

МАТЕМАТИЧЕСКИЕ МЕТОДЫ СТЕГАНОГРАФИИ

УДК 004.94

МОДИФИКАЦИЯ КАРТЫ ХОТЕЛЛИНГА, НИВЕЛИРУЮЩАЯ
ВЛИЯНИЕ ТРЕНДА, И ЕЕ ПРИМЕНЕНИЕ ПРИ ОБНАРУЖЕНИИ
ЦИФРОВЫХ ВОДЯНЫХ ЗНАКОВ

Б. Б. Борисенко

*Институт проблем информационной безопасности МГУ, г. Москва, Россия***E-mail:** fepem@yandex.ru

Предлагается использовать модифицированную статистику Хотеллинга в контрольной карте с целью исключения влияния тренда. Рассматривается вариант применения метода контрольных карт в задаче обнаружения наличия цифровых водяных знаков.

Ключевые слова: компьютерная безопасность, цифровые водяные знаки, карты контроля качества.

Для защиты авторских прав на электронные документы используются цифровые водяные знаки (ЦВЗ). Под ЦВЗ принято понимать специальную метку, встраиваемую в цифровой контент с целью аутентификации и защиты от копирования. Большинство современных методов обнаружения ЦВЗ используют в своей работе принцип обучения [1 – 7]. При этом обучение проводится как на выборке пустых (немаркированных) контейнеров, так и на выборке контейнеров с ЦВЗ. И те и другие генерируются аналитиком самостоятельно. Для создания пустых контейнеров целесообразно использовать фотоаппараты без функции внесения собственного ЦВЗ. Для классификации контейнеров используется, как правило, дискриминантный анализ (ДА) или метод опорных векторов (SVM — Support Vector Machine). Однако для построения выборки контейнеров с ЦВЗ необходимо знать конкретный метод внедрения ЦВЗ (или хотя бы класс методов), что не всегда является возможным.

Предлагаемый метод обнаружения ЦВЗ не требует априорного знания метода внедрения цифровых водяных знаков. Обучение производится только на выборке пустых контейнеров. При этом для каждого тестируемого контейнера на обучающую выборку накладываются определенные требования. Так, например, для изображений объекты обучающей выборки должны максимально походить на тестируемый контейнер, то есть принадлежать к одному типу изображений (пейзаж, средний план, поле), иметь примерно одинаковую гистограмму цветов и одинаковый коэффициент качества (процент сжатия). Поэтому особенностью предлагаемого метода является необходимость создания обучающей выборки в худшем случае для каждого тестируемого мультимедиа-контейнера (если тестируемые контейнеры одного типа, можно использовать одну и ту же обучающую выборку).

Метод основан на отслеживании поведения некоторых заранее определенных признаков контейнеров. От выбора конкретного набора признаков зависит результат работы метода (большая вероятность обнаружения ЦВЗ, меньшая вероятность ошибок

первого и второго рода), но не принцип работы самого метода. Набор признаков выбирается в зависимости от типа (графика, видео, звук) и формата контейнера.

Суть предлагаемого метода заключается в применении элементов метода контрольных карт [8, 9], используемого для выявления производственных сбоев, адаптированного под использование задач обнаружения цифровых водяных знаков. При этом:

- на основе обучающей выборки пустых контейнеров вычисляются контрольные границы для классификации. Пороги находятся из распределения Фишера или подбором по методу Монте-Карло;
- на основе признаков тестируемого контейнера и обучающей выборки пустых контейнеров вычисляется значение модифицированной статистики Хотеллинга для последующей классификации;
- значение статистики сравнивается с вычисленными порогами и выдается результат о наличии или отсутствии ЦВЗ.

Рассмотрим подробнее этапы предлагаемого метода, а также возможность проведения анализа без обучения на конкретном методе сокрытия. Для лучшего понимания приведем основы применения контрольных карт в производстве.

1. Методы статистического контроля технологического процесса

В связи с существенным ростом актуальности ряда практических задач, таких, как автоматическое обнаружение неисправностей, обслуживание оборудования на основе автоматического контроля его состояния, обеспечение безопасности сложных технических и информационных систем, автоматический контроль качества выпускаемой продукции, предсказание естественных катастрофических явлений, мониторинг в биомедицине и финансовой сфере, растет и необходимость выявления момента резкого изменения (*разладки*) некоторых вероятностных характеристик (признаков) у наблюдаемых процессов.

Статистический контроль технологического процесса представляет задачу последовательного обнаружения нарушений в независимой случайной последовательности при неизвестном априорном распределении момента времени, в который произошло нарушение. Вмешательство в технологический процесс для настройки требуется тогда, когда изделие еще удовлетворяет требованиям, но статистические показатели процесса свидетельствуют о наличии неслучайных воздействий. Практический инструмент для решения вопроса о необходимости такого вмешательства — *контрольная карта* — был предложен У. Шухартом [8, 9]: сигнал о разладке процесса подается при выходе контролируемого показателя за некоторую пороговую границу.

Задача последовательного обнаружения изменения среднего уровня технологического процесса для одномерной независимой гауссовской случайной последовательности рассмотрена в работах Е. Пейджа [10, 11]; на основе методов последовательного анализа была предложена контрольная карта кумулятивных сумм (CUSUM — Cumulative Sums). С. Робертсом [12, 13] предложено использование экспоненциального сглаживания для обнаружения нарушений в ходе процесса (метод экспоненциально взвешенных скользящих средних, EWMA — Exponentially Weighted Moving Average). Задача о статистическом контроле процесса при известных вероятностях перехода из налаженного состояния в разлаженное исследована А. Н. Ширяевым [14, 15].

Таким образом, используется три основных подхода к решению задачи статистического контроля процесса и различные их модификации:

- контрольная карта Шухарта (базируется на критерии Неймана — Пирсона);

- контрольные карты кумулятивных сумм (многократное применение последовательного анализа Вальда [16]);
- контрольные карты экспоненциально взвешенных скользящих средних (экспоненциальное сглаживание).

Обобщение контрольных карт Шухарта для независимой последовательности многомерных случайных векторов предложено Г. Хотеллингом [22]. Различные варианты обобщений для многомерного контроля алгоритмов CUSUM и EWMA предложены в [26–29].

Общий подход к контролю качества продукции достаточно прост. В процессе производства проводятся выборки изделий заданного объема. После этого строятся диаграммы (карты контроля качества) изменения выборочных значений параметра изделия (показателя качества). Ход процесса считается удовлетворительным, если найденные значения лежат в некоторых заранее заданных пределах. То есть в каждый момент времени проверяется гипотеза H_0 (процесс статистически управляем) при альтернативной гипотезе H_1 о статистической неуправляемости.

Если оказывается, что выборочные значения находятся на одной из контрольных границ или за ее пределами, то нулевая гипотеза отклоняется и считается, что процесс вышел из-под контроля (произошла разладка), предпринимаются необходимые действия для того, чтобы найти причину его разладки [17, 18]. Обычно контролируется как изменение среднего значения показателя качества, характеризующего уровень настройки процесса, так и изменение технологического рассеивания.

2. Карта Хотеллинга

Качество изделия обычно характеризуется несколькими показателями, которые могут быть коррелированы между собой. В этом случае независимый контроль по отдельным показателям может привести к значительным погрешностям вследствие различия доверительных областей и невозможности определения совместного уровня значимости [21]. В такой ситуации применяют многомерные контрольные карты. Впервые контроль качества с использованием нескольких характеристик был предложен Хотеллингом в [22]. Рассмотрим типы многомерных контрольных карт. Более подробный обзор многомерных карт можно найти, например, в [20, 28].

Предположим, что в технологическом процессе контролируются p показателей качества $\mathbf{X} = (X_1, X_2, \dots, X_p)$, имеющих совместное нормальное распределение. Плотность распределения равна

$$f(\mathbf{X}) = (2\pi)^{-\frac{p}{2}} |\Sigma|^{-\frac{1}{2}} e^{-\frac{(\mathbf{X}-\mu)^T \Sigma^{-1} (\mathbf{X}-\mu)}{2}}, \quad (1)$$

где μ — вектор средних значений; Σ — ковариационная матрица, элементы которой $\sigma_{ii} = \sigma_i^2$ — дисперсии случайных величин X_i , а $\sigma_{ij} = \sigma_{ji} = \rho_{ij} \sigma_i \sigma_j$ — ковариации (ρ_{ij} — коэффициент корреляции между величинами X_i и X_j).

Для проверки гипотезы $H_0: \mu = \mu_0$ в многомерном варианте применяется обобщенная статистика Хотеллинга [23, 24]:

$$T^2 = n(\bar{\mathbf{X}} - \mu_0)^T S^{-1} (\bar{\mathbf{X}} - \mu_0), \quad (2)$$

где S — выборочная оценка ковариационной матрицы Σ . Применение контрольной карты Хотеллинга предполагает расчет для каждой t -й мгновенной выборки ($t = 1, \dots, m$) статистики T_t^2 по формуле (2). При нормальном ходе процесса должно выполняться условие $T_t^2 < T_{кр}^2$, где $T_{кр}^2$ — граница критической области.

Многомерная контрольная карта Хотеллинга, по существу, есть та же карта Шухарта, в которой в качестве контролируемой величины используется обобщенная статистика Хотеллинга. Обратим внимание на то, что эта карта имеет только верхнюю контрольную границу.

Если ковариационная матрица Σ известна, статистика Хотеллинга имеет распределение χ^2 . В этом случае положение контрольной границы на заданном уровне значимости α определяется по таблицам квантилей этого распределения $T_{кр}^2 = \chi_{1-\alpha}^2(p)$ [25]. При неизвестной ковариационной матрице статистика

$$F = \frac{n-p}{p(n-1)} T^2 \quad (3)$$

имеет нецентральное F -распределение Фишера с p и $(n-p)$ степенями свободы и параметром нецентральности

$$\lambda^2 = n(\mu - \mu_0)^T \Sigma^{-1} (\mu - \mu_0). \quad (4)$$

При этом статистика T^2 имеет распределение Хотеллинга (центральное), плотность которого равна

$$f(x) = \frac{\Gamma\left(\frac{n+1}{2}\right) x^{\frac{p}{2}-1} \left(1 + \frac{x}{n}\right)^{-\frac{n+1}{2}}}{\Gamma\left(\frac{n-p+1}{2}\right) \Gamma\left(\frac{p}{2}\right) n^{\frac{p}{2}}}, x > 0. \quad (5)$$

В одномерном случае $p = 1$ и распределение Хотеллинга совпадает с квадратом распределения Стьюдента.

Карта Хотеллинга применима только в условиях совместной нормальности распределения контролируемых показателей. В многомерном случае практически полезным подходом при нарушении нормальности является преобразование (нормализация) данных таким образом, чтобы этот преобразованный набор имел хотя бы приблизительно нормальное распределение [21].

3. Анализ эффективности методов контроля качества в многомерном случае

Вопросы одномерного контроля качества изучены достаточно глубоко, в то время как исследования по многомерному статистическому контролю носят разрозненный характер и не позволяют обеспечить надежный контроль при проведении технологического процесса с коррелированными показателями качества. Чаще и в отечественной [19], и в зарубежной практике многомерный контроль подменяется независимым контролем нескольких показателей.

В случае некоррелированности показателей качества при независимом контроле с использованием соответствующего количества карт Шухарта в [21] было отмечено, что с увеличением числа показателей средняя длина серий несколько уменьшается. Если контролируемые показатели оказываются коррелированными, использование независимого контроля отдельных показателей может привести к серьезным ошибкам [30], связанным с двумя обстоятельствами.

Во-первых, различны доверительные области (acceptance region): при независимом контроле это прямоугольный параллелепипед, стороны которого определяются границами регулирования карт Шухарта. С учетом корреляционных связей в действительности доверительная область при многомерном нормальном распределении показателей представляет эллипсоид, главные оси которого повернуты относительно осей параллелепипеда. Точки, оказывающиеся внутри параллелепипеда, но вне эллипсоида,

свидетельствуют о нормальном ходе процесса, хотя на самом деле процесс статистически неуправляем.

Во-вторых, вычисление совместного уровня значимости невозможно при контроле по отдельным показателям, коррелированным между собой.

Один из возможных вариантов решения проблемы — переход от зависимых показателей к статистически независимым с использованием метода главных компонент [31]. Такое преобразование позволит, во-первых, обеспечить переход к некоррелированным показателям качества. Во-вторых, при определенных условиях после перехода к главным компонентам размерность задачи может снизиться, если у части последних главных компонент будут достаточно малые дисперсии.

Контрольная граница карты Хотеллинга в зависимости от конкретных условий определяется с использованием квантилей распределений χ^2 или Фишера с учетом количества контролируемых показателей, объема мгновенной выборки и количества выборок. При сравнении карт Шухарта на главных компонентах и карт Хотеллинга в [21] отмечается, что при некоррелированных показателях качества средняя длина серий в картах Шухарта меньше. Однако при коррелированных показателях, несмотря на уменьшение средней длины серий для обеих карт, степень уменьшения ARL с увеличением степени коррелированности для карт Хотеллинга гораздо выше. Среди недостатков карт Шухарта на главных компонентах отмечают также и невозможность во многих случаях правильной интерпретации данных, так как преобразование приводит к частичной потере первоначальной информации [32, 33].

Построение других многомерных карт — MCUSUM и MEWMA — является гораздо более трудоемким процессом. К тому же было установлено, что карты MCUSUM и MEWMA в случае больших сдвигов являются менее эффективными, чем карты Хотеллинга [34]. Накопление информации в таких картах происходит вследствие отслеживания истории процесса. Однако чаще всего производится анализ на наличие ЦВЗ абсолютно независимых друг от друга изображений, и, стало быть, для решения поставленных в работе задач необходимость в ведении истории процесса отпадает.

Таким образом, из анализа применения различных контрольных карт можно сделать вывод, что основным инструментом многомерного статистического контроля технологического процесса с коррелированными показателями качества является карта Хотеллинга.

Кроме выхода за границу регулирования к признакам нестабильности процесса также относятся и другие неслучайные структуры; в частности, в [19] упомянут *тренд процесса* — монотонное возрастание или убывание нескольких точек подряд. В одномерных картах Шухарта своевременно обнаружить неслучайные структуры позволяют предупреждающие линии (warning lines), впервые предложенные в [11]. Различные критерии серий с использованием таких линий были исследованы, в частности, в [35]. В картах Хотеллинга критерии нестабильности многомерного процесса вследствие появления тренда *отсутствуют* [21]. В связи с этим при наличии тренда карта Хотеллинга работает менее эффективно (возрастает значение ложных срабатываний, снижается вероятность объявления тревоги).

Поэтому для задач обнаружения ЦВЗ целесообразно разработать модификацию карты Хотеллинга, которая позволила бы повысить вероятность своевременного обнаружения разладки при наличии тренда.

4. Разработка модификации карты Хотеллинга, нивелирующей влияние тренда

На основе анализа применения многомерных контрольных карт был сделан вывод о том, что наиболее эффективным инструментом многомерного контроля качества является карта Хотеллинга. Однако возникает необходимость нивелировать влияние возможного тренда в одном или нескольких показателях (признаках), который может появиться, в частности, в результате незначительного изменения изображения (поворот на небольшой угол, небольшой сдвиг фотоаппарата при панорамной съемке, небольшое изменение яркости/цветности и т. п.). Также, если говорить об обнаружении ЦВЗ в изображениях, небольшое изменение признака, вызванное такой незначительной естественной модификацией изображения, может повлиять на решение анализатора (контрольной карты). В этом случае возрастает количество ложных тревог.

В соответствии с этим целесообразно выделить следующие направления исследования:

- 1) разработка модификации карты Хотеллинга, не чувствительной к тренду (в этом случае контрольная карта не считает тренд причиной нестабильности процесса);
- 2) разработка алгоритма применения модификации карты Хотеллинга к обнаружению ЦВЗ;
- 3) проведение экспериментов и сравнение результатов.

Статистика в карте Хотеллинга вычисляется по формуле (2), где ковариационная матрица обычно оценивается формулой

$$S = \frac{1}{(n-1)} \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T. \quad (6)$$

При тренде такая оценка ковариационной матрицы будет неточной, так как тренд, изменяя значения векторов, будет изменять и значение среднего. Это может привести к ухудшению контроля качества. Следовательно, целесообразно применить другую оценку ковариационной матрицы. Кроме того, в случае анализа изображения на предмет наличия ЦВЗ использование мгновенных выборок для контроля зачастую невозможно (есть одно изображение, которое необходимо проанализировать). В этом случае приходится использовать результаты индивидуальных наблюдений, и для получения несмещенных оценок компонент ковариационной матрицы как при одномерном, так и при многомерном контроле применяются скользящие размахи [36, 37]. Также применение размахов в (9) позволяет дать более точную оценку ковариационной матрицы при наличии тренда в силу того, что в изображениях, содержащих некоторый локальный тренд и шумовую составляющую, ковариация между приращениями различных компонент остается постоянной (исключается влияние тренда) в предположении постоянства коэффициента корреляции между компонентами шумовой составляющей.

Для аппроксимации распределения статистики модифицированной карты одним из известных распределений потребуется нормирующий множитель.

Пусть $\mathbf{x}_1, \dots, \mathbf{x}_n$ — выборка независимых в совокупности одинаково распределенных векторов, где $\mathbf{x}_i \in \mathbb{R}^d$; $\mathbf{x}_i \sim \mathcal{N}(\mu, \Sigma)$; $\mu \in \mathbb{R}^d$ — вектор математических ожиданий; $\Sigma \in \mathbb{R}^{d \times d}$ — ковариационная матрица.

Обозначим через $X \in \mathbb{R}^{n \times d}$ матрицу, составленную из векторов $\mathbf{x}_i^T$, $i = 1, \dots, n$, то есть $X^T = (\mathbf{x}_1, \dots, \mathbf{x}_n)$. Положим $\mathbf{z}_i = \mathbf{x}_{i+1} - \mathbf{x}_i$, где $i = 1, \dots, n-1$.

Очевидно, что $Z = C \cdot X$, где

$$Z = \begin{pmatrix} \mathbf{z}_1^T \\ \vdots \\ \mathbf{z}_{n-1}^T \end{pmatrix} = \begin{pmatrix} \mathbf{x}_2^T - \mathbf{x}_1^T \\ \vdots \\ \mathbf{x}_n^T - \mathbf{x}_{n-1}^T \end{pmatrix}, \quad (7)$$

$$C = \begin{pmatrix} -1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -1 & 1 \end{pmatrix} \in \mathbb{R}^{(n-1) \times n}. \quad (8)$$

В дальнейшем, не ограничивая общности, будем предполагать, что $\mu = \mathbf{0}$. Введем следующие обозначения:

$$S_n = \frac{1}{2(n-1)} \sum_{i=1}^{n-1} \mathbf{z}_i \cdot \mathbf{z}_i^T; \quad (9)$$

$$\bar{x}_n = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i; \quad (10)$$

$$f(n) = \frac{2(n-1)^2}{3n-4}; \quad (11)$$

$$T_n = \frac{f(n) - d + 1}{f(n) \cdot d} \cdot \frac{n}{n+1} (\mathbf{x} - \bar{x}_n)^T S_n^{-1} (\mathbf{x} - \bar{x}_n), \quad (12)$$

где $x \in \mathbb{R}^d$; $\mathbf{x} \sim \mathcal{N}(\mu, \Sigma)$ — вектор, отличный от векторов выборки $\mathbf{x}_1, \dots, \mathbf{x}_n$.

Замечание 1. Если объем выборки n невелик по сравнению с d , то матрица S_n может оказаться вырожденной. В таком случае для подсчета статистики T_n можно поступить следующим образом:

- 1) снизить размерность векторов выборки $\mathbf{x}_1, \dots, \mathbf{x}_n$ при помощи метода главных компонент, оставив компоненты с большей дисперсией [38]. Получим новую выборку $\tilde{\mathbf{x}}_1, \dots, \tilde{\mathbf{x}}_n$;
- 2) подсчитать статистику T_n по выборке $\tilde{\mathbf{x}}_1, \dots, \tilde{\mathbf{x}}_n$.

Лемма 1. При сделанных выше предположениях

$$S_n = \frac{1}{n-1} \cdot X^T A X, \quad (13)$$

где $A = \frac{C^T \cdot C}{2}$.

Доказательство. Имеем

$$S_n = \frac{1}{2(n-1)} \sum_{i=1}^{n-1} \mathbf{z}_i \cdot \mathbf{z}_i^T = \frac{1}{2(n-1)} Z^T Z.$$

Пусть $A = (C^T \cdot C)/2$. Тогда из равенства $Z = C \cdot X$ следует, что $S_n = \frac{1}{2(n-1)} (CX)^T CX = \frac{1}{2(n-1)} X^T (C^T C) X = \frac{1}{n-1} X^T A X$. ■

Лемма 2. S_n — несмещенная и состоятельная оценка матрицы Σ .

Доказательство.

1) Верна цепочка равенств

$$\begin{aligned} ES_n &= \frac{1}{2(n-1)} \sum_{i=1}^{n-1} E(z)_i \mathbf{z}_i^T = \frac{1}{2(n-1)} \sum_{i=1}^{n-1} E(\mathbf{x}_{i+1} - \mathbf{x}_i)(\mathbf{x}_{i+1} - \mathbf{x}_i)^T = \\ &= \frac{1}{2(n-1)} \sum_{i=1}^{n-1} E\mathbf{x}_{i+1}\mathbf{x}_{i+1}^T - \mathbf{x}_i\mathbf{x}_{i+1}^T + \mathbf{x}_i\mathbf{x}_i^T - \mathbf{x}_{i+1}\mathbf{x}_i^T = \\ &= \frac{1}{2(n-1)} \sum_{i=1}^{n-1} (E\mathbf{x}_{i+1}\mathbf{x}_{i+1}^T - E(\mathbf{x}_i\mathbf{x}_{i+1}^T) + E(\mathbf{x}_i\mathbf{x}_i^T) - E(\mathbf{x}_{i+1}\mathbf{x}_i^T)) = \frac{1}{2(n-1)} \sum_{i=1}^{n-1} (\Sigma + \Sigma) = \Sigma, \end{aligned}$$

где $E(\mathbf{x}_i\mathbf{x}_{i+1}^T) = E(\mathbf{x}_{i+1}\mathbf{x}_i^T) = 0$, так как $\mathbf{x}_i$ и $\mathbf{x}_{i+1}$ — независимые случайные векторы.

2) Рассмотрим произвольный элемент s_{ij}^n , $i = 1, \dots, d$, $j = 1, \dots, d$, матрицы S_n . Здесь z_p^i и x_p^i — i -е координаты векторов $\mathbf{z}_p$ и $\mathbf{x}_p$ соответственно:

$$\begin{aligned} s_{ij}^n &= \frac{1}{2(n-1)} \sum_{p=1}^{n-1} z_p^i z_p^j ; \\ D(s_{ij}^n) &= \frac{1}{4(n-1)^2} D\left(\sum_{p=1}^{n-1} z_p^i z_p^j\right) = \frac{1}{4(n-1)^2} \left(\sum_{p=1}^{n-1} D z_p^i z_p^j + 2 \cdot \sum_{1 \leq p < q \leq n-1} \text{cov}(z_p^i z_p^j, z_q^i z_q^j)\right) = \\ &= \frac{1}{4(n-1)^2} \sum_{p=1}^{n-1} D z_p^i z_p^j + \frac{1}{4(n-1)^2} \cdot 2 \cdot \sum_{1 \leq p < q \leq n-1} \text{cov}(z_p^i z_p^j, z_q^i z_q^j) = I + II ; \\ I &= \frac{1}{4(n-1)^2} \left(\sum_{p=1}^{n-1} D(x_{p+1}^i - x_p^i)(x_{p+1}^j - x_p^j)\right) = \\ &= \frac{1}{4(n-1)^2} \left(\sum_{p=1}^{n-1} D(x_{p+1}^i x_{p+1}^j - x_{p+1}^i x_p^j - x_p^i x_{p+1}^j + x_p^i x_p^j)\right) = O\left(\frac{1}{n}\right), \end{aligned}$$

так как $D(x_p^i x_p^j) < \infty$ и $D(x_{p+1}^i x_p^j) < \infty$.

В связи с тем, что $z_p^i z_p^j = (x_{p+1}^i - x_p^i)(x_{p+1}^j - x_p^j)$, имеем

$$\text{cov}(z_p^i z_p^j, z_q^i z_q^j) = \begin{cases} 0 & \text{при } p < q - 1; \\ \neq 0 & \text{при } p = q - 1. \end{cases}$$

Тогда

$$II = \frac{1}{4(n-1)^2} \cdot 2 \sum_{p=1}^{n-2} \text{cov}(z_p^i z_p^j, z_{p+1}^i z_{p+1}^j) = O\left(\frac{1}{n}\right),$$

так как $\text{cov}(z_p^i z_p^j, z_{p+1}^i z_{p+1}^j) < \infty$.

Поскольку $D(s_{ij}^n) \rightarrow 0$ при $n \rightarrow \infty$ и $ES_n = \Sigma$, то S_n — состоятельная оценка ковариационной матрицы. ■

В [39] предлагается способ оценки компонент ковариационной матрицы через скользящие размахи.

В многомерном случае основная идея аппроксимации Welch — Satterthwaite [40, 41] может быть использована для того, чтобы показать, что распределение матрицы S_n может быть аппроксимировано с приемлемой точностью распределением Уишарта $W_d(f(n), \Sigma)$ [42, 43]. Возможность того, что распределение S_n может быть аппроксимировано распределением Уишарта $W_d(f(n), \Sigma)$ в многомерном случае, вытекает из следующего результата.

Теорема 1 [42]. Пусть $(\mathbf{y}_1, \dots, \mathbf{y}_n)$ — выборка независимых в совокупности одинаково распределенных векторов $\mathbf{y}_i \in \mathbb{R}^d$, $\mathbf{y}_i \sim \mathcal{N}(\mathbf{0}, \Sigma_y)$. Обозначим $Y = \begin{pmatrix} \mathbf{y}_1^\top \\ \vdots \\ \mathbf{y}_n^\top \end{pmatrix} \in \mathbb{R}^{n \times d}$ и положим $\xi = Y \cdot \eta$, где $\eta \in \mathbb{R}^d$ — произвольный вектор. Пусть также $B \in \mathbb{R}^{n \times n}$ — произвольная симметричная матрица ранга r . В таком случае

$$Y^\top B Y \sim W_d(r, \Sigma_y) \quad (14)$$

тогда и только тогда, когда

$$\xi^\top B \xi \sim \sigma_\eta^2 \chi_r^2 \quad (15)$$

для любого $\eta \in \mathbb{R}^d$, где $\sigma_\eta^2 = \eta^\top \Sigma_y \eta$, а χ_r^2 — распределение хи-квадрат с r степенями свободы.

Теорема 2. При сделанных выше предположениях квантили распределения статистики T_n и квантили случайной величины с распределением Фишера с $(d, f(n) - d + 1)$ степенями свободы отличаются не более чем на $O\left(\frac{1}{f(n)}\right)$, где $f(n) = O(n)$ при $n \gg 1$.

Доказательство. Обозначим $\xi_i = \mathbf{x}_i^\top \eta$, где $\eta \in \mathbb{R}^d$ — произвольный постоянный вектор, а $\mathbf{x}_i$ — вектор из исходной выборки.

Очевидно, что $\xi_i \sim \mathcal{N}(0, \sigma_\eta^2)$, где $\sigma_\eta^2 = \eta^\top \Sigma \eta$. Пусть

$$\xi = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}; \quad \xi = X \cdot \eta.$$

Подберем такие константы $\alpha > 0$ и $f > 0$, что первые два момента случайных величин $\frac{\xi^\top A \xi}{\alpha}$ и $\sigma_\eta^2 \chi_f^2$ совпадают [40]. Первые моменты случайной величины $\frac{\xi^\top A \xi}{\alpha}$ равны

$$\begin{aligned} \mathbb{E}\left(\frac{\xi^\top A \xi}{\alpha}\right) &= \mathbb{E}\left(\frac{\sum_{i=1}^{n-1} (\xi_{i+1} - \xi_i)^2}{2\alpha}\right) = \frac{n-1}{\alpha} \sigma_\eta^2, \\ D\left(\frac{\xi^\top A \xi}{\alpha}\right) &= \frac{\sum_{i=1}^{n-1} D((\xi_{i+1} - \xi_i)^2)}{4\alpha^2} + \frac{\sum_{1 \leq i < j \leq n-1} 2\text{cov}((\xi_{i+1} - \xi_i)^2, (\xi_{j+1} - \xi_j)^2)}{4\alpha^2} = \frac{\sigma_\eta^4}{\alpha^2} (3n - 4). \end{aligned}$$

В силу того, что χ_f^2 — случайная величина, имеющая χ_f^2 -распределение с f степенями свободы, после приравнивания получим

$$\begin{cases} \frac{n-1}{\alpha} \sigma_\eta^2 = f \cdot \sigma_\eta^2; \\ \frac{\sigma_\eta^4}{\alpha^2} (3n - 4) = 2f \cdot \sigma_\eta^4, \end{cases} \quad (16)$$

откуда

$$\begin{cases} \alpha = \alpha(n) = \frac{3n-4}{2(n-1)}; \\ f = f(n) = \frac{2(n-1)^2}{3n-4} \approx \frac{2}{3}n \text{ при } n \gg 1. \end{cases} \quad (17)$$

При подобранных значениях $\alpha = \alpha(n)$ и $f = f(n)$ из [41] следует, что квантили распределения случайных величин $\frac{\xi^T A \xi}{\alpha}$ и $\sigma_\eta^2 \chi_f^2$ будут отличаться не более чем на $O\left(\frac{1}{f(n)}\right)$. В силу этого и из теоремы 1 следует, что квантили распределения случайных величин $\frac{(n-1)S_n}{\alpha(n)} = f(n)S_n$ и $W_d(f(n), \Sigma)$ будут отличаться не более чем на $O\left(\frac{1}{f(n)}\right)$.

Далее, так как по предположению $\mathbf{x}_i \sim \mathcal{N}(\mu, \Sigma)$, то $\bar{x}_n = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i \sim \mathcal{N}\left(\mu, \frac{\Sigma}{n}\right)$, при этом $\bar{x}_n$ и S_n — независимые величины [42, 43]. Так как $\mathbf{x} \sim \mathcal{N}(\mu, \Sigma)$, то $\mathbf{x} - \bar{x}_n \sim \mathcal{N}\left(\mathbf{0}, \Sigma\left(1 + \frac{1}{n}\right)\right)$ и, аналогично, $\mathbf{x} - \bar{x}_n$ не зависит от S_n .

Из свойств распределения Уишарта [43], из того, что квантили распределения случайных величин $\frac{(n-1)S_n}{\alpha(n)} = f(n)S_n$ и $W_d(f(n), \Sigma)$ отличаются не более чем на $O\left(\frac{1}{f(n)}\right)$, и из [41, 42] следует, что квантили распределения статистики T_n из (12) будут отличаться от квантилей распределения Фишера с $(d, f(n) - d + 1)$ степенями свободы не более чем на $O\left(\frac{1}{f(n)}\right) = O\left(\frac{1}{n}\right)$, где $f(n) = O(n)$. ■

Таким образом, было показано, что, во-первых, оценивать ковариационную матрицу при помощи скользящих размахов можно достаточно точно (данная оценка является несмещенной и состоятельной) и, во-вторых, контрольная граница в модифицированной карте Хотеллинга может быть вычислена из распределения Фишера.

5. Предлагаемый метод обнаружения ЦВЗ

При решении задачи обнаружения ЦВЗ указанные выше методы контроля качества можно использовать для обнаружения разладки (появления ЦВЗ) после наблюдения нескольких пустых контейнеров (чем больше пустых контейнеров, тем выше вероятность правильного обнаружения). Применение таких подходов в классическом виде в стегоанализе не совсем корректно, так как в теории контроля качества подразумевается, что перед разладкой было достаточно много наблюдений при нормальном ходе производства, а после разладки производятся только бракованные изделия. В задаче стегоанализа такая ситуация происходит очень редко. Маркированные и немаркированные контейнеры обычно появляются в перемешанном порядке (если предположить, что, например, абонент маркирует только часть контейнеров, либо заранее неизвестно, какой из нескольких абонентов использует ЦВЗ), при этом контейнер с ЦВЗ может передаваться в самом начале работы канала связи. Поэтому предлагается перед первым наблюдением (изображением, полученным по каналу связи) провести обучение на выборке созданных самостоятельно заведомо пустых контейнеров.

Если детектор признает очередной наблюдаемый контейнер пустым (разладки не произошло), то контейнер добавляется в обучающую выборку. Если же контейнер признан содержащим ЦВЗ, тогда объявляется тревога и обучающая выборка остается прежней до следующего наблюдения.

Каждому изображению поставим в соответствие вектор из d признаков $\mathbf{x}_i \in \mathbb{R}^d$, каждая компонента которого есть некоторый признак контейнера. Кроме того, можно

выбирать из набора признаков наиболее информативные по отношению к конкретному методу внедрения ЦВЗ (изменения которых существеннее остальных при переходе от пустого контейнера к маркированному).

Постановка задачи. Будем считать, что все векторы признаков изображений, в которых нет ЦВЗ, порождены вероятностным распределением, принадлежащим одному классу. Соответственно изображения с ЦВЗ имеют признаки, порожденные другим распределением. При внедрении ЦВЗ в изображение распределение признаков изменяется существенно (гораздо существеннее, чем при любой другой модификации изображения, не влияющей на восприятие системой человеческого зрения). Модификация изображения, вносящая столь же значительные изменения в распределение признаков, например размытие изображения, как правило, визуально заметна. Задача состоит в том, чтобы по наблюдениям $\mathbf{x}_i$, $i = 1, 2, \dots$, выявить, какие из изображений в потоке содержат ЦВЗ.

Поставленную задачу можно решить посредством карт контроля качества. В этом случае технологический процесс представляет собой передачу изображений по каналу связи. Процесс считается стабильным (изображение *не* содержит ЦВЗ), если признаки изображения принадлежат распределению с целевыми параметрами, полученными для обучающей выборки. И наоборот, процесс становится нестабильным (изображение содержит ЦВЗ), если распределение признаков отличается от целевого (верна альтернативная гипотеза).

В связи с тем, что для обнаружения ЦВЗ необходим анализ нескольких признаков, следует использовать многомерные контрольные карты, наиболее эффективной из которых, как было показано выше, является карта Хотеллинга.

Эффективность работы контрольной карты возрастает, если на этапе обучения использовались изображения, максимально похожие на тестовое. Также эффективность работы карты будет тем выше, чем больше объем обучающей выборки. Таким образом, при достаточно больших объемах обучающей выборки изображения, включенные в данную выборку, будут минимально отличаться друг от друга, и, как следствие, будут существовать такие подвыборки обучающих изображений, изменение ряда признаков (например, яркость изображения, цветовые гистограммы, блочность изображения, дисперсия цветового слоя, дисперсия ненулевых коэффициентов ДКП) которых будет иметь монотонный характер.

Присутствие тренда в некоторых признаках изображений (в том числе и из обучающей выборки), вызванного не внедрением ЦВЗ (при внедрении ЦВЗ тренд, как таковой, не появляется, а происходит скачок за пределы контрольной границы), должно быть учтено при работе контрольной карты. Следовательно, влияние тренда на объявление тревоги должно быть минимизировано.

Для этого предлагается использовать карту контроля качества с применением модифицированной статистики Хотеллинга, а именно критерий, основанный на статистике T_n (см. (12)). Пусть n — количество изображений, которые уже были предъявлены до текущего момента времени (допустим, что в этих изображениях ЦВЗ еще не было), а $\mathbf{x}$ — вектор признаков нового изображения, которое было предъявлено для анализа в текущий момент времени. Необходимо решить, содержит ли предъявленное изображение ЦВЗ.

Зададим уровень значимости $\alpha \in (0, 1)$. Вычислим значение контрольной границы $g(n)$:

$$\int_{g(n)}^{\infty} p_{d, f(n)-d+1}(x) dx = \alpha, \quad (18)$$

где $p_{d,f(n)-d+1}(x)$ — плотность распределения Фишера с d и $f(n) - d + 1$ степенями свободы.

К обучающей выборке предъявляются следующие требования:

- 1) объем выборки n должен быть таким, что $f(n) - d + 1 > 0$;
- 2) желательно, чтобы изображения обучающей выборки максимально походили на тестируемое (принадлежали к одному типу изображений (пейзаж, средний план, поле), имели примерно одинаковую гистограмму цветов, одинаковое разрешение, одинаковый коэффициент качества, были созданы однотипным способом (сфотографированы одним фотоаппаратом) и т. п.).

Критерий на основе предложенной статистики можно описать следующим образом:

- 1) накапливаем выборку данных $V_N = \{x_i : i = 1, \dots, N\}$ — векторы признаков пустых изображений. N должно быть таким, что $f(N) - d + 1 > 0$;
- 2) вычисляем $\bar{x}_N$ и S_N ;
- 3) для тестируемого изображения задаем уровень значимости $\alpha \in (0, 1)$ и выбираем $g(N)$, как описано выше (18). Проверяем условие: если $T_N < g(N)$, то принимаем гипотезу, что изображение не содержит ЦВЗ, в противном случае принимаем гипотезу, что содержит ЦВЗ.

Кроме предложенного способа для вычисления порога $g(n)$ можно также применять метод Монте-Карло, однако данный метод требует большого объема обучающей выборки и является достаточно трудоемким по сравнению с предложенным.

Вариант 1. Допустим, требуется оценить величину $g(n)$ для конкретного значения n . Рассматриваем выборку пустых изображений в количестве $5n$. Положим $N \gg 1$. Выделяем N случайных подвыборок объема $n + 1$ из общей выборки.

Для каждой подвыборки, рассматривая в качестве базы n изображений, для оставшегося одного изображения вычисляем значения статистики T_n . Таким образом, получаем N значений статистики T_n .

Строим по полученным значениям гистограмму. Выбираем такое значение $g(n)$, чтобы α значений статистики T_n были больше $g(n)$.

Вариант 2 (bootstrap). Рассматриваем выборку объема $n + 1$. Используя первые n изображений (с номерами $1, \dots, n$), считаем статистику T_n для последнего изображения, далее циклически сдвигаем выборку. То есть теперь используем изображения с индексами $n + 1, 1, 2, \dots, n - 1$ в качестве базы и считаем статистику T_n для изображения с индексом n .

Всего таких циклических сдвигов будет $n + 1$ штук, то есть получим $n + 1$ значений статистики T_n . Далее для вычисления порогов поступаем аналогично предыдущему варианту.

Обнаружение ЦВЗ в графических контейнерах проиллюстрировано при помощи модифицированных контрольных карт Хотеллинга (см. рис. 1–4). Результаты работы метода обнаружения ЦВЗ, внедренных методом Digimap, показаны на рис. 1. При помощи статистики T_n оценивалось изменение выборочного среднего (в качестве набора признаков использовался характеристический вектор из [1]). Обучение проводилось на 200 фотореалистичных изображениях, индексы которых откладываются по оси абсцисс. На карте отображены значения статистики T_n для пустых контейнеров (обозначены точками) и для их маркированных аналогов (обозначены крестиками). Контрольная граница вычисляется из распределения Фишера (по формуле (18)). Ошибки 1-го и 2-го рода достаточно низкие, несмотря на небольшой объем обучающей выборки. Это обуславливается хорошим подбором изображений в обучающей выборке (обучающая

выборка была одной и той же для всех изображений тестовой выборки; обе выборки состоят из изображений, представляющих собой пейзажи летнего леса) и вставкой большого количества копий ЦВЗ в контейнер (зависит от алгоритма внедрения ЦВЗ).

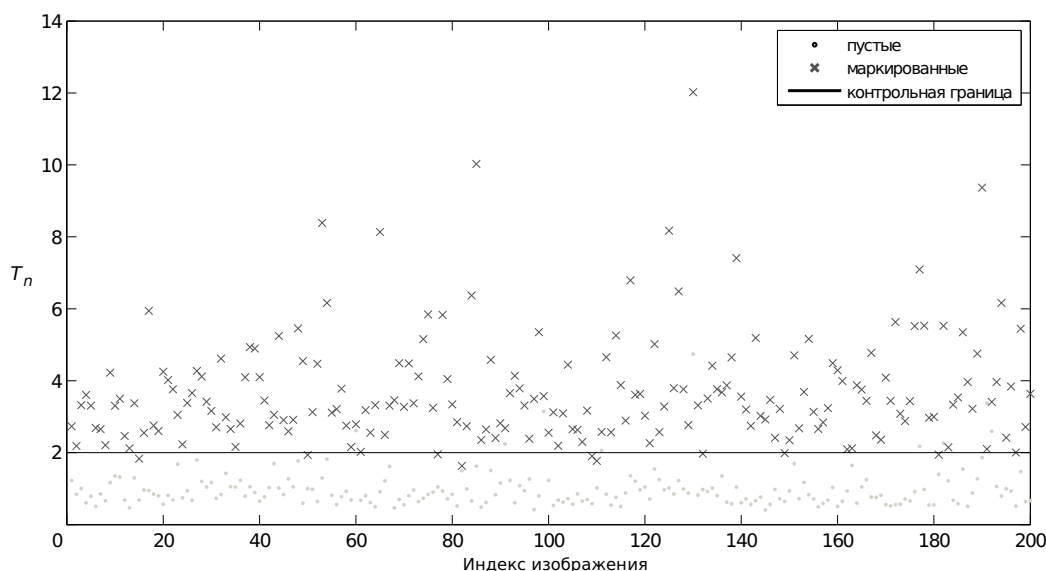


Рис. 1. Обнаружение ЦВЗ, внедренных Digimarc, при помощи модификации контрольной карты. Контрольная граница вычисляется из распределения Фишера

Качество классификации пустых контейнеров и контейнеров, содержащих ЦВЗ, внедренных различными методами, отображено на рис. 2. Тестовая выборка состоит из 300 изображений: 100 пустых контейнеров и 200 маркированных. При этом половина маркированных получена из пустых внедрением одним методом внедрения ЦВЗ, вторая половина получена также из пустых контейнеров другим методом внедрения ЦВЗ. Из рис. 2 видно, что пустые контейнеры (обозначены точками) классифицированы отдельно от их маркированных аналогов (метод внедрения ЦВЗ 1 — кружки, метод внедрения ЦВЗ 2 — крестики). Контрольная граница вычисляется из распределения Фишера (как и прежде, на основе только обучающей выборки пустых контейнеров). В зависимости от параметров и самой сути метода внедрения ЦВЗ соответствующие маркированные контейнеры находятся ближе или дальше от контрольной границы.

Таким образом, можно сделать вывод о том, что данная контрольная карта не привязана к конкретному методу внедрения ЦВЗ и позволяет обнаруживать ЦВЗ, внедренные различными методами.

Схожесть изображений обучающей выборки и тестового изображения имеет важное значение. На рис. 3 и 4 показаны результаты классификации для одной и той же тестовой выборки. При этом на рис. 3 показана классификация с обучением на однотипных с тестовыми контейнерах. На рис. 4 обучение производилось на смешанной выборке (состоящей из контейнеров как похожих на тестовые, так и отличающихся от них). В качестве контейнеров рассматривались изображения двух типов: рассвет и облака. Тестовая выборка содержала 10000 изображений. Обучающие выборки составляли по 3500 изображений каждого типа.

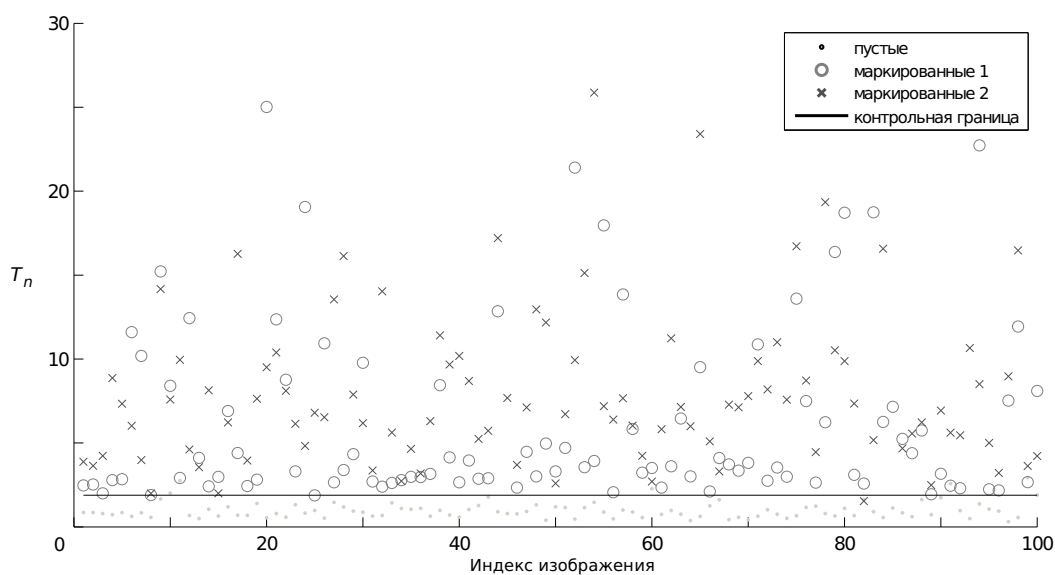


Рис. 2. Результат работы модифицированной контрольной карты при обнаружении ЦВЗ, внедренных различными методами

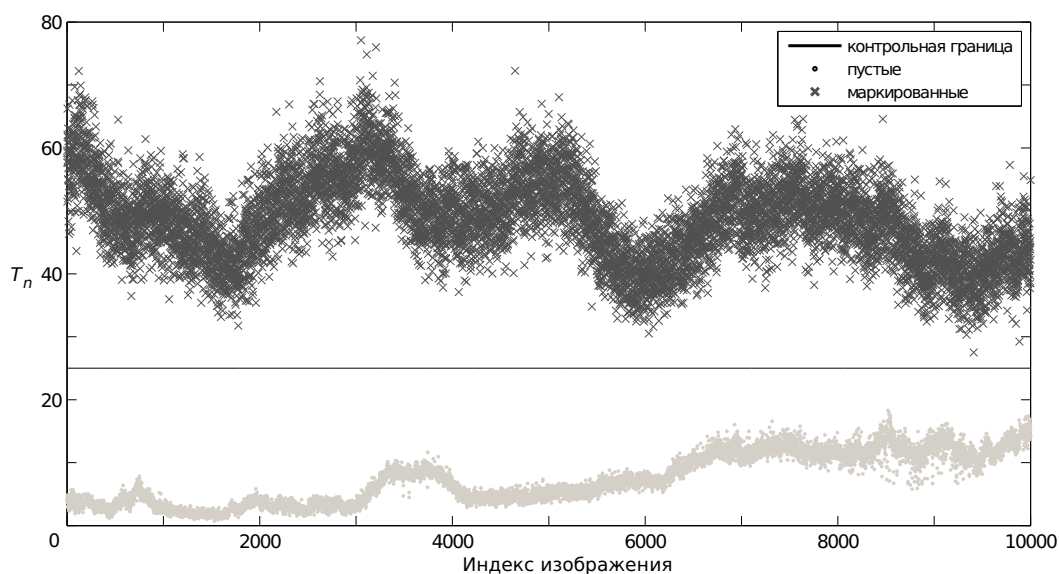


Рис. 3. Обнаружение ЦВЗ, внедренных Digimagc, при помощи модифицированной контрольной карты. Контрольная граница получена из распределения Фишера. Тип изображений обучающей и тестовой выборки совпадает

При обучении на достаточно большой (несколько тысяч) выборке контейнеров, не похожих на тестовые, наблюдается неплохое разделение на классы (ложная тревога и пропуск цели около 0,1), однако вычисленная контрольная граница не соответствует истинной.

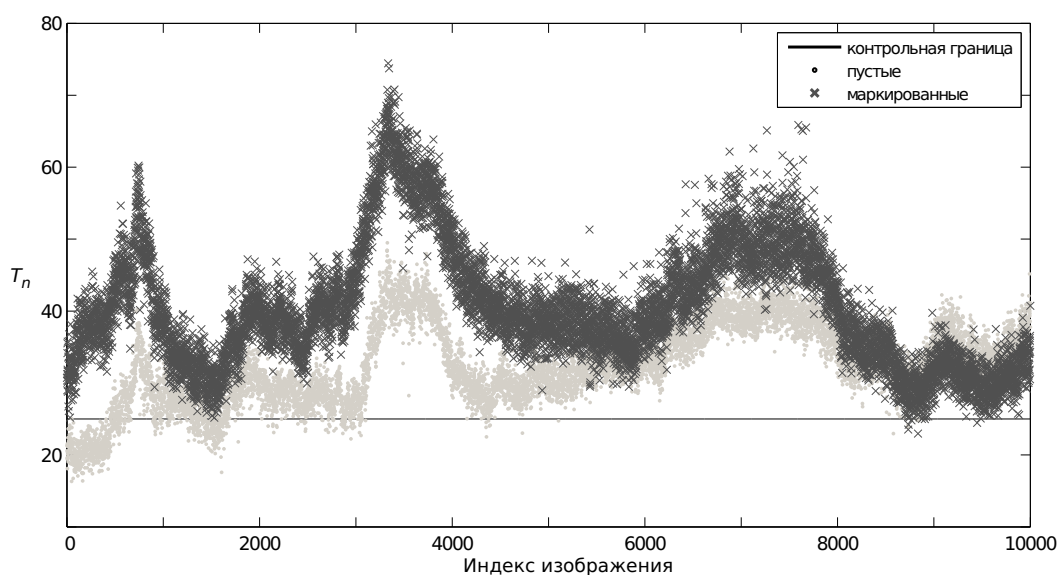


Рис. 4. Обнаружение ЦВЗ, внедренных Digimarc, при помощи модифицированной контрольной карты. Контрольная граница получена из распределения Фишера. Обучающая выборка состоит из изображений двух типов

ЛИТЕРАТУРА

1. *Fridrich J.* Feature-based steganalysis for JPEG images and its implications for future design of steganographic schemes // Proc. 6th Information Hiding International Workshop. LNCS. 2005. V. 3200. P. 67–81.
2. *Lyu S., Farid H.* Steganalysis using higher-order image statistics // IEEE Trans. Inform. Forens. Secur. 2006. V. 1. No. 1. P. 111–119.
3. *Holotyak T., Fridrich J., Voloshynovskiy S.* Blind Statistical Steganalysis of Additive Steganography Using Wavelet Higher Order Statistics // 9th IFIP TC-6 TC-11 Conf. on Communications and Multimedia Security. LNCS. 2005. V. 3677. P. 273–274.
4. *Pevny T., Fridrich J.* Detection of double-compression in JPEG images for applications in steganography // International Workshop on Digital Watermarking. Springer Verlag, 2007.
5. *Pevny T., Fridrich J.* Merging Markov and DCT Features for Multi-Class JPEG Steganalysis // Proc. SPIE Electronic Imaging. V. 6505. Photonics West, 2007. P. 03–04.
6. *Cancelli G., Doerr G., Cox I. J., Barni M.* Detection of LSB steganography based on the amplitude of histogram local extrema // Proc. International Conf. on Image Processing. 2008. P. 1288–1291.
7. www.outguess.org — Программный продукт StegDetect 0.6.
8. *Shewhart W. A.* The application of statistics as an aid in maintaining quality of manufactured product // J. Amer. Statist. Assoc. 1925. P. 546–548.
9. *Shewhart W. A.* Economic control of manufactured product. New York: Von Nostrand Reinhold, 1931.
10. *Page E. S.* Continuous inspection schemes // Biometrika. 1954. V. 41. P. 100–114.
11. *Page E. S.* Control charts with warning lines // Ibid. 1955. V. 42. P. 243–254.
12. *Roberts S. W.* Control charts based on geometric moving average // Technometrics. 1959. V. 1. P. 239–250.
13. *Roberts S. W.* A comparison of some control charts procedures // Ibid. 1966. V. 8. P. 411–430.

14. *Ширяев А. Н.* Обнаружение спонтанно возникающих эффектов // Докл. АН СССР. 1961. Т. 138. № 4. С. 799–801.
15. *Колмогоров А. Н., Прохоров Ю. В., Ширяев А. Н.* Вероятностно-статистические методы обнаружения спонтанно возникающих эффектов // Труды МИАН. 1988. Т. 182. С. 4–23.
16. *Вальд А.* Последовательный анализ. М.: Физматгиз, 1960. 328 с.
17. *Juran J. M.* Juran's quality handbook. McGraw-Hill, 5th ed., 1998. P. 109–117.
18. *Окрепиллов В. В.* Управление качеством. СПб.: Наука, 2000. 912 с.
19. ГОСТ Р 50779.42–99 «Статистические методы. Контрольные карты Шухарта».
20. *Bersimis S., Psarakis S., Panaretos J.* Multivariate statistical process control charts: an overview // Qual. Reliab. Engng. Int. 2007. V. 23. P. 517–543.
21. *Клячкин В. Н.* Модели и методы многомерного статистического контроля технологического процесса: дис. ... докт. техн. наук. Ульяновск, 2003. 285 с.
22. *Hotelling H.* Multivariate Quality Control Illustrated by Air Testing of Sample Bombsights // Techniques of Statistical Analysis. New York: MacGraw-Hill, 1947. P. 111–184.
23. *Hotelling H.* The generalization of Student's ratio // Ann. Math. Statist. 1931. V. 2. P. 360–378.
24. *Williams J. D., Woodall W. H., Birch J. B., Sullivan J. H.* Distribution of Hotelling's T^2 -statistic based on the successive differences estimator // J. Qual. Techn. 2006. V. 38. No. 3. P. 217–229.
25. *Крамер Г.* Математические методы статистики. М.: Мир, 1975. 648 с.
26. *Woodall W. H., Ncube M. M.* Multivariate CUSUM quality control procedures // Technometrics. 1985. V. 27. P. 285–292.
27. *Crosier R. B.* Multivariate generalizations of cumulative sum quality-control schemes // Ibid. 1988. V. 30. P. 291–303.
28. *Клячкин В. Н.* Многомерный статистический контроль технологического процесса. М.: Финансы и статистика, 2003. 192 с.
29. *Lowry C. A., Woodall W. H., Champ C. W., Rigdon S. E.* A multivariate EWMA control chart // Technometrics. 1992. V. 34. P. 46–53.
30. *Montgomery D. C.* Introduction to Statistical Quality Control. 3rd Edition. New York: John Wiley & Sons, 1996.
31. *Jackson J. E., Mudholkar G. S.* Control procedures for residuals associated with principal components analysis // Technometrics. 1979. V. 21. P. 341–349.
32. *Hawkins D. M.* Multivariate quality control based on regression-adjusted variables // Ibid. 1991. V. 33. P. 61–75.
33. *Lowry C. A., Montgomery D. C.* A review of multivariate control charts // IEEE Trans. 1995. V. 27. P. 800–810.
34. *Lucas J. M.* A modified V-mask control scheme // Technometrics. 1973. V. 15. P. 833–847.
35. *Champ C. W., Woodall W. H.* Exact results for Shewhart control charts with supplementary runs rules // Ibid. 1987. V. 29. P. 393–399.
36. *Мутмаг Х., Ринне Х.* Статистические методы обеспечения качества. М.: Машиностроение, 1995. 616 с.
37. *Ryan T. P.* Statistical methods for quality improvement. New York: John Wiley & Sons, 1989. 420 p.
38. Прикладная статистика: Классификация и снижение размерности: справочное издание / под ред. С. А. Айвазяна. М.: Финансы и статистика, 1989. 607 с.
39. *Клячкин В. Н.* Об оценке ковариационной матрицы при многомерном статистическом контроле технологического процесса // ОПиПМ. 2007. Т. 14. № 4. С. 174.

40. *Satterthwaite F. E.* An approximate distribution of estimates of variance components // Biometrics Bulletin. 1946. V. 2. P. 110–114.
41. *Welch B. L.* The generalization of “student’s” problem when several different population variances are involved // Biometrika. 1947. V. 34. P. 28–35.
42. *Seber G. A. F.* Multivariate observations. New York: John Wiley & Sons, 1984.
43. *Bilodeau M., Brenner D.* Theory of multivariate statistics. Springer, 1999.