4.1.СВИ в полноцветные изображения

Видимые ЦВЗ

1. Видимые и невидимые ЦВЗ: особенности

	Неразличимые ЦВЗ	Видимые ЦВЗ
Различимость результата встраивания ЦВЗ	Неразличимый шум	Визуально заметный и интерпретируемый шаблон встраивания
Способ извлечения	Блок извлечения с численным восстановлением встроенной информации	Визуальное наблюдение
Возможные атаки	Явная атака с целью удаления ЦВЗ	Типовая обработка изображения или помимо этого умышленная атака
Учёт особенностей человеческого зрения	Важен	Очень важен
Блочная структура ЦВЗ	Встречается (в Spread Spectrum Watermarking, а также для повышения объёма ЦВЗ)	Используется почти всегда для защиты от вырезания фрагмента

Простой аддитивный видимый ЦВЗ

- Подготовка (циклическое повторение ЦВЗ)

$$W(n_1, n_2) = W_r\,(n_1(mod\ M_1), n_2(mod\ M_2)),$$

- Встраивание информации:

$$C^W(n_1, n_2) = \begin{cases} C(n_1, n_2) + \alpha \cdot \beta(n_1, n_2) \cdot W(n_1, n_2), & C(n_1, n_2) < 128, \\ C(n_1, n_2) - \alpha \cdot \beta(n_1, n_2) \cdot W(n_1, n_2), & C(n_1, n_2) \geq 128. \end{cases}$$

  - Зачем эта схема со сложением и вычитанием?

- $\beta(n_1, n_2)$ может быть рассчитана как средняя яркость локальной окрестности

СВИ-18 (Kankanhalli & Ramakrishnan): Система видимых ЦВЗ в области блочного ДКП

- Контейнер C разбивается на блоки $C_{ij}(n_1, n_2)$ размерами 8×8
- Блочное ДКП, результат - $f_{ij}(m_1, m_2)$
- Встраиваемая информация представляется в виде матрицы W того же размера, что и контейнер
  - если изначально она задана логотипом W_r, то он циклически повторяется
- Далее к ней, как и к контейнеру, применяется блочное ДКП, результатом чего являются компоненты $\Omega_{ij}(m_1, m_2)$

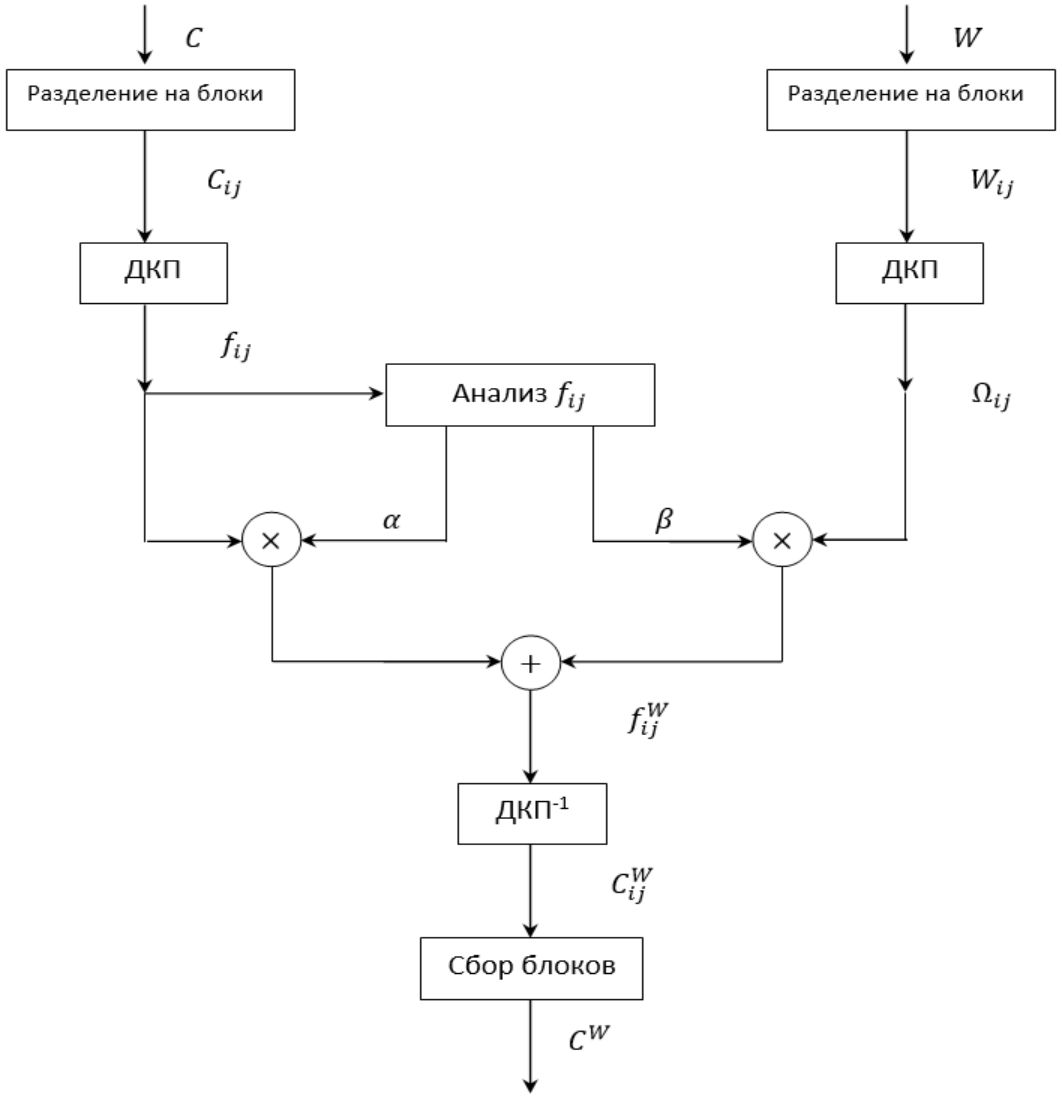
СВИ-18 (Kankanhalli & Ramakrishnan): Система видимых ЦВЗ в области блочного ДКП

- По компонентам $f_{ij}(m_1, m_2)$ производится анализ характеристик:
  - наличие границ и их резкость;
  - наличие однородных областей (яркость которых близка к посто-янной);
  - наличие текстур;
  - средняя яркость блока.
- В результате настраиваются коэффициенты α_{ij} и β_{ij}, которые далее используются в аддитивной процедуре встраивания информации

$$f_{ij}^{W}(m_1, m_2) = \alpha_{ij} \cdot f_{ij}(m_1, m_2) + \beta_{ij} \cdot \Omega_{ij}(m_1, m_2).$$

- Типичные значения α_{ij} варьируются в диапазоне от 0.95 до 0.99, а значения β_{ij} – между 0.01 и 0.15

СВИ-18 (Kankanhalli & Ramakrishnan): Система видимых ЦВЗ в области блочного ДКП



Геометрически стойкие ЦВЗ

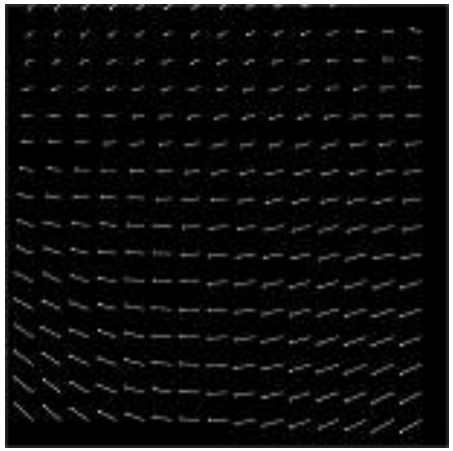
Введение

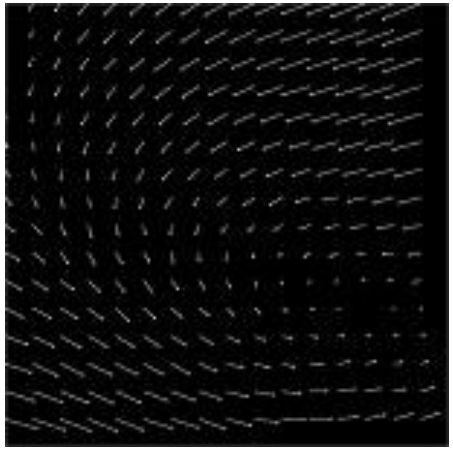
- Как правило, существенное ухудшение качества извлечения ЦВЗ происходит после размытия и геометрических преобразований
  - первое сопровождается существенной деградацией изображения
  - второе может практически не влиять на качество, но разрушать ЦВЗ *а что будет с PSNR?*
- => достижение стойкости к геометрическим преобразованиям – важная и сложная задача при проектировании систем стойких ЦВЗ
- Геометрическое искажение – это атака десинхронизации: ЦВЗ не уничтожается, но декодер из-за десинхронизации не способен извлечь информацию

Примеры геометрических искажений

- Сдвиг
  - Циклический сдвиг
- Поворот
- Масштабирование
- Вырезание фрагмента
- RST
- Аффинное преобразование
- Проективное преобразование
- Локальные деформации

Random bending attack (RBA) – локальные искажения

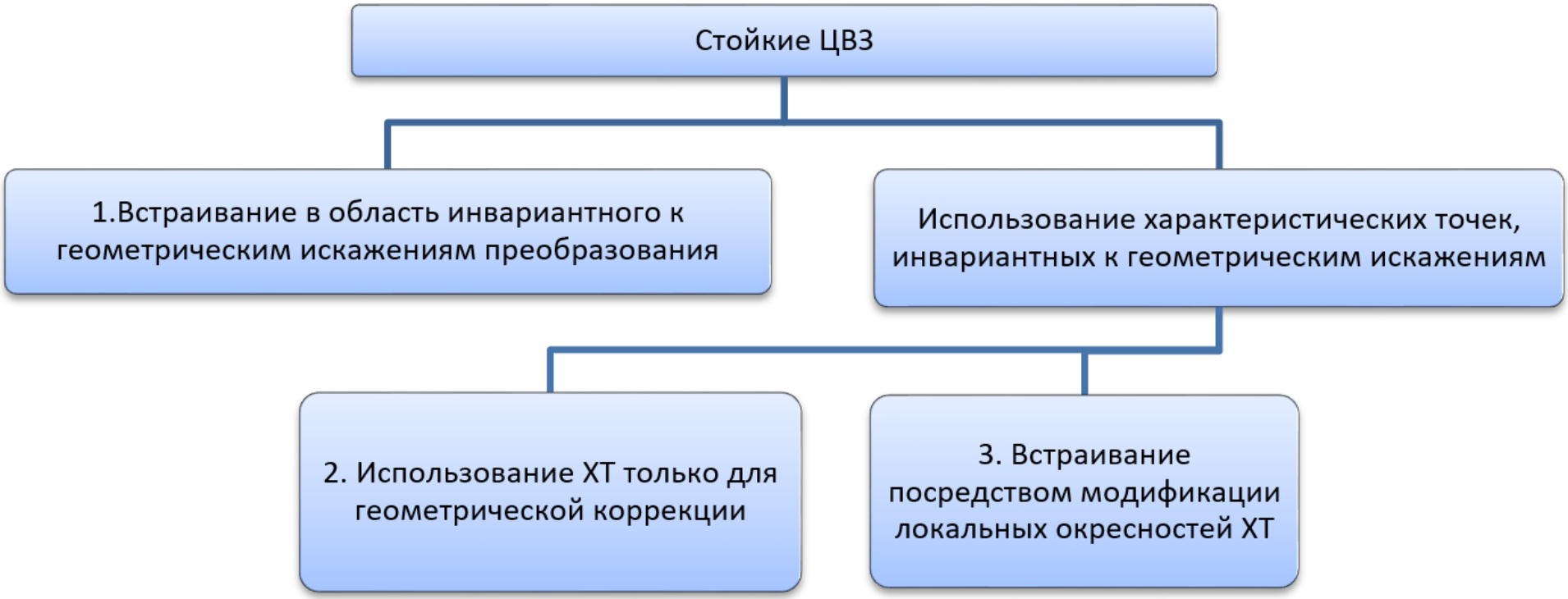




Методы повышения стойкости ЦВЗ к геометрическим искажениям

- Использование исходного контейнера при извлечении информации
- Полный поиск возможных вариантов
- Использование синхронизирующего шаблона (например, второй вспомогательный ЦВЗ)
- Самосинхронизирующиеся ЦВЗ
- …

Варианты, которые мы рассмотрим подробнее



Подход 1: СВИ Zheng & Zhao

- Основан на преобразовании Фурье-Меллина
- Обеспечивает стойкость к RST
- $\text{ДПФ} \rightarrow |\cdot| \rightarrow log.polar \rightarrow \text{ДПФ} \rightarrow |\cdot|$
- Если упрощённо, то далее может быть использована копия СВИ Cox et al. или Piva et al.
- Встраивание информации

$$f^w(m_1, m_2) = f(m_1, m_2) + \alpha \cdot \beta(m_1, m_2) \cdot \Omega(m_1, m_2)$$

- $\Omega(m_1, m_2)$ должна быть положительно определённая.
  - Если $\Omega(m_1, m_2)$ формируется из последовательности, реализующей нормальное распределение, то используется удвоение:

$$x \geq 0 => (x, 0);\ x < 0 => (0, x)$$

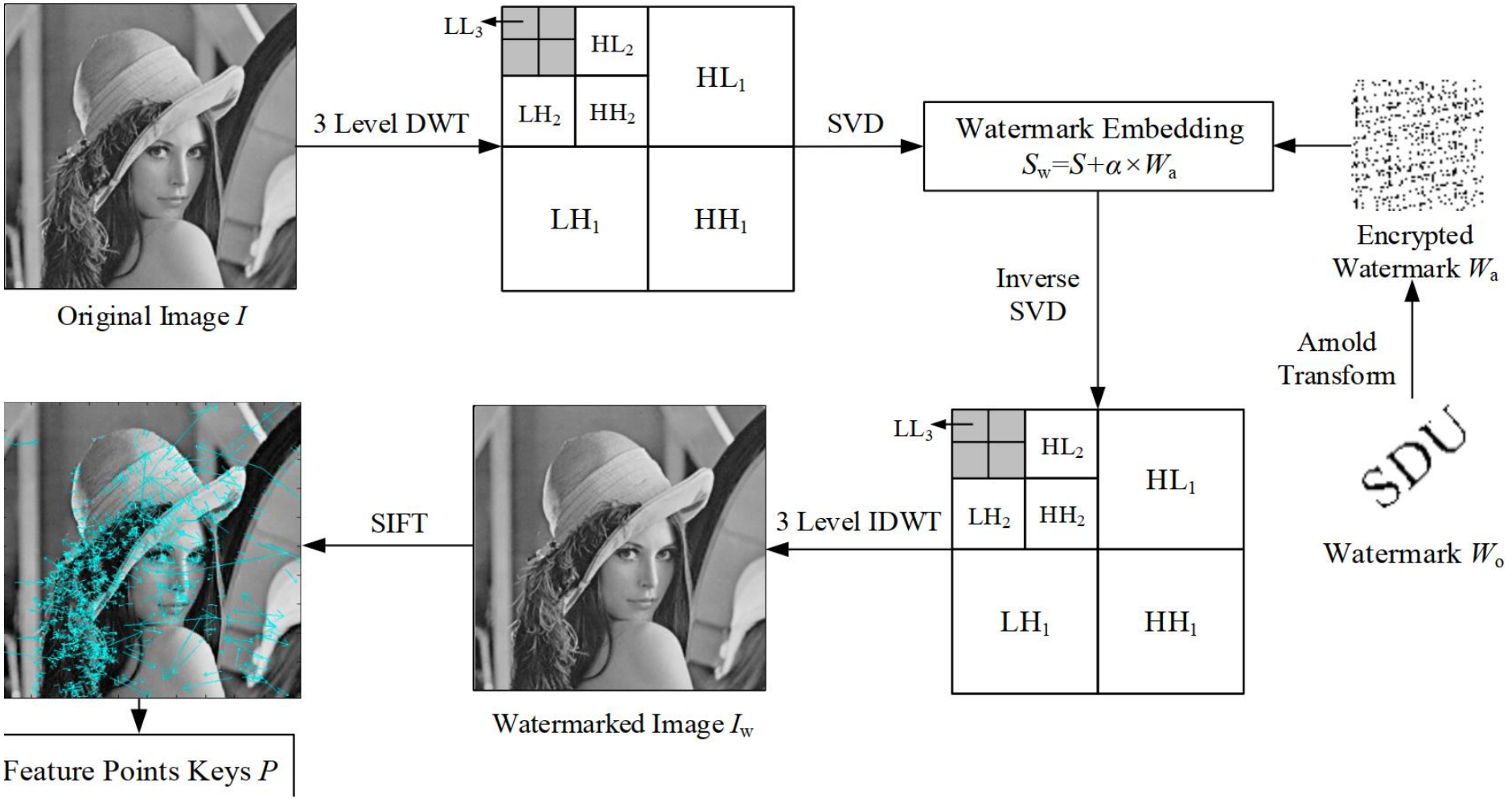
Характеристические точки

- Материал представлен в отдельной презентации “LocalFeatures.pdf”

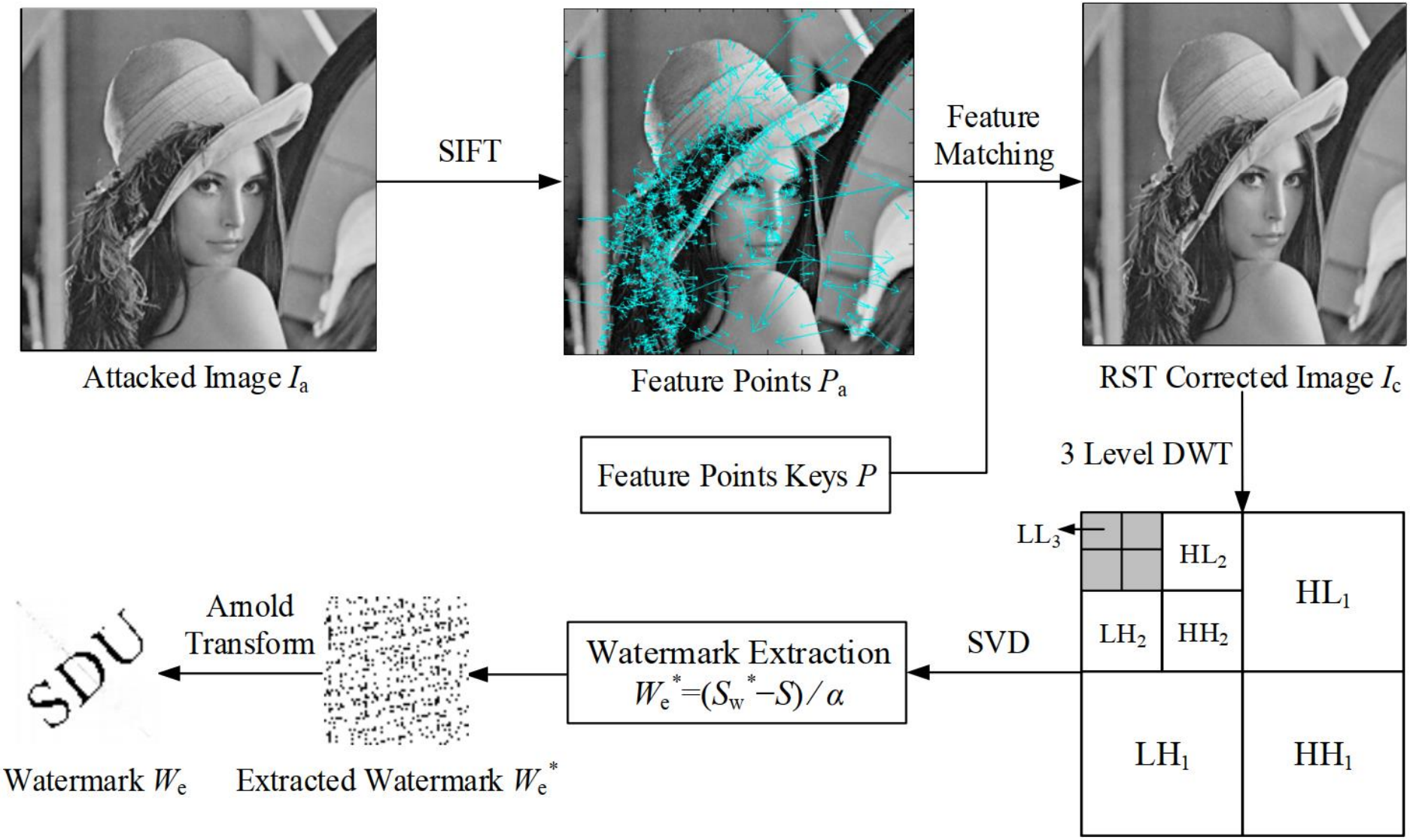
Подход 2: Zhang et al.

- Используется SIFT
- Расчёт характеристических точек происходит после вейвлет-преобразования (для фильтрации незначительной информации)
- Мультипликативное встраивание
- На выходе – детектор (нет извлечения неизвестной встроенной информации)

Подход 1: СВИ Zheng & Zhao



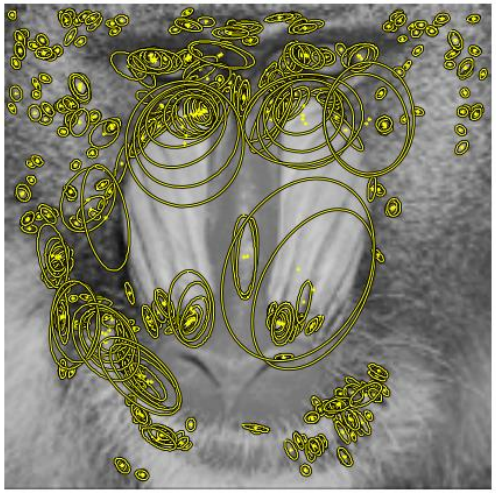
Подход 1: СВИ Zheng & Zhao

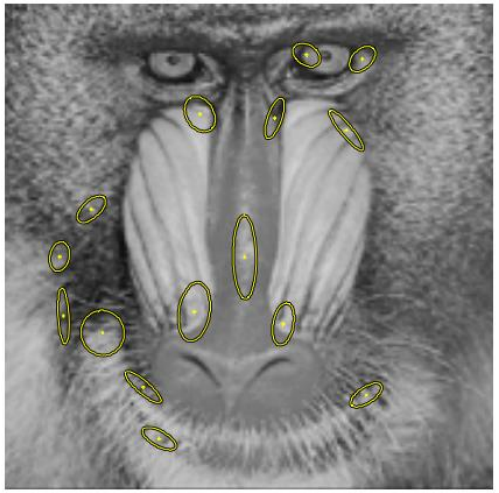


Подход 3: СВИ Deng et al.

- Ищутся точки Харриса-Лапласа, их дескрипторы, опирающиеся на эллиптические области.
- Исключаются точки с накладывающимися друг на друга локальными областями дескрипторов.
- Определяется направление эллипса θ.
- Эллиптические регионы нормируются до кругов радиуса R (рассчитывается из матрицы вторых моментов как корень из суммы квадратов собственных чисел)

Подход 1: СВИ Deng et al.





Исходные характеристические области

Отобранные характеристические области

Подход 3: СВИ Deng et al.

- W – двоичная матрица с нулевым средним: $W(n_1, n_2) \in \{-1,1\}$
- Отображается в окружность (по секторам и сегментам)
- $\beta(n_1, n_2) = \frac{1}{1+\frac{\alpha \cdot \sigma^2(n_1,n_2)}{\sigma^2_{max}}}$, D –весовой множитель.
- Аддитивное добавление W во все нормированные регионы
- $C^W(n_1, n_2) = \alpha_1 \beta(n_1, n_2) C(n_1, n_2) + \alpha_2 \big(1 - \beta(n_1, n_2)\big) W(n_1, n_2)$
- Извлечение: те же шаги → Винеровский фильтр → нормированная линейная корреляция (W, W^*).

Аутентификация изображений при помощи ЦВЗ-систем

1. Задачи аутентификации изображений

- В ряде задач существует необходимость проверки подлинности или целостности мультимедийных данных
- Вопросы, определяющие уровень детальности решения задачи аутентификации
  - Был ли объект каким-либо образом изменён?
  - Если да, то были ли эти изменения значительными?
  - Какие фрагменты подверглись изменению?
  - Может ли объект быть восстановлен?

1. Основные подходы к решению задач аутентификации

- Два подхода:
  - Пассивные
    - Условие применения: изображение доступно один раз при проверке
    - Заключаются в расчёте ряда характеристик объекта или его фрагментов и выявлении статистических или иных аномалий:
      - Глобальные характеристики и глобальные аномалии: на ансамбле типовых изображений
      - Локальные характеристики и локальные аномалии: внутри анализируемого изображения
      - Несоответствия физических или геометрических характеристик
  - Активные
    - Условие применения: изображение доступно два раза: до возможных искажений и после
    - Криптографические методы и методы на основе ЦВЗ

1. Методы, используемые в рамках обоих подходов

- Пассивные
  - Как правило, ориентированы на конкретное искажение (Copy-Move,..)
  - Могут вылавливать факт двойного сохранения (в JPEG, например)
  - Могут выявлять физические или геометрические несоответствия:
    - Тени и блики в разные стороны
    - Не та камера
    - Несоответствие даты съёмки историческим реалиям
- Активные
  - Криптографические методы (цифровая подпись): высокая точность, но требуется отдельный канал для её передачи
  - ЦВЗ-системы
    - Удаляемые ЦВЗ
    - Хрупкие и полухрупкие ЦВЗ, ЦВЗ с локализацией
    - Self-embedding
  - Возможно сочетание двух методов (ЦП как ЦВЗ)

1. Преимущества и недостатки каждого подхода

- Пассивные
  - + более широкие условия применения
  - - часто необходим большой объём данных
  - - как правило, ниже качество решения задачи (ошибки первого и второго рода; как следствие хуже точность локализации изменений)
  - - нельзя решить задачу восстановления искажённых областей
- Активные
  - - ограниченные возможности применения
  - - возможность подвергнуться умышленной атаке против применяемого метода аутентификации
  - В остальном одни плюсы
- Выбор подхода и конкретного метода определяется спецификой задачи и сценарием работы.
- Наилучший вариант – комбинация различных средств защиты

2. Точная аутентификация

- Требуется выявить любые изменения, произошедшие с изображением
- Недопустимы даже искажения изображения, вызванные встраиванием ЦВЗ
- Таким требованиям удовлетворяют так называемые *удаляемые ЦВЗ (reversible watermarking)*

2. Сценарий точной аутентификации

- Алиса вычисляет одностороннюю хэш-функцию изображения и встраивает её результат в это изображение в качестве ЦВЗ.
- Алиса отправляет результирующий носитель ЦВЗ Бобу.
- Боб извлекает ЦВЗ из полученного изображения.
- Боб удаляет ЦВЗ из изображения. Теперь оно должно быть эквивалентно исходному в случае отсутствия изменений при его передаче.
- Боб вычисляет одностороннюю хэш-функцию полученного изображения и сравнивает её результат с ЦВЗ.
- Изображение признаётся подлинным тогда и только тогда, когда результаты хэширования полностью совпадают.

- Вывод: эффективность решения задачи точной аутентификации сводится к эффективности построения систем удаляемых ЦВЗ

2. СВИ-19 (E_MOD/D_LC): удаляемые ЦВЗ на основе модульной арифметики

- Модификация системы СВИ-12 (E_BLIND/D_LC) (Spread Spectrum)
- $W_{mod}(n_1, n_2) \sim N(0,1)$, строится на основе ключа $\mathbf{k}$
- Изменённая формула встраивания

$$C^W(n_1, n_2) = \left(C(n_1, n_2) + \alpha \cdot W_{mod}(n_1, n_2)\right) (\mathrm{mod}\ 256),$$

- Проверка подлинности: расчёт линейной корреляции $\rho\left(\widetilde{C^W}, W_{mod}\right)$ и сравнение с порогом
- *Удаление ЦВЗ*

$$\tilde{C}(n_1, n_2) = \left(\widetilde{C^W}(n_1, n_2) - \alpha \cdot W_r(n_1, n_2)\right)\ (mod\ 256)$$

2. СВИ-19 (E_MOD/D_LC): недостатки

- Шум типа «соль-и-перец», вызванный изменением формулы встраивания:
  - ухудшает визуальное качество носителя информации
  - отрицательно сказывается на качестве детектирования
- Корреляционный детектор будет неработоспособен, если исходный контейнер C содержит значения, равномерно распределённые на отрезке от 0 до 255
- Данная система предполагает детектирование ЦВЗ, что не позволяет полноценно использовать её в предложенном выше сценарии точной аутентификации

2. СВИ-20 (Lossless-LSB): Удаляемые ЦВЗ за счёт сжатия НЗБ

- $C^W := C$
- Сжимаем без потерь одну из битовых плоскостей $C_k, k = \{3, 4\}$
- Обнуляем C_k^W
- Вместо исходной C_k^W записываем:
  - Архив C_k
  - Метка окончания архива
  - Хеш исходного изображения C

3. Избирательная аутентификация

- Основные характеристики
  - 1) Требуется выявить явные изменения, произошедшие с изображением
  - 2) Искажения изображения, вызванные встраиванием ЦВЗ, допустимы
  - 3) Дополнительное условие: задано подмножество допустимых преобразований
- Выводы и особенности:
  - (2) и (3) отличают этот случай от точной аутентификации
  - Если (3) не требуется, это частный случай, который можно назвать неточной аутентификацией. Решается при помощи хрупких ЦВЗ
  - Если (3) требуется – классическая избирательная аутентификация. Решается при помощи полухрупких ЦВЗ

3. Хрупкие ЦВЗ

- Примеры методов: НЗБ, QIM,…
- $W(n_1, n_2)$, как правило, известен на этапе извлечения, не несёт в себе информации и определяется частично на основе ключа
- Может быть вариант, когда часть ЦВЗ выступает в этой роли, а часть несёт некую полезную информацию
- Если найдётся хотя бы одна точка (n_1, n_2), в которой $W^R(n_1, n_2) \neq W(n_1, n_2)$, то считаем, что в изображение внесены изменения

3. Полухрупкие ЦВЗ

- Основа – хрупкие методы
- Далее частичная стойкость достигается за счёт преобразования данных, способа формирования ЦВЗ, формулы встраивания, значений параметров
- Для каждого искажения, к которому нужно обеспечить стойкость, формируется индивидуальное решение
- *Пример:* стойкость к повороту на угол, кратный 90°
  Решение: формирование ЦВЗ согласно условию

$$W(n_1, n_2) = W(N - n_1, n_2) = W(n_1, N - n_2) = W(N - n_1, N - n_2)$$

- *Пример:* стойкость равномерному шуму амплитудой $\Delta/2$
  Решение: QIM со шкалой квантования с шагом Δ

3. ЦВЗ, полухрупкие к JPEG

- Цель – сохранять ЦВЗ при JPEG-сжатии с высоким качеством, разрушать ЦВЗ при сжатии с низким качеством
- QF – параметр качества JPEG-сжатия, стойкость к которому требуется обеспечить. QF^* - параметр качества будущего сжатия.
- $QF^* \geq QF$ – ЦВЗ сохраняется.
- СВИ-21 (Lin & Chang) по версии книги *Watermarking Systems Engineering*
  - Блоки 8x8, ДКП. Результат – $f_{ij}(m_1, m_2)$
  - В каждый блок встраиваем 4 бита информации $b_{ij,k}$, где $k = \overline{0,3}$
  - Берём множество D из 28 коэффициентов $f_{ij}(m_1, m_2)$, расположенных ниже побочной диагонали: $D = \{(m_1, m_2): m_1 + m_2 > 7\}$,
  - Разбиваем D на 4 равных подмножества D_k по 7 коэффициентов в соответствии с ключом встраивания $\mathbf{k}$.

3. Коэффициенты при сжатии JPEG

- $p_{ij}(m_1, m_2) = round\left(\frac{f_{ij}(m_1,m_2)}{\eta_{QF}(m_1,m_2)}\right)$. Далее p_{ij} в код Хаффмана

$$\eta_{QF}(m_1, m_2) = k(QF) \cdot \eta_{50}(m_1, m_2),$$

$$k(QF) = \begin{cases} round\left(\frac{50}{QF}\right), & QF < 50, \\ 2 - 0.02 \cdot QF, & QF \geq 50. \end{cases}$$

- $\eta_{50}(m_1, m_2)$:

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

3. Основная процедура встраивания в системе СВИ-21 (Lin & Chang)

- Для встраивания каждого бита $b_{ij,k}$ повторяем шаги:
  - Осуществить деление коэффициентов $f_{ij}(m_1, m_2)$, где $(m_1, m_2) \in D_k$, на элементы матрицы η_{QF}

$$f'_{ij}(m_1, m_2) = round\left(\frac{f_{ij}(m_1, m_2)}{\eta_{QF}(m_1, m_2)}\right),$$

  - Вычислить двоичное значение

$$\beta = \underset{(m_1,m_2)\in D_k}{XOR}\left(f'_{ij}(m_1, m_2)(mod\ 2)\right).$$

  - Если $\beta \neq b_{ij,k}$, то инвертировать младший бит $f'_{ij}(m_1, m_2)$ у того коэффициента (m_1, m_2), которому соответствует наибольшее значение $\eta(m_1, m_2)$.
  - Умножить обратно значения $f'_{ij}(m_1, m_2)$ на элементы матрицы η_{QF}

$$f^W_{ij}(m_1, m_2) = \eta_{QF}(m_1, m_2) \cdot f'_{ij}(m_1, m_2).$$

3. Процедура извлечения и принцип работы СВИ-21 (Lin & Chang)

- Извлечение:

$$b_{ij,k}^{R} = \underset{(m_1,m_2)\in D_k}{XOR} \left(\widetilde{f_{ij}'}(m_1, m_2)(mod\ 2) \right)$$

- Функция переквантования в JPEG:

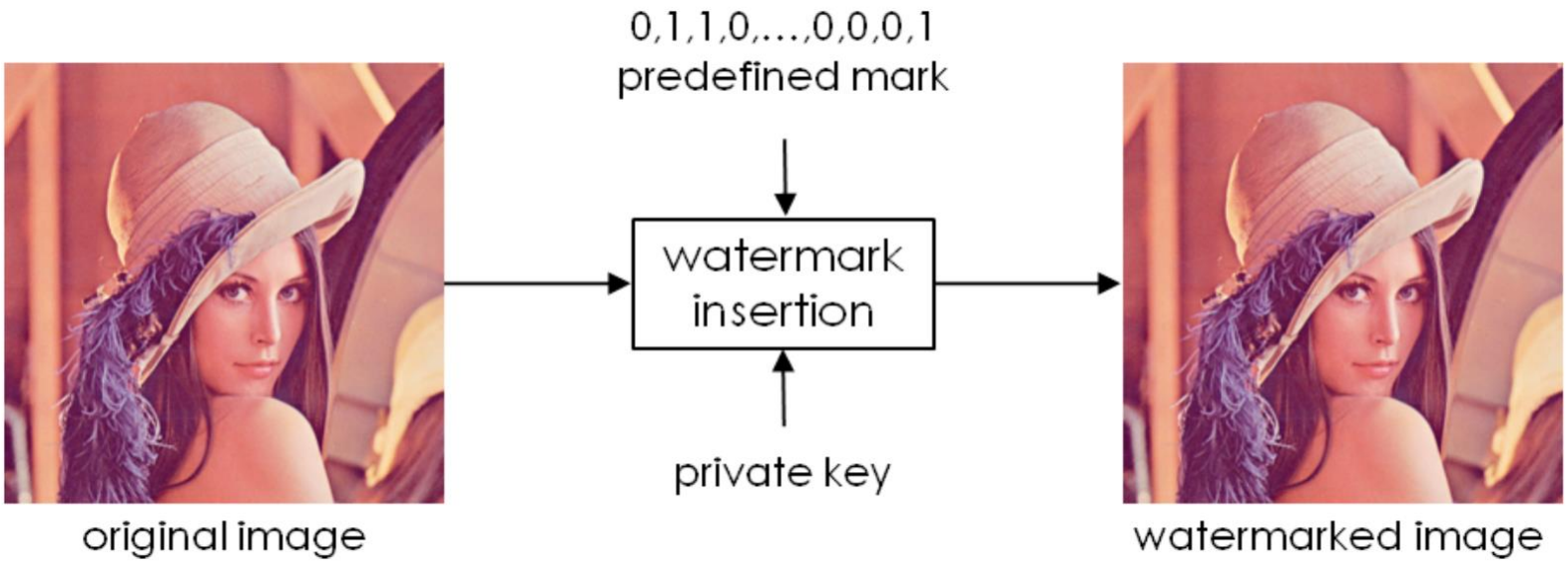
$$Q(x, \Delta) = \Delta \cdot round\left(\frac{x}{\Delta}\right).$$

- Принцип работы метода обеспечения полухрупкости:
  Пусть Δ_1 – некоторый из шагов квантования коэффициентов ДКП, определяемый значением QF, а $\Delta_2 \leq \Delta_1$ – пользовательский шаг квантования. Тогда справедливо равенство

$$Q(Q(Q(x, \Delta_1), \Delta_2), \Delta_1) = Q(x, \Delta_1).$$

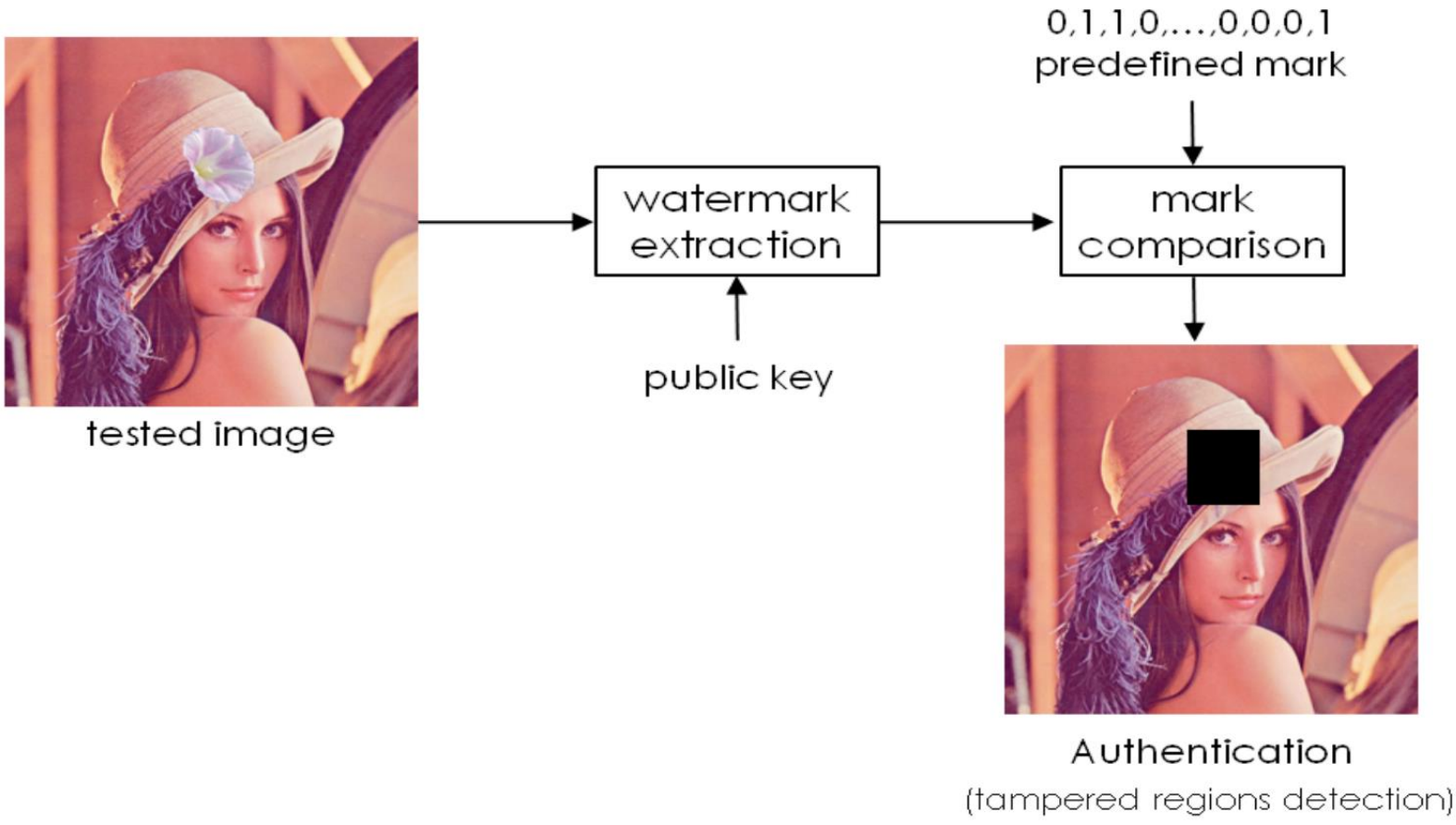
3. Локализация изменений: разные подходы

- Локализация изменений при помощи (полу)хрупких ЦВЗ-систем:
  - Защита изображения на предварительном этапе (встраивание информации)



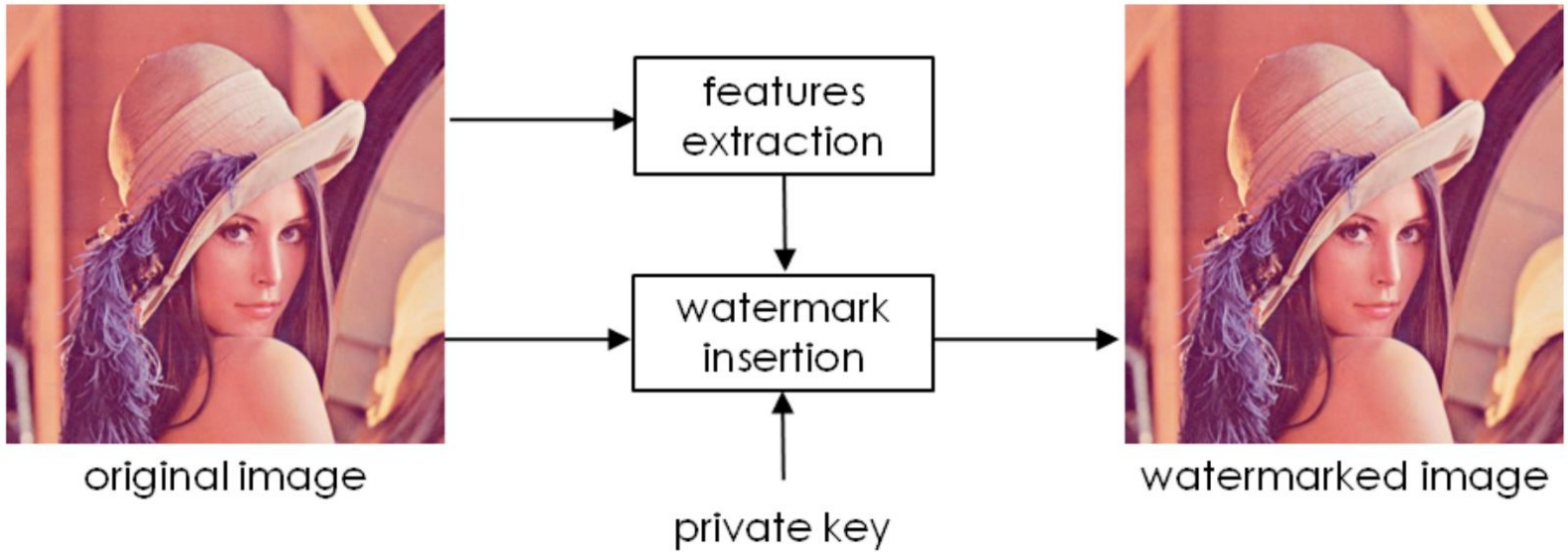
3. Локализация изменений: разные подходы

- Локализация изменений при помощи (полу)хрупких ЦВЗ-систем:
  - Проверка подлинности изображения (извлечение информации)



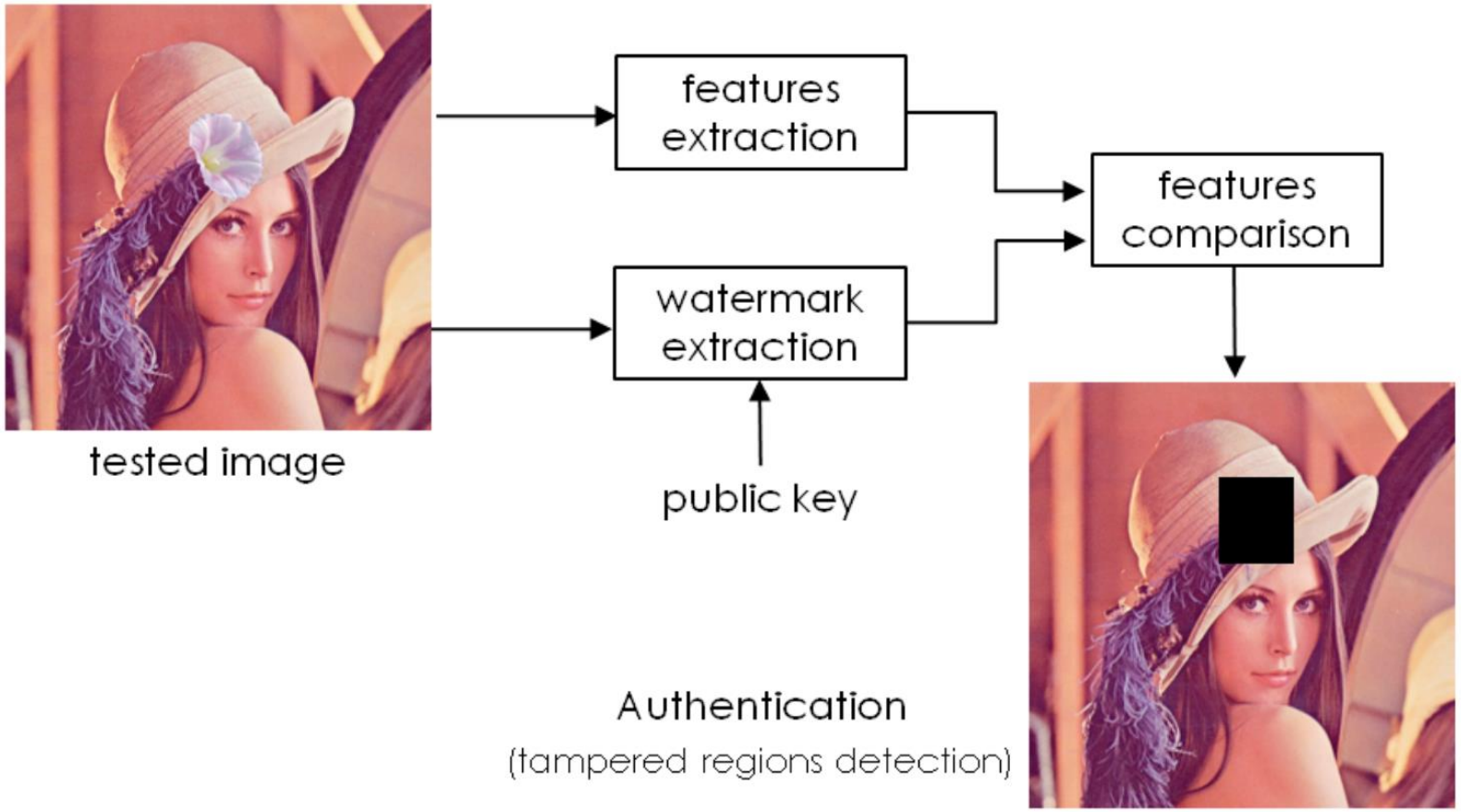
3. Локализация изменений: разные подходы

- Локализация изменений при помощи стойких ЦВЗ-систем:
  - Защита изображения на предварительном этапе (встраивание информации)



3. Локализация изменений: разные подходы

- Локализация изменений при помощи стойких ЦВЗ-систем:
  - Проверка подлинности изображения (извлечение информации)



3. Локализация изменений: разные подходы

- Чаще всего применяется метод на основе (полу)хрупких ЦВЗ-систем
- Почему?
  - Первая причина?
  - Вторая причина?
- Можно сказать, что аутентификация без локализации не особенно нужна. Стандарт де факто – сразу использовать средства локализации искажений
- Основная цель – не извлечение встроенной неизвестной информации, а поиск маски изменений

3. СВИ-23 (Yeung & Mintzer)

- На основе ключа $\mathbf{k}$ формируется отображение
$$\mu: \mathbb{N}_0 \cap [0,255] \longmapsto \{0,1\},$$
- Например, $\mu(x) = x(\mathrm{mod}\ 2)$.
- Для встраивания формируется бинарный шаблон ЦВЗ W_r размерами $M_1 \times M_2$
- В результате встраивания должно быть выполнено условие
$$\mu(C^W(n_1, n_2)) = W_r\big(n_1(mod\ M_1), n_2(mod\ M_2)\big).$$
- Проверка для каждого пикселя (n_1, n_2): если условие уже выполняется у контейнера, то $C^W(n_1, n_2) = C(n_1, n_2)$

3. СВИ-23 (Yeung & Mintzer)

- В противном случае находится такое $v \in \mathbb{N}_0 \cap [0,255]$, что
$$\mu(v) = W_r\left(n_1(mod\ M_1), n_2(mod\ M_2)\right)$$
- и v – ближайшее к $C(n_1, n_2)$ число, удовлетворяющее этому соотношению. То есть
$$v = arg \min_{x:\ \mu(x)=W\left(n_1(mod\ M_1), n_2(mod\ M_2)\right)} |x - C(n_1, n_2)|$$
- Далее, $C^W(n_1, n_2) = v$.
- Маска изменений
$$E(n_1, n_2) = \begin{cases} 1, & \mu\left(\widetilde{C^W}(n_1, n_2)\right) \neq W_r\left(n_1(mod\ M_1), n_2(mod\ M_2)\right), \\ 0 & \text{иначе.} \end{cases}$$
- По статистике половина изменённых точек будет иметь нулевое значение. Выход – постобработка. Например, морфологическое замыкание

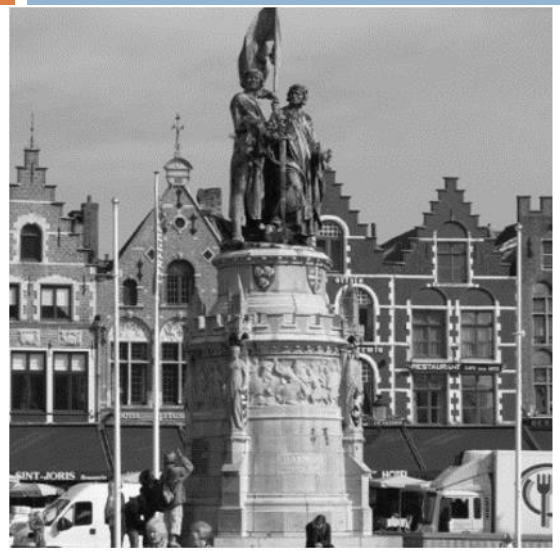
4. Восстановление искажённых фрагментов изображения

- Основная идея – self-embedding – встраивание в изображение информации о нём самом
- Реализация – либо через (полу)хрупкие системы, либо через стойкие – аналогично задаче локализации
- Два подхода:
  - Встраивание всей информации об изображении в сжатом виде
    - + полная картина по пространству
    - - низкое разрешение
  - Встраивание информации только о значимой области изображения (Region of Interest, ROI)
    - - Сначала нужно отыскать эти области
    - - Возможно, придётся закодировать смещения для восстановления ROI
    - + выше разрешение

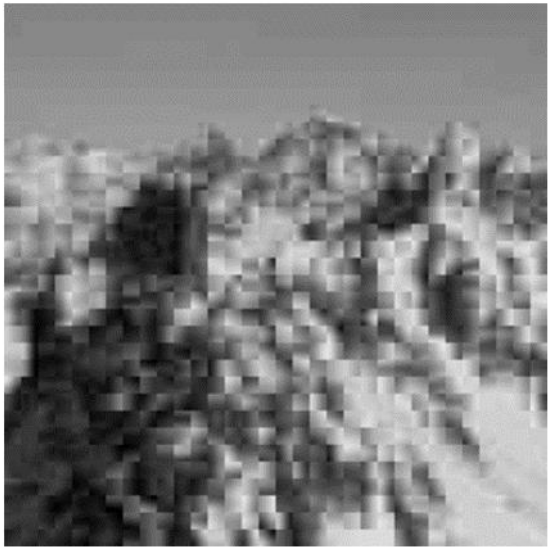
4. Средства восстановления в СВИ-21 (Lin & Chang)

- Реализует первый подход
- В каждый блок 8x8 встраивается 6 бит для восстановления в дополнение к 4 битам для локализации изменений.
- Основные этапы:
  - При кодировании контейнера в ЦВЗ изображение масштабируется в два раза по обеим осям (итого в 4 раза меньше точек)
  - Разделение на блоки 8x8, ДКП.
  - Каждому такому блоку соответствует 4 блока носителя информации. Значит суммарно $6 \cdot 4 = 24$ бита информации для его восстановления мы можем встроить.
  - Квантование с $QF = 25$, далее кодирование Хаффмана. 24 первых символа и есть искомый ЦВЗ

4. Пример восстановления в СВИ-21 (Lin & Chang)

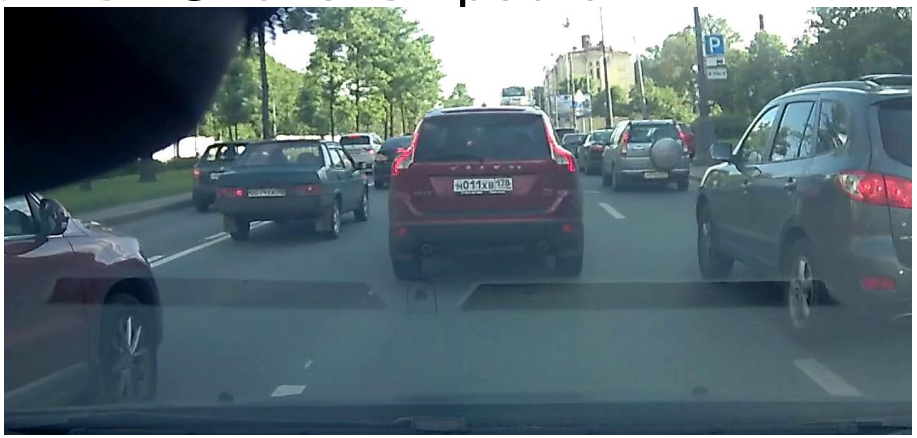







4. Средства восстановления в СВИ-21 (Lin & Chang)

- Второй подход на основе ROI ранее применялся только для медицинских снимков. Там выделить ROI очень просто
- [Egorova & Fedoseev, 2020]:
  - Снимки дорожной обстановки
  - ROI – область номерного знака
  - Стойкость к вырезанию фрагмента



- Основные этапы:
  - По умолчанию 5 проверочных бит; длина кода Хаффмана до 32 бит
  - Масштабирования не производится, $QF = 10$
  - Есть возможность выполнить избыточное встраивание информации для восстановления
  - Отдельные метки определяют высоту ROI и номер блока ROI
  - Широкое использование ключа

Сравнение двух методов на примере восстановления ROI

Исходные ROI

b) *PSNR*=19.2 *dB* c) *PSNR*=22.7 *dB* d) *PSNR*=21.3 *dB*

Восстановление контейнера из ЦВЗ, Egorova & Fedoseev

b) *PSNR*=14.8 *dB* c) *PSNR*=17.1 *dB* d) *PSNR*=16.0 *dB*

Восстановление контейнера из ЦВЗ, Lin & Chang