

Некоторые алгоритмы обнаружения объектов на основе двухмасштабной модели изображения

П. А. Чочиа

Институт проблем передачи информации им. А. А. Харкевича РАН, Москва, Россия
e-mail: chochia@iitp.ru

Поступила в редколлегию 19.05.2014

Аннотация—Формулируется двухмасштабная модель изображения, приводится алгоритм декомпозиции изображения и выделения сглаженной компоненты, несущей основную информацию об объектах. Излагается алгоритм обнаружения объектов по их площади. Предлагается новый подход к обнаружению контуров на изображении. Рассматриваются аспекты обнаружения различных множеств объектов на изображениях. Решается задача разработки алгоритма обнаружения дефектов электронных микросхем.

КЛЮЧЕВЫЕ СЛОВА: обработка изображений, анализ, обнаружение объектов, модель изображения, дефектоскопия.

1. ВВЕДЕНИЕ

Разработка методов обработки и анализа видеoinформации базируется на использовании математических моделей сигнала. Изображение, как двумерный сигнал, отображает некоторую сцену из множества объектов и представляет собой набор областей разнообразной формы, яркостно-цветовых характеристик и текстуры. Среди значительного количества моделей дискретного изображения распространены: авторегрессионные (каузальные), некаузальные, статистические, марковские, основанные на визуальном восприятии и др. [1–4]. Большинство моделей достаточно хорошо описывает вероятностные связи близлежащих элементов, находящихся внутри одной области изображения на расстоянии нескольких шагов дискретизации. Проблемы как правило возникают при необходимости совместного описания групп, составленных из представителей статистически различных множеств элементов — внутренних и контурных, или нескольких соседних областей одновременно.

Одним из важнейших критериев выбора модели является возможность построения на ее основе тех или иных методов и алгоритмов. Соответственно, для разработки алгоритмов, способных работать в условиях сложной смеси статистически различных сигналов, требуется модель, позволяющая описывать подобную ситуацию. В качестве таковой предлагается двухмасштабная модель изображения [5], учитывающая его свойства как на масштабе малого размера — группы из нескольких соседних элементов, так и на масштабе большого размера — группы, состоящей из элементов нескольких разных объектов.

На базе двухмасштабной модели был разработан метод декомпозиции изображения [6], разделяющий его, согласно модели, на две компоненты, содержащие детали разных информационных классов и имеющие различные статистические свойства. Первая, сглаженная компонента, несет информацию о протяженных деталях (объектах) изображения и контурных перепадах между ними, а вторая — текстурная — содержит информацию о мелких деталях, текстуре и шуме.

Сглаженная компонента используется в настоящей работе для построения алгоритма обнаружения объектов и для рассмотрения аспектов обнаружения различий в наборе объектов на паре изображений. Также на основе сглаженной компоненты предлагается альтернативный подход к одной из классических задач предварительной обработки — обнаружению контуров на изображении. В последней части статьи рассматривается использование двухмасштабной модели для решения практической задачи — обнаружения дефектов на снимках электронных микросхем, получаемых микроскопом высокого разрешения.

2. ПРЕДПОСЫЛКИ

2.1. Двухмасштабная статистическая модель изображения

В дискретном виде изображение представляется матрицей $\mathbf{X} = [x_{mn}]$ размерами $M \times N$. Значение каждого элемента x_{mn} есть квантованное на K уровней значение логарифма яркости $0 \leq x_{mn} \leq K - 1$, для краткости называемое просто *яркостью*. Изображение, как средство двумерного отображения информации о некоторой наблюдаемой сцене или предмете, можно рассматривать состоящим из отдельных связанных областей (объектов), плотно покрывающих изображение и соответствующих различаемым деталям сцены или предмета. *Областью* или *объектом* называют максимальное по размеру связанное множество элементов изображения, имеющих близкие, возможно плавно меняющиеся значения яркости. Области могут соприкасаться произвольным образом, в том числе одна область может быть полностью окружена другой. Яркости соседних областей должны заметно различаться. Не соприкасающиеся области могут иметь произвольные, в том числе и совпадающие яркости. Границы между соседними областями, как объектами различающейся яркости, называют *контурами*.

Согласно двухмасштабной модели [5] изображение $\mathbf{X} = [x_{mn}]$ представляется в виде суммы статистически независимых компонент:

$$x_{mn} = S_{mn} + t_{mn} + \xi_{mn}. \quad (1)$$

Первый член суммы — кусочно-гладкая компонента S_{mn} , определяющая средние уровни яркости протяженных областей изображения; t_{mn} — текстурно-детальная компонента, несущая информацию о текстуре и мелких деталях; ξ_{mn} — шумовая компонента, определяемая шумами регистратора, аналого-цифрового преобразователя и др. Все компоненты предполагаются независимыми и аддитивными, а t_{mn} и ξ_{mn} — нормально распределенными и несмещенными.

На масштабе малого размера (*масштабе элементов окрестности*) рассматривается сравнительно небольшое связанное множество элементов, расположенных на расстоянии нескольких шагов дискретизации. Вводится понятие окрестности V_{mn} элемента x_{mn} как группы из R элементов $x_{mn}^r \in V_{mn}, r = 1, \dots, R$, ближайших к x_{mn} и попадающих в то же множество (контурное или фоновое), что и x_{mn} . Методом наименьших квадратов проводится плоскость, наиболее близкая значениям элементов из V_{mn} , составляющая с горизонтальной плоскостью некоторый двугранный угол. Величина и направление данного угла характеризуется вектором \mathbf{g}_{mn} , который приписывается точке (m, n) .

В точке r окрестности проведенная плоскость отличается от значения x_{mn}^r на случайную величину γ_{mn}^r . Такое представление позволяет связать значения элементов $x_{mn}^r \in V_{mn}$ окрестности формулой:

$$x_{mn}^r = \mu_{mn} + \rho^r g_{mn}^r + \gamma_{mn}^r, \quad (2)$$

где μ_{mn} — значение проведенной плоскости в центральной точке окрестности (m, n) , ρ^r — расстояние между центральным элементом x_{mn} и x_{mn}^r , g_{mn}^r — величина проекции \mathbf{g}_{mn} на вектор из x_{mn} в x_{mn}^r , а γ_{mn}^r — случайная величина.

Элементы изображения делятся на два непересекающихся множества: попадающие на контурные участки (контурные) и не попадающие (фоновые), составляющие вместе полное изображение. Вводится понятие контурной маски, совпадающей по размерам с изображением, $\mathbf{E} = [e_{mn}] : e_{mn} = 1$ для контурных и $e_{mn} = 0$ для фоновых элементов. Обозначая для контурных и фоновых элементов g_{mn}^r через φ_{mn}^r и ψ_{mn}^r , а γ_{mn}^r через ζ_{mn}^r и η_{mn}^r соответственно, представим g_{mn}^r и γ_{mn}^r в виде сумм $g_{mn}^r = e_{mn}^r \varphi_{mn}^r + (1 - e_{mn}^r) \psi_{mn}^r$ и $\gamma_{mn}^r = e_{mn}^r \zeta_{mn}^r + (1 - e_{mn}^r) \eta_{mn}^r$. В результате получим окончательную формулу модели окрестности [5], описывающую статистические связи ее элементов:

$$x_{mn}^r = \mu_{mn} + e_{mn}^r (\varphi_{mn}^r \rho^r + \zeta_{mn}^r) + (1 - e_{mn}^r) (\psi_{mn}^r \rho^r + \eta_{mn}^r). \quad (3)$$

Здесь ζ_{mn}^r — стохастическое возбуждение в точке r окрестности для контурных, а η_{mn}^r — для фоновых элементов. Случайные величины φ_{mn} , ψ_{mn} , ζ_{mn} , и η_{mn} считаются некоррелированными и несмещенными, а шумовые составляющие ζ_{mn} , и η_{mn} — нормально распределенными.

На масштабе большого размера (*масштабе объектов фрагмента*) предполагается, что гладкие составляющие S^v тех частей областей u^v ($v = 1, \dots, V$), которые попадают во фрагмент W_{mn} , могут быть представлены полиномом степени не выше, чем ω . Тогда составляющая S_{ij} внутри фрагмента W_{mn} может быть описана формулой:

$$S_{ij}^v(W_{mn}) = \sum_{v=1}^V \delta_{uv} \sum_{p=0}^{\omega} \sum_{q=0}^p a_{pq}^v i^{p-q} j^q; \quad (4)$$

здесь (i, j) — точка области u^v во фрагменте W_{mn} ; $\delta_{uv} = 1$, если точка $(i, j) \in u^v$ и $\delta_{uv} = 0$ в остальных случаях. Добавляя в (4) текстурную t_{ij} и шумовую ξ_{ij} составляющие, получим выражение для значения элемента изображения:

$$x_{ij}^v = \sum_{v=1}^V \delta_{uv} \left(\sum_{p=0}^{\omega} \sum_{q=0}^p a_{pq}^v i^{p-q} j^q + t_{ij}^v + \xi_{ij}^v \right). \quad (5)$$

Это общая формула модели, описывающей значения элементов областей внутри фрагмента. Значения t_{ij}^v считаются нормально распределенными, но возможно с различными значениями дисперсий, различающимися от области к области.

На многих реальных изображениях области, в пределах типичного фрагмента анализа, имеют приблизительно постоянные средние локальные яркости, не меняющиеся заметно. По этой причине во многих случаях допустимо выбрать минимальную степень полинома: $\omega = 0$. Тогда $S_{ij}^v(W_{mn}) = S_{mn}^v$ и (5) преобразуется к виду [5]:

$$x_{ij}^v = \sum_{u=1}^R \delta_{uv} (S_{mn}^v + t_{ij}^v + \xi_{ij}^v). \quad (6)$$

Это формула кусочно-постоянной модели фрагмента для представления участков областей изображения, попадающих во фрагмент W_{mn} .

2.2. Алгоритм декомпозиции (сглаживания) изображения

Рассмотренная модель (6) достаточно проста и хорошо описывает статистику элементов во фрагменте для большинства реальных изображений. На ее основе был разработан ранговый алгоритм сглаживания изображения [6, 7], названный *алгоритмом декомпозиции*, т.е. разделения x_{mn} в (1) на компоненты S_{mn} и $(t_{mn} + \xi_{mn})$.

Алгоритм декомпозиции (сглаживания) [6] при заданных размерах $l \times l$ и $L \times L$ ($l < L$) окрестности V_{mn} и фрагмента W_{mn} , центрированных в точке (m, n) , ширине яркостного интервала анализа Δ^V и Δ^W , а также ранговых параметрах $n^V < l^2/2$ и $n^W < L^2/2$ соответственно, позволяет находить значение S_{mn} в (6). Он включает следующие операции, выполняемые для каждой точки (m, n) изображения:

1. Подсчитываются гистограммы по окрестности H_{mn}^V и фрагменту H_{mn}^W с центром в (m, n) .
2. По гистограмме окрестности H_{mn}^V и заданному значению n^V находятся ранговые параметры $R_1^V = R^V(n^V/l^2)$ и $R_2^V = R^V(1 - n^V/l^2)$, где $R(x)$ определяется как решение уравнения $\sum_{i=0}^{R(x)} h_{mn}^V(i) = x$. Сравнением значения центрального элемента x_{mn} с R_1^V и R_2^V определяется промежуточное усеченное значение \tilde{x}^V :

$$\tilde{x}^V = \begin{cases} x_{mn}, & \text{если } R_1^V \leq x_{mn} \leq R_2^V; \\ R_1^V, & \text{если } x_{mn} < R_1^V; \\ R_2^V, & \text{если } x_{mn} > R_2^V. \end{cases}$$

3. С помощью сигма-фильтра [8], являющегося частным случаем парзеновского окна [9], находится урезанное среднее \bar{x}_{mn} . Для этого из элементов окрестности V_{mn} выбираются те n значений $x_{mn}^r \in V_{mn}$ ($r = 1, \dots, n$), которые попадают в интервал $(\tilde{x}_{mn}^V - \Delta^V, \tilde{x}_{mn}^V + \Delta^V)$, где Δ^V — полуширина интервала. По значениям x_{mn}^r , попадающим в данный интервал, подсчитывается среднее:

$$\bar{x}_{mn} = A(V_{mn}, x_{mn}, n^V, \Delta^V) = \frac{1}{n} \sum_{r=1}^n x_{mn}^r, \quad \tilde{x}^V - \Delta^V \leq x_{mn}^r \leq \tilde{x}^V + \Delta^V. \quad (7)$$

4. Аналогично п. 2, по гистограмме фрагмента H_{mn}^W и заданному n^W находятся ранговые значения $R_1^W = R^W(n^W/L^2)$ и $R_2^W = R^W(1 - n^W/L^2)$. Сравнением значения \bar{x}_{mn} с R_1^W и R_2^W определяется усеченное значение \tilde{x}^W :

$$\tilde{x}^W = \begin{cases} \bar{x}_{mn}, & \text{если } R_1^W \leq \bar{x}_{mn} \leq R_2^W; \\ R_1^W, & \text{если } \bar{x}_{mn} < R_1^W; \\ R_2^W, & \text{если } \bar{x}_{mn} > R_2^W. \end{cases}$$

5. Сглаженное значение S_{mn} находится по гистограмме фрагмента H_{mn}^W как медиана значений H_{mn}^W , попадающих в интервал $(\tilde{x}^W - \Delta^W, \tilde{x}^W + \Delta^W)$:

$$S_{mn} = \text{med}(W_{mn}, \bar{x}_{mn}, n^W, \Delta^W). \quad (8)$$

Полученное значение считается искомой сглаженной компонентой S_{mn} . Возможно дополнительное уточнение значения S_{mn} несколькими шагами итераций, используя в качестве \tilde{x}^W значение S_{mn} , получаемое на предыдущем шаге. Следует отметить, что в (8) вместо медианы в принципе также можно использовать урезанное среднее, как и в (7), но вблизи контурных перепадов медиана дает более точную оценку. Касаясь вычислительной эффективности приведенного алгоритма можно отметить, что для его реализации достаточно в среднем $C_1(l+L) + C_2$ операций на один элемент изображения, где C_1 и C_2 — константы, а l и L — число строк в окрестности V и фрагменте W .

Применение описанного алгоритма показано на Рис. 1 на примере сглаживания изображения „Лепна“ (Рис. 1,а), ставшего классическим. На Рис. 1,б представлена выделенная сглаженная компонента S_{mn} , полученная алгоритмом (8); в правом верхнем углу изображения показан размер фрагмента сглаживания.

Упрощенный вариант построения алгоритма сглаживания изображения на основе метода парзеновского окна с использованием только одного уровня оценивания S_{mn} (по фрагменту, без предварительной оценки по окрестности) и без введения ранговых ограничений стал впоследствии известен под названием *билатеральная фильтрация* [10].



Рис. 1. Декомпозиция изображения: а) исходное изображение; б) сглаженная компонента S_{mn} .

3. ОБНАРУЖЕНИЕ ОБЪЕКТОВ ЗАДАННОЙ ПЛОЩАДИ

Можно сформулировать три варианта задачи обнаружения объектов на изображении по их площади: обнаружение объектов площади больше заданной, меньше заданной и обнаружение объектов, имеющих площадь в заданном интервале. Другими словами, задается некоторое пороговое значение Q (или пара значений: $Q_1 < Q_2$) и требуется найти все области изображения, число элементов N^j внутри которых либо больше, либо меньше Q , либо находится в интервале $Q_1 < N^j < Q_2$.

а) Обнаружение объектов с $N^j > Q$

В данной постановке предполагается, что изображение представляет собой достаточно ровный фон (одна большая область U^0), на которой имеется определенный набор сравнительно небольших областей U^1, \dots, U^J . Воспользуемся алгоритмом декомпозиции (8). Предположим, что области U^1, \dots, U^J отстоят друг от друга достаточно далеко и можно выбрать некоторый размер фрагмента L ($L^2/2 > Q$) такой, что в любой фрагмент W_{mn} попадает не более одной области с $N^j > Q$, либо несколько меньших, но при условии $\sum N^j < Q$ ($U^j \in W$). Тогда в п. 5 алгоритма декомпозиции можно выбрать $n^W = Q$, а R_1^W и R_2^W , соответственно, как $R(Q/L^2)$ и $R(1 - Q/L^2)$. Обработкой исходного изображения по алгоритму (8) с указанными значениями R_1^W и R_2^W , получим [6]:

$$y_{mn} = S_{mn}, \quad (9)$$

т.е. сглаженную компоненту исходного изображения, на которой остались области с $N^j > Q$. Эти области уже легко обнаруживаются детектором со значением порога $S(U^0) \pm \delta$, где $S(U^0)$ — средняя яркость фона, а $\delta < \min_j |S(U^j) - S(U^0)|$; (U^j — яркости соответствующих областей).

Данная задача может иметь важное приложение в неразрушающем контроле (дефектоскопии) при обнаружении дефектов в исследуемых объектах. На Рис. 2,а представлена рентгенограмма детали с внутренними кавернами. Согласно правилам контроля, дефектами считаются лишь те каверны, размеры которых превышают некоторый порог. Обработка данного

изображения вышеописанным алгоритмом с указанием параметра n^W равным площади максимального из допустимых дефектов позволила отфильтровать все мешающие мелкие пятна (изображения мелких каверн). Результат такой обработки (9) показан на Рис. 2,б; на котором, однако, осталась большая область, вызванная особенностями формы детали, и которая будет устранена ниже.

б) Обнаружение объектов с $N^j < Q$

Сглаженная компонента S_{mn} , в (9) содержит лишь области с $N^j > Q$, а значит области с $N^j < Q$ будут содержаться в текстурной компоненте $t_{mn} = (x_{mn} - \xi_{mn}) - S_{mn}$. Для ликвидации возможных шумовых выбросов полезно осуществить фильтрацию t_{mn} , которая в нашем случае фактически происходит на шаге 3 алгоритма декомпозиции (7): $\bar{x}_{mn} \approx x_{mn} - \xi_{mn}$. Собственно выделение объектов с $N^j < Q$ достигается пороговым их обнаружением в тех точках, где $|t_{mn}| \geq \delta$ (δ — порог обнаружения):

$$y_{mn} = \begin{cases} x_{mn}, & \text{если } |\bar{x}_{mn} - S_{mn}| \geq \delta; \\ C, & \text{если } |\bar{x}_{mn} - S_{mn}| < \delta. \end{cases} \quad (10)$$

в) Обнаружение объектов с $Q_1 < N^j < Q_2$

Возможны два варианта решения.

В первом случае сначала выберем $n^W = Q_1$. Тогда сглаженная компонента S_{mn} в (9) будет содержать объекты с $N^j > Q_1$ и не будет содержать объекты с $N^j < Q_1$. Осуществим повторную ее обработку алгоритмом (10), но с $n^W = Q_2$ ($Q_2 > Q_1$). Очевидно, y_{mn} будет содержать лишь искомые объекты, попадающие в диапазон $Q_1 < N^j < Q_2$. Недостаток данного решения — алгоритм получается двухпроходовым.

Второй вариант. Обратим внимание, что при анализе гистограмм по окрестности и фрагменту используются два разных порога (n^V и n^W). Выберем размеры окрестности l и фрагмента L больше обычного — такими, чтобы $l^2 > 2Q_1$ и $L^2 > 2Q_2$. Тогда, задав R_1^V и R_2^V как $R_1^V = R^V(Q_1/l^2)$ и $R_2^V = R^V(1 - Q_1/l^2)$, после операции (7) будем иметь \bar{x}_{mn} , которое уже не содержит области с $N^j < Q_1$. Далее в п. 5 алгоритма декомпозиции, при анализе H_{mn}^W , зададим R_1^W и R_2^W как $R_1^W = R^W(Q_2/L^2)$ и $R_2^W = R^W(1 - Q_2/L^2)$. Получив значение S_{mn} в (8), воспользуемся решающим правилом, аналогичным (10):

$$y_{mn} = \begin{cases} \bar{x}_{mn}, & \text{если } |\bar{x}_{mn} - S_{mn}| \geq \delta; \\ C, & \text{если } |\bar{x}_{mn} - S_{mn}| < \delta. \end{cases} \quad (11)$$

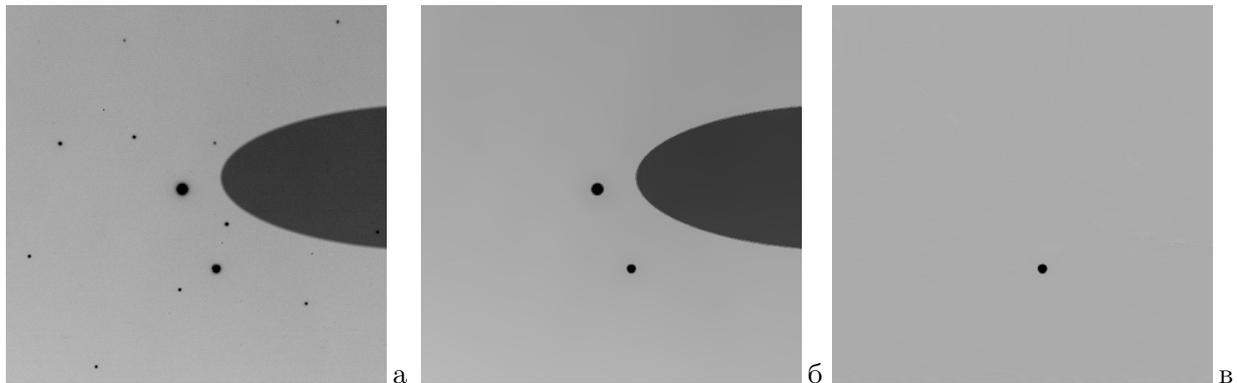


Рис. 2. Обнаружение каверн на рентгенограмме детали: а) исходное изображение; б) обнаружение объектов с площадью больше Q ; в) обнаружение объектов с площадью в интервале от Q_1 до Q_2 .

Как видно, (11) почти полностью совпадает с (10). Действительно, они отличаются лишь тем, что в (11) выбирается значение \bar{x}_{mn} , получаемое после п. 3 алгоритма декомпозиции. Результат показан на Рис. 2,в.

Другой пример обнаружения областей на изображении данным алгоритмом показан на Рис. 3. Здесь аэрофотоснимок (Рис. 3,а) обработан по формуле (11) с различными парами значений Q_1 и Q_2 . В первом случае диапазон выбран в области малых значений и обнаружены объекты малой площади (Рис. 3,б); во втором параметры увеличены и в результате обнаружены объекты большой площади (Рис. 3,в). Отметим, что понятие площади здесь используется несколько в необычном смысле — как „локальная“ площадь, т.е. площадь той части объекта, которая попадает внутрь фрагмента W_{mn} . По этой причине на Рис. 3,б обнаружены протяженные тонкие линии, общее число элементов в которых достаточно велико, но одновременно во фрагмент W_{mn} их попадает сравнительно немного (приблизительно $d \times L$, где d — толщина линии).



Рис. 3. Обнаружение объектов: а) исходный аэрофотоснимок; б) объекты малой локальной площади; в) объекты большой локальной площади.

4. ОБНАРУЖЕНИЕ КОНТУРОВ НА ИЗОБРАЖЕНИИ

Обнаружение контуров является одной из традиционных задач предварительного анализа изображений. Как правило, ее решают путем обработки изображения каким-либо контурным фильтром (локальным оператором) и последующим квантованием полученного отклика на два уровня по выбираемому порогу бинаризации δ . Вопрос выделения контуров на изображении многократно рассматривался в литературе (см., например, [11, 12]). Существенным недостатком стандартного подхода является высокая вероятность ошибки, значительно возрастающая при наличии на изображении шума или текстуры. Ошибки вызывают появление ложных точек, идентифицируемых как контурные, а также разрывы истинных протяженных контурных линий. Причина подобных неудач заложена в самом подходе, поскольку для принятия решения о наличии контура в каждой точке изображения анализируется фактически лишь малая окрестность этой точки.

Опираясь на двухмасштабную модель, можно предложить следующий подход к решению задачи обнаружения контуров. Как уже говорилось ранее, изображение считается состоящим из набора отдельных областей U^1, \dots, U^N , соответствующих объектам реальной сцены, при этом границы между ними являются контурами. Информацию о данных границах несет составляющая S_{mn} в (1), которая считается гладкой внутри отдельных областей изображения и разрывной на контурах. Текстура t_{mn} и шумовая ξ_{mn} компоненты в данном случае являются мешающими, поскольку в них подобной информации не содержится, и, следовательно, они могут лишь вызывать появление дополнительных искажений при обнаружении контуров.

Из сказанного с очевидностью вытекает, что из изображения X следует предварительно выделить сглаженную компоненту S_{mn} , а затем уже на ней обнаруживать контуры. Вопрос выделения компоненты S_{mn} был изложен выше, в разделе 2.2. Алгоритм декомпозиции (8) устраняет текстурную и шумовую составляющие, приводит к значительному сглаживанию изображения и даже к некоторому повышению резкости контурных перепадов [6]. На таком изображении проведение границ между областями достигается практически любым известным оператором выделения контуров. Меняя в алгоритме декомпозиции (8) значение параметра n^W , можно задавать минимальный размер областей, которые выделятся на изображении как отдельные и будут оконтурены. Заметим, что данный подход аналогичен рассмотренному выше в разделе 3.

В качестве примера на Рис. 4 демонстрируются результаты обнаружения контуров на изображении, представленном на Рис. 3,а, используя для формирования контурного отклика простейший дифференциальный оператор Робертса [13]:

$$y_{mn} = ((x_{mn} - x_{m+1,n+1})^2 + (x_{m,n+1} - x_{m+1,n})^2)^{1/2} \quad (12)$$



Рис. 4. Обнаружение контуров: а) по исходному изображению (Рис. 3,а); б) по сглаженной компоненте S_{mn} .



Рис. 5. Обнаружение контуров: а) по исходному изображению (Рис. 1,а); б) по сглаженной компоненте S_{mn} (Рис. 1,б).

с одним и тем же значением порога бинаризации δ . На Рис. 4,а показано выделение контуров по исходному, а на Рис. 4,б — по сглаженному алгоритмом декомпозиции (8). На (а) отчетливо видна разрывность некоторых протяженных контурных границ, наличие ложных контурных линий и отдельных точек, которые вызваны шумом и текстурой. Изображение (б) отличается слитностью контурных линий, меньшей их толщиной и отсутствием помех.

Еще один пример сглаживания и обнаружения контуров показан на Рис. 5. На Рис. 5,а представлено обнаружение контуров по исходному изображению (Рис. 1,а), а на Рис. 5,б — по сглаженной компоненте S_{mn} (Рис. 1,б). Контурные выделялись тем же оператором Робертса (12) с одинаковыми значениями порога бинаризации δ .

5. ОБНАРУЖЕНИЕ РАЗЛИЧИЙ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ

Задача ставится как обнаружение различий двух изображений \mathbf{X}^1 и \mathbf{X}^2 , отображающих сцену, на которой произошли некоторые изменения. Сформулируем постановку задачи. Пусть имеется некоторая исходная сцена \mathbf{V}_0 , состоящая из объектов v_1, \dots, v_M . Зарегистрированные изображения \mathbf{X}^1 и \mathbf{X}^2 отображают в действительности сцены \mathbf{V}^1 и \mathbf{V}^2 , отличающиеся от \mathbf{V}_0 добавлением в одном случае объектов v_{11}, \dots, v_{1K} , а в другом — v_{21}, \dots, v_{2L} . Отметим, что отсутствие объекта на одном изображении можно трактовать как добавление его на другом. Подобным же образом можно трактовать и изменение формы объектов.

На тех участках изображения, где не произошло появления новых объектов, \mathbf{X}^1 и \mathbf{X}^2 являются двумя реализациями одной и той же сцены \mathbf{V}_0 ; обозначим их через \mathbf{X}_0^1 и \mathbf{X}_0^2 . Тем самым \mathbf{X}^1 и \mathbf{X}^2 отличаются от \mathbf{X}_0^1 и \mathbf{X}_0^2 добавлением областей u_{11}, \dots, u_{1k} и u_{21}, \dots, u_{2l} , соответствующих v_{11}, \dots, v_{1k} и v_{21}, \dots, v_{2l} .

Введем некоторые ограничения. Будем считать, что на изображениях \mathbf{X}^1 и \mathbf{X}^2 требуется обнаружить области, площади которых лежат в пределах $N_1 \leq N(u_{ij}) \leq N_2$. Представим изображение \mathbf{X}^1 как сумму изображения \mathbf{X}_0^1 , областей u_{11}, \dots, u_{1p} , площади которых находятся в пределах $N_1 \leq N(u_{ij}) \leq N_2$, и областей w_{11}, \dots, w_{1P} , имеющих площади вне указанных пределов: $N(w_{ij}) < N_1$ или $N(w_{ij}) > N_2$ ($P + p = K$):

$$\mathbf{X}^1 = \mathbf{X}_0^1 \oplus (w_{11}, \dots, w_{1P}) \oplus (u_{11}, \dots, u_{1p}), \quad (13)$$

где операция $a \oplus b$ означает, что в тех точках, где определены элементы как множества a , так и множества b , берутся элементы из b . Аналогично для изображения \mathbf{X}^2 ($Q + q = L$):

$$\mathbf{X}^2 = \mathbf{X}_0^2 \oplus (w_{21}, \dots, w_{2Q}) \oplus (u_{21}, \dots, u_{2q}). \quad (14)$$

Можно считать, что имеются промежуточные изображения \mathbf{X}_w^1 и \mathbf{X}_w^2 :

$$\mathbf{X}_w^1 = \mathbf{X}_0^1 \oplus (w_{11}, \dots, w_{1P}) \quad \text{и} \quad \mathbf{X}_w^2 = \mathbf{X}_0^2 \oplus (w_{21}, \dots, w_{2Q}),$$

отличающиеся наличием областей w_{11}, \dots, w_{1P} и w_{21}, \dots, w_{2Q} , но не различимые в нашей постановке задачи. Отсюда:

$$\mathbf{X}^1 = \mathbf{X}_w^1 \oplus (u_{11}, \dots, u_{1p}) \quad \text{и} \quad \mathbf{X}^2 = \mathbf{X}_w^2 \oplus (u_{21}, \dots, u_{2q}).$$

Выделить на изображении области, имеющие площадь в диапазоне $[N_1, N_2]$, можно при помощи алгоритма (11), рассмотренного в разделе 3,в. Преобразовав таким способом \mathbf{X}^1 и \mathbf{X}^2 , получим сглаженные изображения $\bar{\mathbf{X}}^1$ и $\bar{\mathbf{X}}^2$. На них будут выделены как искомые области u_{11}, \dots, u_{1p} , и u_{21}, \dots, u_{2q} , так и области $u_{11}^0, \dots, u_{1r}^0$, и $u_{21}^0, \dots, u_{2r}^0$ компонент \mathbf{X}_w^1 и \mathbf{X}_w^2 , отображающие повторяющиеся объекты исходной сцены с площадью в том же диапазоне; в данном

случае они являются мешающими. Из (11) следует, что $\bar{\mathbf{X}}^1$ и $\bar{\mathbf{X}}^2$ должны представлять собой ровный фон яркости C , на котором будут содержаться сглаженные области \bar{u}_{ij}^0 и \bar{u}_{ij} :

$$\bar{\mathbf{X}}^1 = C \oplus (u_{11}^0, \dots, u_{1r}^0, u_{11}, \dots, u_{1p}) \quad \text{и} \quad \bar{\mathbf{X}}^2 = C \oplus (u_{21}^0, \dots, u_{2r}^0, u_{21}, \dots, u_{2q}). \quad (15)$$

Согласно модели (1):

$$x_{mn}^1 = S_{mn}^1 + t_{mn}^1 + \xi_{mn}^1 \quad \text{и} \quad x_{mn}^2 = S_{mn}^2 + t_{mn}^2 + \xi_{mn}^2.$$

Для \bar{x}_{mn} в (11) можно показать, что

$$\bar{x}_{mn} = S(u_{mn}^i) + o(D), \quad (16)$$

где $S(u_{mn}^i)$ — средняя яркость той части области u^i во фрагменте анализа, в которую попадает элемент x_{mn} , а D — общий диапазон яркостей.

Пусть изображения \mathbf{X}^1 и \mathbf{X}^2 согласованы по градационным характеристикам, т.е. для областей u_{1i}^0 и u_{2i}^0 , отображающих один и тот же реальный объект v_i на разных изображениях, выполняется:

$$S(u_{1i}^0) = S(u_{2i}^0) + o(D).$$

Тогда разность изображений \mathbf{X}^1 и \mathbf{X}^2 , $z_{mn} = \bar{x}_{mn}^1 - \bar{x}_{mn}^2$, согласно (11) будет выглядеть:

$$z_{mn} = \begin{cases} o(D), & \text{если объекты совпадают,} \\ (\bar{x}_{mn}^1 - \bar{x}_{mn}^2), & \text{если объекты не совпадают.} \end{cases} \quad (17)$$

Таким образом, $z_{mn} = o(D)$ в точках, где \mathbf{X}^1 и \mathbf{X}^2 отображают одни и те же объекты исходной сцены, и, согласно (16), $z_{mn} \approx S(u_{mn}^1) - S(u_{mn}^2)$, если x_{mn}^1 и x_{mn}^2 соответствуют разным объектам. Для наиболее частого случая, когда объект возникает лишь на одном из изображений:

$$z_{mn} = \begin{cases} o(D), & \text{если объекты совпадают,} \\ (\bar{x}_{mn}^1 - C), & \text{если объект возник на изображении 1,} \\ (C - \bar{x}_{mn}^2), & \text{если объект возник на изображении 2.} \end{cases} \quad (18)$$

Для удобства обнаружения объектов положим $C = 0$. Тогда $z_{mn} \leq o(D)$ в точках, где объекты интересующих размеров либо отсутствуют, либо совпадают; $z_{mn} = x_{mn}^1$, если объект возник на изображении 1; и $z_{mn} = -x_{mn}^2$, если объект возник на изображении 2. Выбрав значение порога $\delta \approx \sigma_t + \sigma_\xi$, где σ_t и σ_ξ — стандартные отклонения текстурной t_{mn} и шумовой ξ_{mn} компонент в (1), очевидным становится решающее правило обнаружения объекта на одном или другом изображении в зависимости от факта попадания z_{mn} в интервал $[-\delta, \delta]$, либо знака z_{mn} при выходе за рамки данного интервала.

Тем самым мы получили способ отыскания различий на изображениях. Однако он может давать некоторые погрешности, вызванные следующим. Пусть u_{1i}^0 и u_{2i}^0 — две области изображений \mathbf{X}^1 и \mathbf{X}^2 , отображающие один и тот же объект. В силу неточностей дискретизации и совмещения изображений возможны несовпадения границ данных областей. Из-за этого на разностном изображении \mathbf{Z} могут возникнуть отдельные точки или контурные линии на границах областей. Избавиться от них нетрудно, обнаружив предварительно контурные участки на $\bar{\mathbf{X}}^1$ и $\bar{\mathbf{X}}^2$, и воспользовавшись следующим решающим алгоритмом:

1. Определяется значение разности $d_{mn} = |\bar{x}_{mn}^1 - \bar{x}_{mn}^2|$.
2. Если $d_{mn} < \delta$, то x_{mn}^1 и x_{mn}^2 — элементы одной и той же области и в этом случае $z_{mn} = C$.
3. Если $d_{mn} \geq \delta$, то: а) если (m, n) — контурная точка как на \mathbf{X}^1 , так и на \mathbf{X}^2 , то $z_{mn} = C$; б) если (m, n) — контурная точка на \mathbf{X}^1 и фоновая на \mathbf{X}^2 , то $z_{mn} = \bar{x}_{mn}^1 - C$, если же наоборот, то $z_{mn} = C - \bar{x}_{mn}^2$; в) если (m, n) — фоновая точка и на \mathbf{X}^1 и на \mathbf{X}^2 , то z_{mn} равно тому

значению \bar{x}_{mn}^1 или \bar{x}_{mn}^2 , которое не попадает в диапазон $C \pm \delta$; если же оба выходят за этот диапазон, то значит соответствующие области на изображениях (объекты сцены) перекрываются. В этом случае можно в качестве значения выбрать либо \bar{x}_{mn}^1 , либо \bar{x}_{mn}^2 в зависимости от априорного приоритета изображений, либо взять их разность. В некоторых случаях может потребоваться вместо сглаженных \bar{x}_{mn}^1 или \bar{x}_{mn}^2 в соответствующих точках брать значения элементов исходных изображений x_{mn}^1 или x_{mn}^2 .

6. ОБНАРУЖЕНИЕ ДЕФЕКТОВ НА СНИМКАХ ЭЛЕКТРОННЫХ МИКРОСХЕМ

6.1. Постановка задачи

Исходными данными для обнаружения дефектов электронных микросхем являются снимки участков их поверхности, получаемые микроскопом высокого разрешения. Делаются снимки двух повторяющихся участков — один с вероятным дефектом, второй без дефекта. Изображения имеют одно и то же увеличение и отображают одинаковые структуры. Между объектами на снимках, как правило, имеется сдвиг, но нет поворота. Первое изображение (дефектное) с высокой вероятностью содержит один или несколько дефектов неопределенной формы, второе (эталонное) дефектов не содержит. Задача состоит в нахождении местоположения и формы дефектов на первом изображении.

Виды дефектов.

Реальные дефекты на электронных микросхемах возникают либо вследствие появления посторонних частиц (по сути, „грязи“) на одном из этапов их изготовления, либо из-за нарушения формы деталей схемы. На изображениях они выглядят следующим образом:

- дефекты в виде одиночных или множественных пятен разнообразной формы, размеров и яркости на ровном фоне или перекрывающие несколько первоначальных объектов;
- разрывы протяженных объектов (т.е. проводников и пр.);
- нарушения формы контуров объектов.

Диапазон возможных размеров дефектов — от нескольких элементов до изображения целиком.

Особенности изображений микросхем.

Формы и размеры объектов на поверхности электронных микросхем могут быть совершенно различными, однако разнообразие одновременно возможных их характеристик (имея в виду яркость или отражающую способность) весьма невелико; они сохраняются на протяжении всего объекта и повторяются в других частях изображения. Поэтому изображения микросхем характеризуются постоянством яркостей каждой из областей. Такое представление вполне соответствует рассмотренной выше кусочно-постоянной модели изображения (6). Более того, поскольку текстурная составляющая на самих деталях практически отсутствует, то модель (6) может быть упрощена до следующей:

$$x_{ij}^v = \sum_{u=1}^V \delta_{uv} (S_{mn}^v + \xi_{ij}^v). \quad (19)$$

Согласно такой модели фон может восприниматься как один (пусть и достаточно большой) объект. Шум ξ предполагается нормально распределенным с нулевым средним и дисперсией σ_ξ^2 : $N(0, \sigma_\xi^2)$. Полагаем также, что разность яркостей p и q соседних объектов удовлетворяет неравенству:

$$|S^p - S^q| > 2\sigma_\xi. \quad (20)$$

Из особенностей реальных изображений и дефектов важно отметить следующие:

- шум на изображениях весьма заметный: $\sigma_\xi \approx 2, 5\%$ и более диапазона яркостей;

- градационные характеристики и контрасты объектов на парах изображений (дефектном и эталонном) могут существенно различаться;
- имеются неточности повторения форм объектов на эталонном и дефектном изображениях, поэтому возможно отличие контуров объектов на совмещенных изображениях в 2–3 элемента;
- возможна неравномерность (неконстантность) яркости протяженных объектов;
- дефекты часто имеют размытые очертания, что затрудняет точное определение их границ;
- контраст пятна дефекта может быть весьма мал и близок уровню шума.

Примеры пар изображений (дефектное–эталонное) представлены на Рис. 6 и 7.

6.2. Подход к построению алгоритма обнаружения дефектов

Поскольку изображения поверхности электронных микросхем удовлетворяют достаточно простой кусочно-постоянной модели (19), то для обнаружения различия пары изображений хотелось бы применить подход, описанный в разделе 5. К сожалению ряд особенностей как дефектов (низкий контраст и большие возможные размеры дефектов, превышающие допустимые величины), так и самих изображений (неточности повторения форм объектов) не позволяют им воспользоваться.

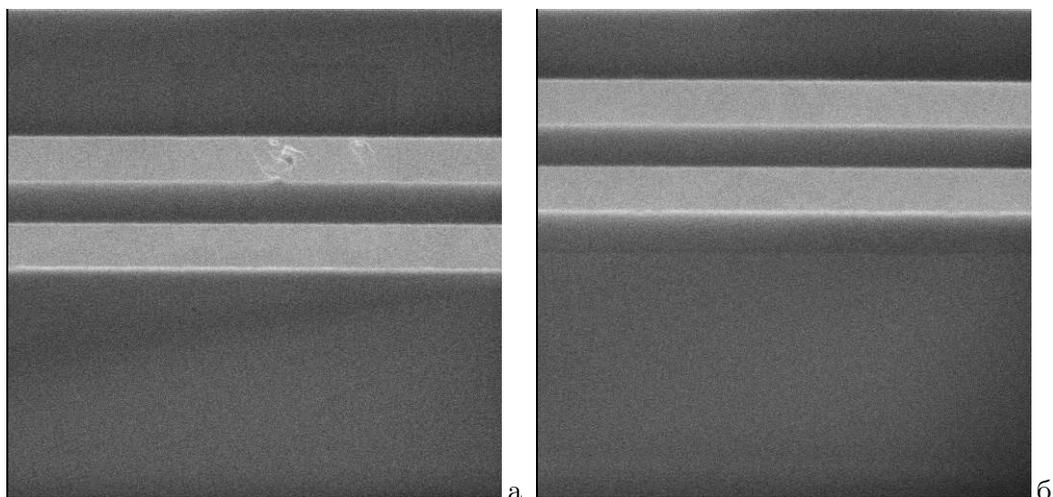


Рис. 6. Фрагменты изображений микросхем: а) с дефектом; б) без дефекта.

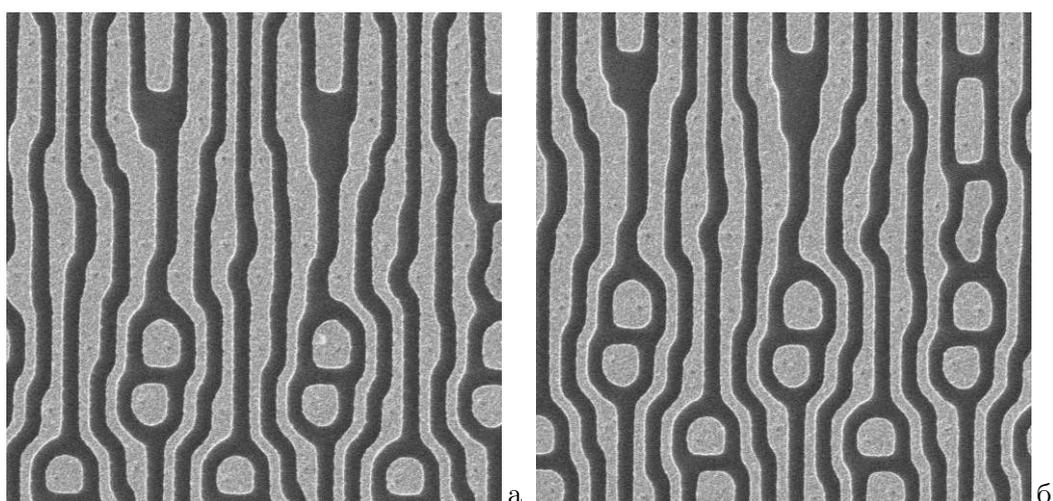


Рис. 7. Фрагменты изображений микросхем: а) с дефектом; б) без дефекта.

Указанные причины, а также требования к вычислительной скорости алгоритма анализа (число операций, требуемое процедурой декомпозиции, несколько превышает допустимое), привели к решению использовать алгоритм, основанный на прямом поиске различий эталонного и дефектного изображений. При таком подходе алгоритм обнаружения дефектов по паре изображений может быть разделен на два последовательных и независимых этапа. Первый — совмещение и предобработка пары изображений, и второй — нахождение области дефекта на уже совмещенных дефектном и эталонном изображениях.

Совмещение изображений.

Первым этапом сопоставления изображений является их совмещение. При всем разнообразии возможных форм контуров объектов удалось подметить две морфологические особенности, такие, что для каждого из анализируемых изображений выполняется хотя бы одна из них: а) значительная часть контуров объектов имеет горизонтальную и/или вертикальную направленность; б) объекты расположены на изображении периодически, и периодичность также имеет вертикальную и горизонтальную составляющие. Совокупность подобных особенностей будем называть *квазирегулярной структурой* объектов изображений, или просто *квазирегулярностью*.

Средством автоматического совмещения изображений, не требующим предварительной семантической интерпретации объектов, является их корреляционное совмещение [14, 15]. В нашем случае удастся воспользоваться свойством квазирегулярности и разработать быстрый алгоритм совмещения изображений [16]. Он сводится к предварительному формированию перпендикулярных проекций изображений, нахождению двумерного набора точек возможного совмещения изображений на основе одномерного корреляционного совмещения проекций, выбора оптимальной точки и окончательному уточнению позиции совмещения.

Возможности анализа совмещенных изображений.

Предполагаем, что дефектное \mathbf{X}_D и эталонное \mathbf{X}_E изображения совмещены. Однако надежно обнаружить дефект по разности совмещенных изображений простым пороговым сравнением не удастся — на разностном сигнале возникает довольно много ложных пятен. Причины их появления кроются в указанных выше особенностях изображений, а именно: в различии форм объектов дефектного \mathbf{X}_D и эталонного \mathbf{X}_E изображений, в неравномерностях яркостей объектов; в высоком уровне шума, в наличии мелких пятен, по своим характеристикам похожих на искомые дефекты. Примеры получаемых таким путем разностных изображений показаны на Рис. 8. Отыскать области дефектов на подобных изображениях без морфологического анализа удастся далеко не всегда.

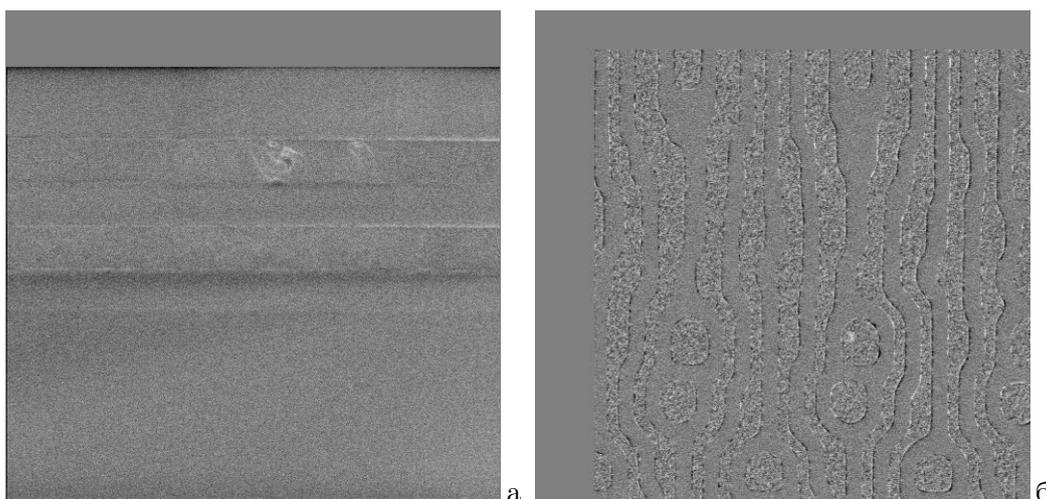


Рис. 8. Разности пар совмещенных изображений: а) на Рис. 6; б) на Рис. 7.

Обнаружение области дефекта становится возможным при совместном анализе пары изображений. Для этого предлагается алгоритм, включающий следующие стадии:

1. Анализ и предобработка исходных изображений. На данной стадии осуществляется взаимное приведение градационных характеристик пары изображений, оценивается уровень шума и выполняется градационная коррекция.
2. Формирование двумерного массива несовпадений. Определяются точки различий пары изображений, формируется двумерная маска несовпадений и осуществляется первичная фильтрация мелких пятен.
3. Формирование и фильтрация списка областей несовпадений. Осуществляется переход от исходного представления данных, как двумерного массива пикселей, к списку описаний каждой отдельной области. Сформированный список областей подвергается последовательным этапам анализа, фильтрации, и упрощения.

6.3. Предобработка изображения

Градационная коррекция.

Для согласования амплитудных характеристик дефектного и эталонного изображений используется следующий алгоритм коррекции. Пусть M_E и σ_E^2 — среднее и дисперсия эталонного, а M_D и σ_D^2 — дефектного изображения. Для согласования изображений выполняется поэлементное преобразование:

$$x_{mn} = \frac{\sigma_E}{\sigma_D}(\tilde{x}_{mn} - M_D) + M_E. \quad (21)$$

Здесь \tilde{x}_{mn} — элемент исходного, а x_{mn} — преобразованного дефектного изображения.

Определение параметров шума.

Подсчитываются некоторые требуемые статистические параметры. Через статистику распределения разности соседних элементов изображения оценивается значение дисперсии шума, имеющегося на дефектном и эталонном изображениях:

$$\Delta x_{mn} = x_{mn} - x_{m,n+1}. \quad (22)$$

Согласно модели (19), шум на изображении независимый и имеет нормальное распределение: $N(0, \sigma_\xi^2)$. Тогда распределение разности элементов (22), находящихся внутри одной области, будет также нормальным: $N(0, 2\sigma_\xi^2)$. К сожалению, изначально нам не известны границы отдельных областей изображения, а значит, при подсчете статистики распределения Δx заметное влияние будут оказывать элементы, попадающие на границы областей (контурные участки). Как следует из неравенства (20), это влияние будет особенно сильным на „хвостах“ распределения. Решить задачу удастся следующим образом [5].

Будем считать, что Δx есть смесь двух нормальных распределений Δx^1 и Δx^2 со значительно различающимися дисперсиями, причем вероятности элементов $P\{\Delta x^2\} \ll P\{\Delta x^1\}$, что справедливо для соотношения числа контурных и фоновых элементов. Выберем некоторые два значения F_1 и F_2 , вблизи средних значений: $0,3 < F_1 < F_2 < 0,7$. Подсчитаем функцию распределения Δx в (22): $F(z) = P\{\Delta x < z\}$ и найдем u_1 и u_2 такие, что $F(u_1) = F_1$, а $F(u_2) = F_2$. Зная функцию нормированного нормального распределения $\Phi(z)$, найдем v_1 и v_2 такие, что $\Phi(v_1) = F_1$, а $\Phi(v_2) = F_2$. Тогда при вышеуказанном соотношении вероятностей Δx^1 и Δx^2 можно оценить дисперсию распределения Δx^1 как: $\sigma_{\Delta x^1} \approx (u_2 - u_1)/(v_2 - v_1)$. При реализации были выбраны $v_1 = -0,6$ и $v_2 = 0,6$; отсюда $F_1 \approx 0,274$ и $F_2 \approx 0,726$, следовательно точки u_1 и u_2 находятся по функции распределения $F(\Delta x)$ как удовлетворяющие равенствам:

$F(u_1) = 0,274$ и $F(u_2) = 0,726$. С учетом того, что дисперсия Δx равна $2\sigma_\xi^2$, получим:

$$\sigma_\xi \approx \frac{u_2 - u_1}{\sqrt{2}(v_2 - v_1)}. \quad (23)$$

Формирование разностного сигнала выравнивания.

Как отмечалось, на реальных изображениях наблюдается заметная неравномерность (неконстантность) фона изображения, которая может привести к возникновению ложных пятен. Для устранения таких артефактов при формировании массива несовпадений учитывается сигнал выравнивания, который вычисляется как разность дефектного и эталонного изображений $d_{mn} = x_D(m, n) - x_E(m, n)$, которая подвергается фильтрации низких частот путем двумерного сглаживания по фрагменту большого размера [7]:

$$\bar{d}_{mn} = \frac{1}{(2h+1)^2} \sum_{i=m-h}^{m+h} \sum_{j=n-h}^{n+h} (x_D(i, j) - x_E(i, j)). \quad (24)$$

Здесь $h \times h$ — размеры фрагмента усреднения ($h \approx 1/6$ размеров изображения).

6.4. Формирование двумерного массива несовпадений

Формирование двумерного массива несовпадений является фундаментальным этапом обнаружения различий пары изображений. Выше уже были приведены причины, мешающие использовать для этого значение их разности. Для уменьшения влияния указанных факторов был разработан следующий алгоритм.

Формирование первичного массива несовпадений.

Основная идея алгоритма формирования массива несовпадений в следующем. Осуществляется поэлементный анализ дефектного \mathbf{X}_D и эталонного \mathbf{X}_E изображений. В каждой точке (m, n) вычисляются интервалы допустимых значений яркости: $[l_D, r_D]$ для дефектного и $[l_E, r_E]$ для эталонного изображений. Считается, что в точке (m, n) наблюдается совпадение, если значения элементов $x_D(m, n)$ и $x_E(m, n)$ каждого из изображений \mathbf{X}_D и \mathbf{X}_E оказываются внутри допустимого интервала, вычисленного по второму изображению.

Опишем алгоритм более подробно. Для каждой точки (m, n) по окрестности 3×3 , окружающей обрабатываемый элемент, находятся оценки: $\min_D(m, n)$ и $\max_D(m, n)$ — минимальное и максимальное значения дефектного изображения; $\min_E(m, n)$ и $\max_E(m, n)$ — то же для эталонного изображения. По значению \bar{d}_{mn} сглаженного разностного сигнала выравнивания (24) определяются инкременты d^+ и d^- :

$$d^+ = \begin{cases} \bar{d}_{mn}, & \text{если } \bar{d}_{mn} > 0, \\ 0, & \text{если } \bar{d}_{mn} \leq 0; \end{cases} \quad \text{и} \quad d^- = \begin{cases} 0, & \text{если } \bar{d}_{mn} \geq 0, \\ \bar{d}_{mn}, & \text{если } \bar{d}_{mn} < 0; \end{cases}$$

Значения границ допустимых интервалов l_D, r_D, l_E и r_E вычисляются по следующим формулам, в которых коэффициенты K_1 и K_2 определяются исходя из величины измеренного шума σ на изображении (раздел 6.3):

$$\begin{aligned} l_D &= \min_D(m, n) - K_1[\max_D(m, n) - \min_D(m, n)] - K_2 - d^+; \\ r_D &= \max_D(m, n) + K_1[\max_D(m, n) - \min_D(m, n)] + K_2 - d^-; \\ l_E &= \min_E(m, n) - K_1[\max_E(m, n) - \min_E(m, n)] - K_2 + d^-; \\ r_E &= \max_E(m, n) + K_1[\max_E(m, n) - \min_E(m, n)] + K_2 + d^+. \end{aligned}$$

Считается, что в точке (m, n) имеется совпадение, если одновременно выполняются условия:

$$\begin{cases} l_D(mn) \leq x_E(mn) \leq r_D(mn); \\ l_E(mn) \leq x_D(mn) \leq r_E(mn). \end{cases} \quad (25)$$

Иначе точка (m, n) относится к области возможного дефекта. По окончании процедуры формируется первичная маска несовпадений, элементы которой равны нулю в точках совпадений и $(x_D - x_E)$ в точках несовпадений.

Удаление шумовых выбросов.

Высокий уровень шума на исходных изображениях приводит к появлению на первоначальной маске несовпадений большого количества мелких всплесков, состоящих из одного или нескольких элементов. Для их удаления используется следующий простой алгоритм.

Для каждого ненулевого элемента по маске несовпадений проверяются 8 его ближайших соседей. Если число окружающих ненулевых элементов, имеющих тот же знак, что и центральный элемент, меньше или равно некоторому заданному порогу, то текущий элемент удаляется из маски несовпадений и обнуляется.

6.5. Формирование списка областей несовпадений

Во многих случаях пятна дефектов не являются ни самыми яркими, ни самыми большими по площади областями несовпадений. Для их обнаружения требуется анализ участков несовпадений, определение и удаление ложных областей, и окончательный выбор одной или нескольких областей, идентифицируемых как дефектные области.

На данной стадии осуществляется переход от анализа исходных данных как двумерного изображения, к описанию и анализу данных как набора отдельных объектов (пятен), являющихся кандидатами на искомые области дефекта. При этом, конечно же, остаются доступными исходные двумерные данные (изображения) но их использование становится вспомогательным, а основным материалом анализа становится одномерный набор дескрипторов, характеризующий области несовпадений.

Операции производятся со следующими объектами:

- а) **пятно** — дескриптор, описывающий отдельную область на изображении и содержащий всю требуемую информацию об области несовпадений;
- б) **список пятен** — структура, объединяющая все пятна на изображении и обеспечивающая их упорядочение, нумерацию и ассоциативные связи;
- в) **список ассоциаций** — промежуточная таблица, объединяющая нескольких объектов в один.

Формирование первичного списка пятен.

Формирование списка пятен несовпадений включает следующие операции:

- анализ двумерной маски несовпадений, построение первичного списка пятен и списка ассоциаций для объединения соприкасающихся пятен;
- объединение ассоциированных пятен и корректировка списка пятен;
- подсчет требуемых характеристик пятен.
- удаление пустых мест в списке пятен и сортировка списка по качеству.

При формировании списка пятен для каждого пятна формируются и сохраняются следующие характеристики:

N — число точек в пятне (площадь пятна);

m_t, m_b, n_l, n_r — координаты прямоугольника, описывающего пятно;

ζ^+ — сумма положительных значений разностей элементов $(x_D - x_E) > 0$;

ζ^- — сумма отрицательных значений разностей элементов $(x_D - x_E) < 0$;

Q — „качество“ пятна, определяемое функционалом

$$Q = \frac{(\zeta^+ - \zeta^-) \log(k + 32)}{N + 16}.$$

Здесь $k = N$ для малых ($N < 64$), и $k = 64$ для больших N .

Удаление периферийных и мелких пятен.

Первоначально удаляются пятна, отстоящие от краев изображения на расстояние менее $1/20$ размеров изображения, а также пятна, отстоящие от краев изображения на расстояние менее $1/10$ размеров изображения и не попадающие в число первых 16 по качеству. Затем удаляются пятна низкого качества, расположенные достаточно далеко от пятен высокого качества. Находится пятно с максимальным уровнем качества и определяется порог, выше которого пятно считается приемлемым, а ниже — слабым. Удаляются те слабые пятна, которые не находятся вблизи пятен приемлемого качества. Оставшиеся пятна малого размера либо ассоциируются с соседними, если таковые есть, либо удаляются. Результат показан на Рис. 9.

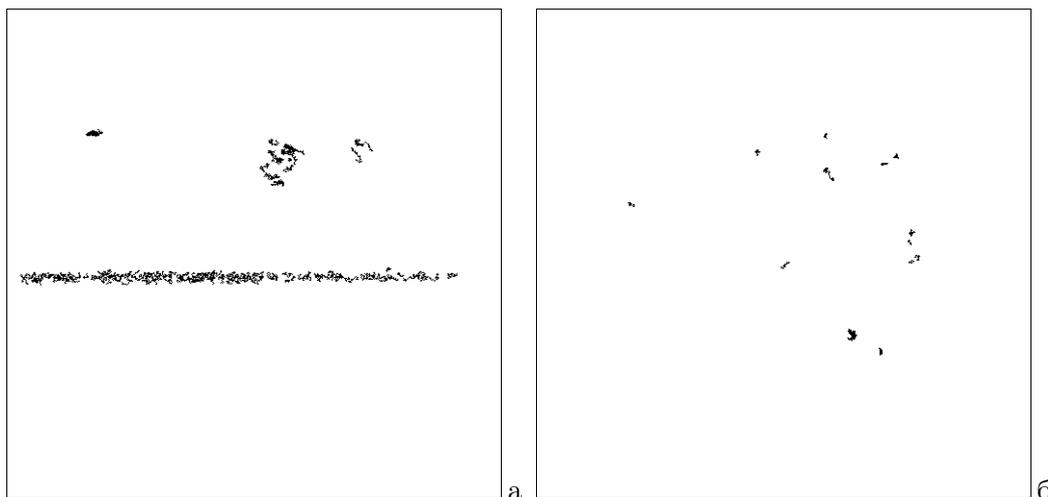


Рис. 9. Вид маски после удаления периферийных и мелких пятен.

Затем над маской пятен выполняется морфологическая операция „замыкание“, состоящая в последовательном выполнении операций дилатации и эрозии [12]. Эта пара операций достаточно часто используется в задачах обнаружения объектов и формирования масок. После чего производится еще один этап обнаружения и удаления пятен малого размера и формируется обновленный список пятен.

6.6. Удаление ложных пятен

После выполнения вышеуказанных процедур на сформированной маске еще остается заметное число ложных пятен. Их удаление происходит на финальном этапе. Для этого подсчитываются дополнительные характеристики оставшихся пятен и формируются некоторые новые признаки. Сочетание этих признаков позволяет принять решение о том, какие пятна действительно соответствуют дефектам, а какие являются ложными и должны быть удалены.

Формируются следующие дополнительные характеристики пятен:

L — длина границы пятна, измеряемая как суммарное число сторон пограничных элементов пятна, контактирующих с фоном;

\sum_D^+ — сумма положительных разностей элементов $(x_d - x_b) > 0$ на всех граничных элементах пятна по дефектному изображению (x_d — элемент пятна, x_b — соседний элемент фона);

\sum_D^- — сумма отрицательных разностей элементов $(x_d - x_b) < 0$ на всех граничных элементах пятна по дефектному изображению;

\sum_E^+ — сумма положительных разностей элементов $(x_d - x_b) > 0$ на всех граничных элементах пятна по эталонному изображению;

\sum_E^- — сумма отрицательных разностей элементов $(x_d - x_b) < 0$ на всех граничных элементах пятна по эталонному изображению;

L_M — число контурных элементов с совпадающим знаком разности $(x_d - x_b)$;

R — удаленность пятна от центра анализируемого участка изображения ($0 \leq R \leq 1$);

D — разность яркостей, определяемая как $D = (\zeta^+ - \zeta^-)(1 - R^2)/N$;

I_{true}, I_{false} — два набора битовых индикаторов, в которых пятно помечается как „правильное“ или „ложное“ по тем или иным признакам.

C — „контраст“ пятна, определяемый как функционал

$$C = \frac{K \left| \left| \sum_D^+ - \sum_D^- \right| - \left| \sum_E^+ - \sum_E^- \right| \right|}{L} + \frac{\zeta^+ - \zeta^-}{N};$$

здесь K — коэффициент, выравнивающий соотношение между средней контурной разницей и средней разницей элементов по всему пятну дефектного и эталонного изображений (оцениваемый по тестовым изображениям как $K = 6$).

F — фактор пятна (функционал характеристик, $0 \leq F \leq 1$), определяющий его „видность“; изначально устанавливается в 1 и изменяется при анализе.

При подсчете характеристик осуществляется нормализация значений качества Q и контраста C по их максимальным значениям: $Q = Q_0/Q_{max}$, $C = C_0/C_{max}$. Таким образом, как качество, так и контраст принимают значения в диапазоне от 0 до 1.

Стирание пятен, ограниченных совпадающими контурами.

Проверяется соотношение $L_M \geq 0,5L$ для каждого из оставшихся пятен. Если оно выполняется, то считается, что данный контур очерчивает повторяющиеся по форме детали, находящиеся на анализируемом и эталонном изображениях, но различающиеся по контрасту с фоном. Поскольку разница в контрастах деталей не считается дефектом, такие объекты удаляются из маски дефектов.

Подавление пятен, удаленных от центра изображения.

Подсчитывается расстояние r между центром пересекающейся части изображений и центром каждого пятна (под этим понимается центр прямоугольника, описывающего пятно). Затем формируется фактор R удаленности пятна от центра изображения, который нормируется на размер половины диагонали изображения: $R = 2r/d$. Формируется значение разности яркостей D . Подавление пятен, удаленных от центра, состоит в уменьшении значения фактора пятна F по формуле $F = (1 - R)F_0$.

Выбор главного пятна и подавление удаленных от него пятен.

Главное пятно определяется как то, на котором достигается максимум функционала

$$\varphi = QCF \log N, \quad (26)$$

являющегося произведением следующих характеристик пятна: качества, контраста, фактора, и логарифма его размеров. Далее также будут использоваться значения размера N_{main} и функционала φ_{main} главного пятна.

Как и выше, подавление пятен сводится к модификации (уменьшению) значений факторов F пятен, удаленных от главного пятна. Для этого первоначально находится базовый допустимый радиус R_0 как функция линейных размеров главного пятна и выбранного минимального значения R_{min} :

$$R_0 = 2(m_b - m_t + n_r - n_l) + R_{min};$$

здесь $2(m_b - m_t + n_r - n_l)$ — периметр прямоугольника, описывающего главное пятно. Если оказывается, что $R_0 > 2R_{min}$, то значение ограничивается $R_0 = 2R_{min}$.

Осуществляется проверка всех оставшихся пятен. Первоначально задается допустимый радиус R . Для обычных пятен он выбирается как $R = R_0$. Для мощных пятен, удовлетворяющих условиям: $N > 0,3N_{main}$ и $\varphi > 0,5\varphi_{main}$ и не имеющих маркированных битовых полей в I_{false} , выбирается большее значение: $R = 3R_0$.

Затем подсчитывается расстояние r между центром анализируемого пятна и центром главного пятна. Фактор пятна F подвергается преобразованию:

$$F = \begin{cases} F, & \text{если } r < R, \\ F(2R - r)/R, & \text{если } R < r < 2R; \\ 0, & \text{если } r > 2R. \end{cases}$$

Удаление оставшихся лишних пятен и формирование результата.

Значения контраста C , качества Q , размера N , а также разности яркостей D для каждого из пятен проверяются на условия: $C \geq k_c$, $Q \geq k_q$, $N \geq k_N N_{max}$ и $D \geq k_d D_{max}$. В зависимости от выполнения этих условий устанавливаются соответствующие индикаторы I_{true} или I_{false} . Все пятна, имеющие ненулевые значения индикатора I_{false} , а также те, у которых низкий фактор пятна ($F < 0,1$), удаляются.

Формируется двухградационное изображение, на котором оставшиеся пятна составляют области дефектов. Полученное изображение считается окончательным результатом — маской пятен дефектов (Рис. 10).

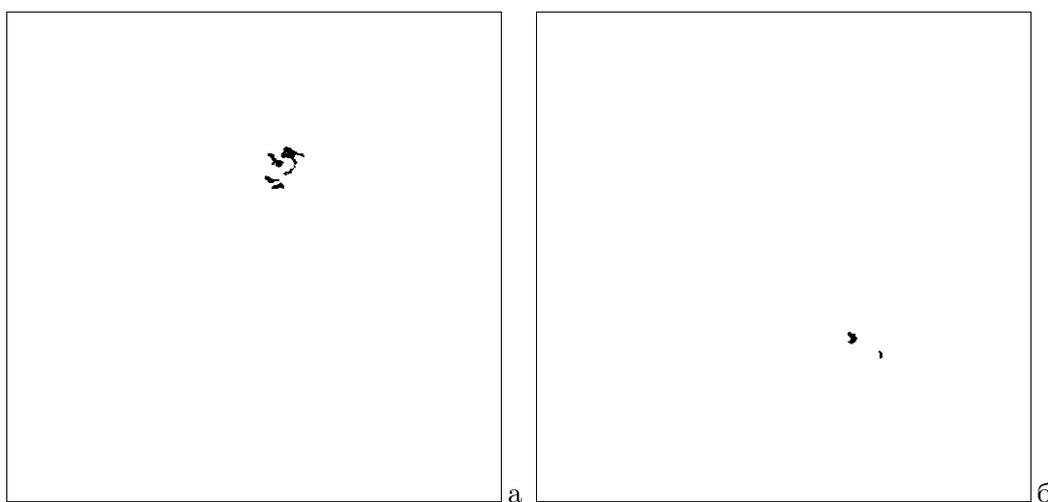


Рис. 10. Сформированные маски пятен дефектов по изображениям: а) Рис. 6; б) Рис. 7.

7. ЗАКЛЮЧЕНИЕ

На основе использования двухмасштабной модели изображения и алгоритма декомпозиции разработаны эффективные методы обнаружения объектов и выделения контуров на изображении. Рассмотрены теоретические аспекты обнаружения различий изображений с малыми изменениями.

Разработанный алгоритм обнаружения дефектов на снимках электронных микросхем был проверен на сотнях образцов снимков и использован в процедуре диагностики при промышленном производстве интегральных микросхем.

Показано, что двухмасштабная модель изображения может являться подходящим фундаментом для построения эффективных алгоритмов обработки и анализа видеoinформации.

СПИСОК ЛИТЕРАТУРЫ

1. Джайн А. К. Успехи в области математических моделей для обработки изображений. *ТИИЭР*, 1981, т. 69, № 5, с. 9–39.
2. Розенфельд А., Дэвис Л. С. Сегментация и модели изображений. *ТИИЭР*, 1979, т. 67, № 5, с. 71–81.
3. Li S. Z. *Markov Random Field Modeling in Image Analysis*. Springer, 2009.
4. Стокхэм Т. Дж. мл. Обработка изображений в контексте моделей зрения. *ТИИЭР*, 1972, т. 60, № 7, с. 93–108.
5. Чочиа П. А. Двухмасштабная модель изображения. В кн. *Кодирование и обработка изображений*. М.: Наука, 1988, с. 69–87.
6. Чочиа П. А. Сглаживание изображения при сохранении контуров. В кн. *Кодирование и обработка изображений*. М.: Наука, 1988, с. 87–98.
7. Chochia P. A. Image Enhancement Using Sliding Histograms. *Computer Vision, Graphics, Image Processing*, 1988, vol. 44, no. 2, pp. 211–229.
8. Lee J.-S. Digital Image Smoothing and the Sigma Filter. *Computer Vision, Graphics, Image Processing*, 1983, vol. 24, no. 2, pp. 255–269.
9. Parzen E. On the estimation of a probability density function and mode. *Annals of Mathematical Statistics*, 1962, vol. 33, pp. 1065–1076.
10. Tomasi C., Manduchi R. Bilateral filtering for gray and color images. *Proc. IEEE 6th Int. Conf. on Computer Vision, Bombay, India*, IEEE, 1998, pp. 839–846.
11. Прэтт У. *Цифровая обработка изображений*. М.: Мир, 1982, т. 1, 2.
12. Гонсалес Р., Вудс Р. *Цифровая обработка изображений*. М.: Техносфера, 2012.
13. Робертс Л. Автоматическое восприятие трехмерных объектов. *Интегральные роботы*. М.: Мир, 1973, с. 162–208.
14. Gruen A. W. Adaptive least squares correlation: a powerful image matching technique. *S. Afr. J. of Photogrammetry, Remote Sensing and Cartography*. 1985, no. 14(3), pp. 175–187.
15. Ackermann F. Digital image correlation: Performance and potential application in photogrammetry. *Photogrammetric Record*. 1984, no. 11(64), pp. 429–439.
16. Чочиа П. А. Быстрое корреляционное совмещение квазирегулярных изображений. *Информационные процессы*. 2009, т. 9, № 3, с. 117–120.

Статью представил к публикации член редколлегии В. И. Венец

Several Algorithms for Object Detection on the Base of Two-Scale Image Model

Chochia P. A.

Two-scale image model is formulated. The algorithm for image decomposition and extracting the smoothed component is argued. The object detection algorithm is expounded. New approach for contour detection is proposed. The theoretical aspects of detecting the differences of object sets in two images is considered. The practical task of developing the algorithm for defect detection in the photographs of electronic microcircuits is decided.

KEYWORDS: image processing, analysis, object detection, image model, nondestructive testing.