  $k$  оказывается точнее оценки, полученной прямым моделированием.

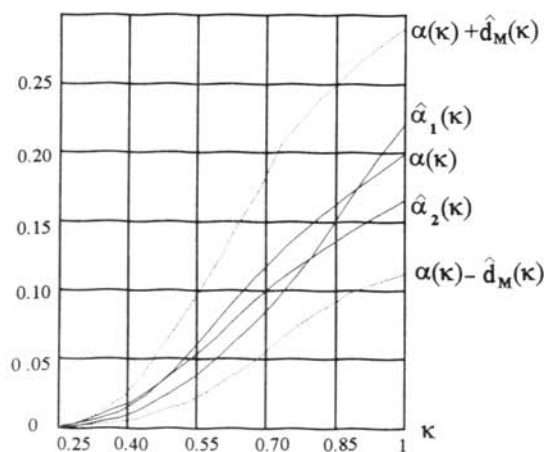


Рис.1

Результаты оценивания вероятности  $\alpha(k)$  в скалярном случае при  $N=100$ :

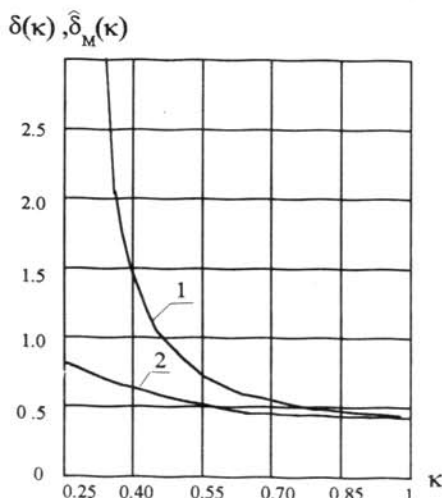


Рис.2

Относительная погрешность оценки  $\hat{\alpha}(k)$  в скалярном случае при  $N=100$ :

1 - прямое статистическое моделирование; 2 - МЗВ.

Этот выигрыш особенно велик при малых значениях оцениваемой вероятности. Заметим, что по результатам  $N=100$  экспериментов над системой  $G$  оценивание вероятностей, меньших  $0.01$ , с помощью прямого моделирования практически невозможно, в то время как МЗВ дает вполне удовлетворительный ответ.

Вместе с тем с увеличением размерности входного вектора точность МЗВ быстро снижается и становится значительно хуже, чем при прямом моделировании. Данное обстоятельство и стимулировало разработку модификаций МЗВ, эффективных в случаях многомерных входных воздействий.

Покажем, какие результаты оценивания вероятности  $\alpha(k)$  дает описанный в п.3 подход, основанный на аппроксимации системы обнаружения пороговой системой. Для использования этого подхода нужно решить следующие вспомогательные задачи:

- выбрать значения параметров  $a$  и  $b$  плотности вероятностей (18) рандомизирующей случайной величины  $\xi$ ;
- определить начальное значение порога  $c_0$ ;
- определить число и расположение значений аргумента  $k$  в интервале  $[0,1]$ , используемых для расчета "опорного" решения  $\alpha_c(k)$  с учетом значений оценки  $\hat{g}(k)$ .

Поиск наилучшего (по критерию точности оценки) решения этих задач требует проведения большого объема экспериментальных исследований. Ниже мы ограничимся подбором требуемых величин, обеспечивающих удовлетворительные результаты оценивания для рассматриваемого демонстрационного примера.

Пусть для обнаружения объекта на изображении применяется корреляционный алгоритм [1,2], действие которого заключается во взвешенном суммировании отсчетов функции яркости (компонент многомерного входного сигнала) и в сравнении суммы с порогом. Очевидно, рандомизация многомерного входного воздействия случайной величиной  $\xi$  эквивалентна рандомизации суммарного сигнала. Следовательно, в данном случае можно упростить вычислительные эксперименты с системой обнаружения, рассматривая ее в простейшей форме (28) со скалярным сигналом  $X$  на входе.

Для входного воздействия  $X$ , имеющего гауссовское распределение, моделирование процедуры оценивания зависимости  $\alpha(k)$  по модифицированной схеме МЗВ с дополнительной рандомизацией состояло в следующем. Сигнал  $X$  умножался на случайную величину  $\xi$  с распределением (18), параметры  $a$  и  $b$  которого предполагались связанными соотношением:

$$a + b + 1 = 0. \quad (32)$$

Условие (32) равносильно тому, что  $M\xi = 1$ , и оно гарантирует равенство средних значений входного вектора до рандомизации и после нее:  $MX\xi = MXM\xi = MX$ . Подобное предположение является вполне оправданным. Кроме того, оно существенно упрощает нахождение оптимальных значений параметров  $a$  и  $b$  и, в частности, приводит к отсутствию в уравнении (22) второго слагаемого.

Далее произведение  $X\xi$  подавалось на вход системы (28), имеющей тот же порог обнаружения, что и выше. По результатам  $N=500$  экспериментов с помощью формул (15) и (23) строилась зависимость  $\hat{g}(k)$ . Начальное значение порога  $c_0$  при построении "опорного" решения выбиралось из условия:

$$g_{c_0}(1) \approx \hat{g}(1).$$

Затем производилась коррекция  $g_{c_0}(k)$ , основанная на значениях оценки  $\hat{g}(k)$  в трех точках:  $k_1 = 0.25$ ,  $k_2 = 0.65$ ,  $k_3 = 1$  (см. формулу (25)). Строилось "опорное" решение  $\alpha_c(k)$ , а поправка к нему с учетом условия (32) и в предположении линейного сглаживания погрешности  $\Delta\hat{g}(k)$  полагалась равной

$$\Delta\hat{\alpha}(k) = L(\Delta\hat{g}(k)) = \Delta\hat{g}(k).$$

Окончательно искомая оценка  $\hat{\alpha}(k)$  вычислялась по формуле:

$$\hat{\alpha}(k) = \alpha_c(k) + \Delta\hat{g}(k).$$

Всего было выполнено  $M=50$  серий экспериментов. После этого были вычислены характеристики точности оценивания (29)-(31) и установлено, что наилучшие показатели этих характеристик соответствуют значениям параметров  $a=0.875$ ,  $b=-1.875$ . На рис. 3, 4 представлены полученные при этом результаты оценивания.

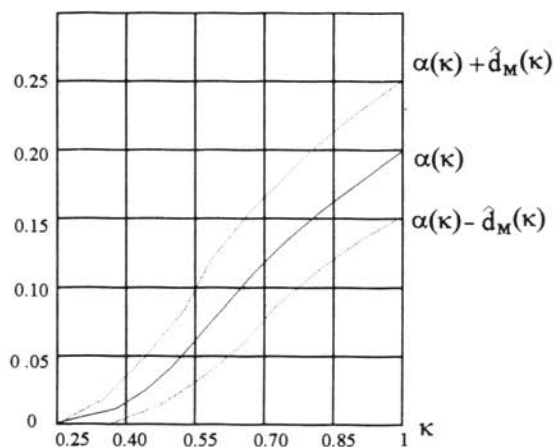


Рис.3

Диапазон оценок зависимости  $\alpha(k)$  для модифицированного МЗВ ( $N=500$ ).

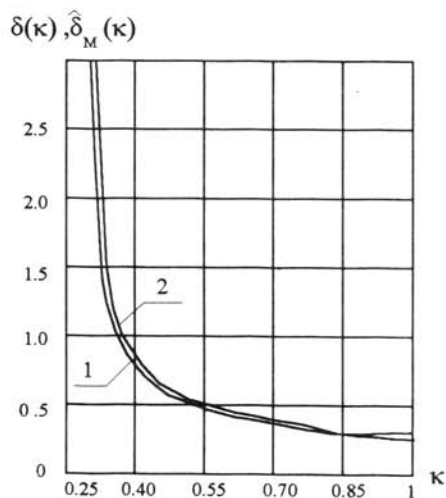


Рис.4

Относительная погрешность оценки  $\hat{\alpha}(k)$ :

- 1 - прямое статистическое моделирование ( $N=250$ );
- 2 - модифицированный МЗВ ( $N=500$ ).

Анализ зависимостей на рис.4 показывает, что относительная погрешность оценивания здесь близка к той, которая получается для прямого статистического моделирования при  $N=250$ , т.е. модифицированный МЗВ проигрывает при оценке конкретного значения  $\alpha$ . Однако этот проигрыш, обусловленный дополнительной рандомизацией, не зависит от размерности входного воздействия и может полностью компенсироваться возможностью оценивания всей функции  $\alpha(k)$  на основе одного цикла моделирования системы обнаружения. Данное обстоятельство свидетельствует о

целесообразности применения предложенной процедуры оценивания, базирующейся на модифицированном МЗВ, для проведения статистических экспериментов с системой обнаружения объектов на изображениях.

## 5. ЗАКЛЮЧЕНИЕ

Для проведения статистических экспериментов по оцениванию малых вероятностей ошибок перспективен метод значимой выборки, позволяющий сократить необходимое число экспериментов с системой обнаружения. Однако стандартная схема МЗВ неработоспособна при многомерном сигнале на входе системы. Предложена и экспериментально исследована информационная технология проведения статистических экспериментов, основанная на схеме МЗВ, модифицированной для многомерного входного воздействия. Эффективность технологии обусловлена возможностью получения всей зависимости оцениваемой вероятности от варьируемого параметра распределения сигнала за один цикл моделирования.

## Л и т е р а т у р а

1. *Прэнтт У.К.* Цифровая обработка изображений. - М.: Мир, 1982. - Кн.1, Кн.2. - 312 с., 480 с.
2. *Ярославский Л.П.* Цифровая обработка сигналов в оптике и голографии: Введение в цифровую оптику. - М.: Радио и связь, 1987. - 296 с.
3. *Ермаков С.М., Михайлов Г.А.* Статистическое моделирование. - М.: Наука, 1982. - 296 с.
4. *Клейнен Дж.* Статистические методы в имитационном моделировании. - М.: Статистика, 1978. - Том 1. - 223 с.
5. *Cheng R.C.H.* Variance Reduction Methods // Winter Simulation Conference, Washington, D.C., Dec.8-10, 1986: Proceedings. - 1986. - pp.60-68.
6. *Wilson J.P.* Variance Reduction Techniques for Digital Simulation // American Journal of Mathematical and Management Sciences. - 1984. - N 3-4. - pp.277-312.
7. *Кузьмин С.З., Костина С.С.* Применение значимой выборки для исследования систем обнаружения сигналов методом имитационного моделирования // Известия вузов. Радиоэлектроника. - 1984. - N 9. - С.8-14.
8. *Коломиец Э.И., Сергеев В.В.* Метод значимой выборки в задаче исследования процедуры обнаружения объектов // Тезисы докладов IV Всесоюзной конференции "Математические методы распознавания образов" (ММРО-IV). - Рига, 1989. - Часть 2. - С.69-70.
9. *Белюсова М.П., Коломиец Э.И., Сергеев В.В.* Исследование качества функционирования системы обнаружения с помощью имитационного моделирования // "Проблемы информатики": Тезисы докладов научно-

практической конференции с международным участием. - Самара, 1991. - С.77-78.

10. *Стечкин С.Б., Субботин Ю.Н.* Слайды в вычислительной математике. - М.: Наука, 1976. - 248 с.

11. *Вайникко Г.М., Веретенников А.Ю.* Итерационные процедуры в некорректных задачах. - М.: Наука, 1986. - 184 с.

**МЕЖДУНАРОДНЫЙ ЦЕНТР  
НАУЧНОЙ И ТЕХНИЧЕСКОЙ ИНФОРМАЦИИ**  
предлагает специальный номер журнала

**"ПРОБЛЕМЫ МАШИНОСТРОЕНИЯ И АВТОМАТИЗАЦИИ"**

№ 5-6/1994 г.,

подготовленный совместно с финскими фирмами

**"ТЕХНОЛОГИИ И ОБОРУДОВАНИЕ ФИНСКИХ ФИРМ"**

В этом номере представлены фирмы **MESAMAT**, **HOMMA ARCTIC**, **KEVYTHALLI**, **RAUMASTER**, **KUITUMYYNTI**, **SAIMATEC**, **REINLUFT**, которые являются типичными представителями малого и среднего бизнеса Финляндии.

**MESAMAT OY** – это автоматизированная технология и оборудование для производства бревенчатых стеновых блоков промышленным способом с сохранением естественной формы бревен. Базовая разработка фирмы – автоматизированная линия МЕКАЛОГ 640. Линия МЕКАЛОГ 640 особенно хорошо подходит для условий России.

**HOMMA ARCTIC OY**, на заводе которой в Финляндии установлена автоматизированная линия МЕКАЛОГ 640, представляет технологию сборки бревенчатых домов из готовых стеновых блоков, изготовленных на указанной линии.

**KEVYTHALLI OY** представляет рентабельное, многоцелевое, логическое решение **VINYHALL** для производства арочных конструкций из дерева (типа ангаров) различного назначения.

**RAUMASTER OY** специализируется на автоматизированных технологиях транспортировки материалов, особенно для лесоперерабатывающей промышленности.

**KUITUMYYNTI OY** представляет теплоизоляционную ленту **RIVE-LINE** для бревенчатых домов, реконструируемых зданий и новостроек.

**SAIMATEC ENGINEERING OY** – это специализированное инженерное бюро, занимающееся машиностроением в области лесной промышленности. Основной деятельностью фирмы является проектирование конструкций и автоматизации машинного оборудования, а также техническая реализация систем "под ключ".

**REINLUFT OY** представляет оборудование для сушки древесины.

Технологии и оборудование, разработанные финскими фирмами, запатентованы в различных странах и широко известны на западном рынке.

Целью выпуска специального номера журнала "Проблемы машиностроения и автоматизации" является представление продукции этих фирм на рынке России и других стран СНГ.

**Цена специального номера 5 USD.**

# Information technology for analyzing the quality of the system for detecting objects in the images

*E.I. Kolomiets, V.V. Sergeev*

## *Abstract*

As a rule, the quality of the system for detecting objects in the images is determined by criteria that have the meaning of probabilities: errors in detecting objects, incorrect classification of the image, the system's failure to achieve its purpose, etc. [1,2]. In situations of practical interest, the values of such quality indicators are rather small, and therefore their experimental measurement is very complicated. Additional difficulties are caused by the high dimension of the impact at the input of the detection system dealing with images. Below we describe the method of conducting statistical experiments with the system for detecting objects in images, which allows to reduce significantly the number of experiments needed to obtain the required accuracy of determining its quality characteristics as compared with the traditional methods [3,4]. We should note that all of the above results can be used to assess the probabilities of rare events that have a fairly arbitrary meaning.

*Citation:* Kolomiets EI, Sergeev VV. Information technology for analyzing the quality of the system for detecting objects in the images. Computer Optics 1995; 14-15(1): 166-178.

## *References*

- [1] Pratt WK, Digital Image Processing; Moscow: Mir Publisher; 1982; 1: 312; 2: 480.
- [2] Yaroslavsky LP. Digital signal processing in optics and holography: Introduction to digital optics; Moscow: Radio i svyaz; 1987; 296.
- [3] Ermakov SM, Mikhailov GA. Statistical Modeling; Moscow: Nauka; 1982; 296.
- [4] Kleinen J. Statistical methods in simulation modeling; Moscow: Statistics; 1978; 1: 223.
- [5] Cheng RCH. Variance Reduction Methods. Winter Simulation Conference. Washington D.C.: 1986; Proceedings; 60-68.
- [6] Wilson JP. Variance Reduction Techniques for Digital Simulation. American Journal of Mathematical and Management Sciences; 1984; 3-4: 277-312.
- [7] Kuzmin SZ, Kostina SS. Application of the significant-sample method to the study of signal-detection systems on the basis of simulation. News of universities. Radioelektronika; 1984; 9: 8-14.
- [8] Kolomiets EI, Sergeev VV. The method of significant sampling in the task of studying the procedure for detecting objects. Abstracts of the IV All-Union Conference "Mathematical Methods of Pattern Recognition"; Riga: 1989; 2: 69-70.
- [9] Belousova MP, Kolomiets EI, Sergeev VV. Study of the detection system performance using simulation modeling. Problems of Informatics: Abstracts of the scientific-practical conference with international participation; Samara: 1991; 77-78.
- [10] Stechkin SB, Subbotin YN. Splines in computational mathematics. - Moscow: Nauka; 1976; 248.
- [11] Vainikko GM, Veretennikov AY. Iterative procedures in ill-posed problems; Moscow: Nauka; 1986; 184.