

## АНАЛИЗ ОСОБЕННОСТЕЙ СИНГУЛЯРНЫХ ЧИСЕЛ МАТРИЦ ЦИФРОВЫХ ИЗОБРАЖЕНИЙ ПРИ РАЗНЫХ СТЕПЕНЯХ СЖАТИЯ ДЛЯ ВЫЯВЛЕНИЯ ФОТОМОНТАЖА

Для обеспечения возможности автоматизированного анализа состояния произвольных информационных систем с учетом вероятного возникновения как естественных, так и искусственных каналов утечки информации, автоматической обработки результатов такой утечки, выбора пути восстановления системы после атаки, а также для решения вопросов синтеза систем защиты информации с привлечением современного математического аппарата актуальным и своевременным явился разработанный в [1] общий математический подход к анализу свойств и технологии функционирования информационных систем, основанный на теории возмущений и матричном анализе, позволивший уйти от их разнородности и плохой формализуемости путем несложного в вычислительном смысле анализа лишь определенной группы математических параметров, описывающих систему, - сингулярных чисел (СНЧ) и сингулярных векторов соответствующих матриц. В [2-4] было положено начало адаптации этого подхода к проблеме идентификации фальсификаций цифровых сигналов. Настоящая работа является продолжением исследований в этом направлении.

Поскольку любой цифровой сигнал может быть формализован в матричном виде [1,3], после чего отличия в его математической обработке и анализе практически пропадают, для определенности и наглядности далее будем рассматривать как исследуемый сигнал - цифровое изображение (ЦИ).

На сегодняшний день одним из наиболее распространенных форматов, используемых для ЦИ, является JPEG с потерями. Так большинство современных цифровых видеокамер сохраняет изображение именно в этом формате; хранение и передача сигналов по каналам телекоммуникаций в связи со значительным увеличением объемов информации осуществляется в сжатом состоянии; обработка изображения, в том числе и при фальсификации, как правило, проводится с использованием общедоступных на сегодняшний день программных средств, таких как, например, Adobe Photoshop, после чего ЦИ вновь сохраняется в JPEG. Таким образом, значительный интерес вызывает анализ особенностей математических параметров, определяющих ЦИ в формате JPEG.

Одним из основных шагов рассматриваемого формата является квантование коэффициентов дискретного косинусного преобразования (вейвлет-преобразования) для  $8 \times 8$ -блоков матрицы ЦИ, на которые она разбивается предварительно [5]. Данная операция является необратимой, ее результаты зависят от используемых матриц квантования (для определенности везде ниже будем рассматривать матрицы квантования Adobe Photoshop). Так при сохранении изображения в JPEG в Photoshop, который может использоваться в стандартном, оптимизированном и прогрессивном (3 варианта по количеству шагов) виде, возможны 65 вариантов такого сохранения в зависимости от значения параметра «качество», который далее обозначается  $Q$  ( $Q$  принимает значения  $0,1,2,\dots,12$ ). Именно квантование коэффициентов приводит к определенным особенностям математических параметров, характеризующих  $n \times m$ -матрицу  $F$  изображения [2].

*Цель работы* – выявление таких характеристик параметров, определяющих ЦИ, - сингулярных чисел соответствующих матриц, которые бы позволили выявить фальсификацию изображения при его хранении в формате JPEG с потерями.

Для достижения поставленной цели необходимо решить следующие задачи:

- сузить область исследований различных вариантов JPEG Photoshop с учетом специфики рассматриваемой задачи;
- выявить и обосновать особенности СНЧ матриц, отвечающих изображению, при разных вариантах JPEG, используемых для его хранения;
- проанализировать возможные значения  $Q$ , которые целесообразно использовать на практике при построении фотомонтажа;

- выявить пороговые значения для разницы параметров  $Q$ , отвечающих подобластям ЦИ, дающие возможность для определения фотомонтажа.

Доказательство подлинности цифровых сигналов является важной задачей во многих областях человеческой деятельности: медицине, прессе, научных публикациях, судебных разбирательствах и т.д. Конкретизируем интересующую нас область возможных фальсификаций судебными разбирательствами. В силу специфики задачи качество фотографии, предоставляемой, например, как вещественное доказательство, должно быть приемлемым. Это означает, что на фотографии недопустимо наличие артефактов, которые могут возникнуть после сжатия. Кроме того, плохое качество фотографии само по себе вызывает сомнение в подлинности, что не позволит рассматривать ее как вещественное доказательство.

Путем субъективного ранжирования для ЦИ, сохраняемых в формате JPEG (стандартный) в Photoshop, было установлено, что при  $Q \in \{0,1,\dots,7\}$  артефакты в большей или меньшей степени заметны на изображении, кроме того, как показывает практика, хорошее визуальное восприятие достигается уже при  $Q = 9, Q = 10$ , а качественная картина для анализируемых ниже математических параметров для  $Q = 11, Q = 12$  не будет отличаться от характеристик при  $Q = 10$ . Поэтому далее исследования проводятся для  $Q \in \{8,9,10\}$ .

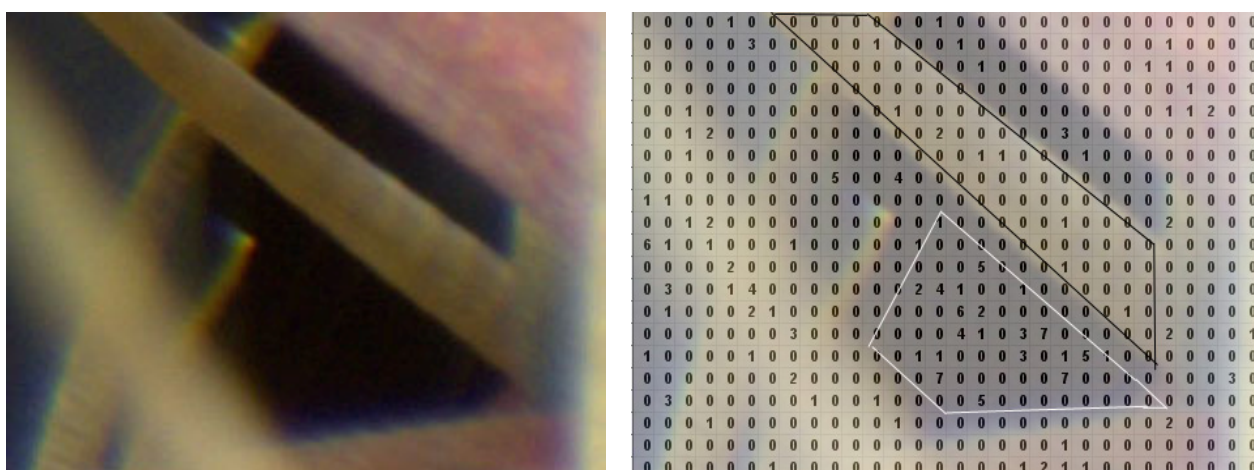
Для достижения поставленной цели матрица ЦИ  $F$  разбивается стандартным образом на блоки  $8 \times 8$ , фигурирующие при реализации JPEG, для каждого из которых определяется множество СНЧ (сингулярный спектр) [1]. Изображению ставится в соответствие матрица нулевых СНЧ блоков (МНСЧБ) размерности  $[n/8] \times [m/8]$ , каждый элемент которой равен количеству нулей в сингулярном спектре соответствующего блока [2]. Для  $Q = 9, Q = 10$  МНСЧБ содержит практически только нули (возможны единичные «вкрапления» отличных от нуля значений, равных 1,2, что подробно объясняется в [2]). Для  $Q = 8$  картина будет отличаться: для фоновых изображений возникает разброс значений элементов от 0 до 7, для контурных – от 0 до 4. Причем этот разброс возникает на МНСЧБ практически равномерно (рис.1(б) – МНСЧБ для наглядности наложена на изображение; белым цветом выделена фоновая область, черным - контурная). Этот результат не случаен. Матрицы квантования в Photoshop для  $Q = 9, Q = 10$  качественно отличаются от соответствующей матрицы для  $Q = 8$  (рис.2): коэффициенты, находящиеся над и на побочной диагонали для  $Q = 8$  приведут к обнулению не только коэффициентов дискретного косинусного преобразования, отвечающих высоким, но и частично средним частотам. Причем, если для изображений (блоков изображений), содержащих контуры, коэффициенты при средних и высоких частотах по значениям явно отличаются друг от друга, что позволит оставить ненулевыми некоторые (все) средние частоты даже при возрастании соответствующих коэффициентов квантования (рис.2(в)), то для фоновых изображений коэффициенты при средних и высоких частотах малы и сравнимы между собой (рис.3(б)): действительно, определяющей здесь является низкочастотная составляющая сигнала, матрица яркости блока фонового изображения без потерь имеет очень малый разброс значений, ее столбцы (строки) близки к линейной зависимости или являются линейно зависимыми (рис.3(а)), что приведет к обнулению большего количества коэффициентов при высоких и средних частотах, чем в контурном изображении. В соответствии с [6] такое обнуление приведет к обнулению сингулярных чисел (к росту соответствующих значений элементов МНСЧБ), причем для фоновых областей, в силу вышесказанного, таких нулевых значений будет не меньше, чем для контурных (для фоновых областей возможна ситуация, когда весь сингулярный спектр станет нулевым).

Заметим, что, как правило, при создании фотомонтажа итоговое ЦИ подвергается дополнительной обработке средствами Photoshop (или какой-либо аналогичной программы), после чего сохраняется в JPEG с использованием одной из конечного числа предусмотренных матриц квантования. Большинство из современных широко распространенных фотокамер, как показывает практика, имеют матрицы квантования, отличные от имеющихся в Photoshop. Таким образом, определение матрицы квантования

анализируемого ЦИ, которое легко сделать с использованием существующих программных средств, явно укажет на обработку этого изображения, а значит косвенно – на возможную фальсификацию, если матрица совпадет с какой-либо из Photoshop. Однако такой способ, во-первых, детектирует вероятную фальсификацию лишь косвенно, а во-вторых, принципиально не может локализовать ее область.

Пусть часть фотоснимка (формат JPEG), который будем называть основным изображением (ОИ), заменяется частью ЦИ (формат JPEG), далее называемой вклейкой, или замещающей областью (ЗО), причем для большей наглядности получаемых ниже выводов никакая последующая обработка изображения не производится. Такое фальсифицированное ЦИ (фотомонтаж) сохраняется в формате JPEG. Поскольку для рассматриваемой нами области использования фотомонтажа, представляемые итоговые ЦИ должны быть хорошего качества, то при различных значениях  $Q$  для ОИ и ЗО, сохранение фотомонтажа целесообразно проводить с бóльшим из двух имеющихся  $Q$ , что и делается ниже.

Значительное отличие в значениях  $Q$  ОИ и ЗО легко обнаруживается при помощи МНСЧБ фотомонтажа. Иллюстрация представлена на рис.4: для подматрицы МНСЧБ, отвечающей ЗО, явно виден разброс значений от 0 до 7, причем этот разброс не носит систематический характер, а локализован, что совершенно не свойственно МНСЧБ для формата JPEG с  $Q = 10$  и говорит о наличии вклейки.



а

б

Рис.1. ЦИ в формате JPEG (стандартный) при  $Q = 8$  (а), соответствующая МНСЧБ (б)

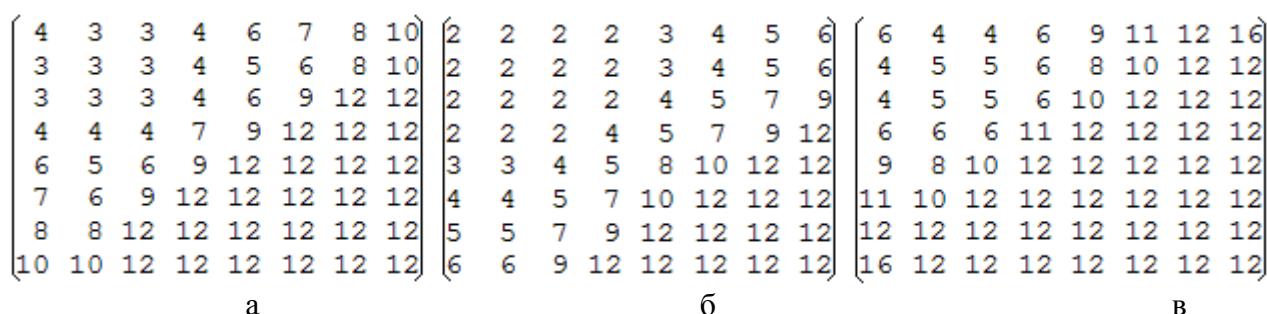


Рис.2. Матрицы квантования, соответствующие показателю  $Q = 9$  (а),  $Q = 10$  (б),  $Q = 8$  (в)

107	108	107	106	99	101	102	107
109	106	108	107	103	102	103	110
107	106	110	110	106	107	107	120
106	107	108	108	108	108	108	114
105	108	109	109	108	106	107	110
105	108	109	110	108	108	109	109
108	109	109	109	108	109	110	109
107	107	108	108	107	110	110	109

а

861.2500	-3.0946	4.3024	-10.0126	2.7500	-0.3334	2.5475	-2.4361
-7.5487	3.8656	5.6269	-5.8941	4.3543	-0.8037	2.6387	-1.3123
-5.5993	5.3250	2.2071	1.8366	-1.9038	2.5010	-0.0000	-0.9627
-5.1741	5.4716	-2.3326	0.4579	-2.6087	0.6486	-2.9587	-0.1493
-3.2500	0.7709	-0.3827	1.2010	-1.7500	-0.5247	-0.9239	-0.2509
2.1159	-1.3862	1.0089	0.2521	-0.8046	-0.4756	-0.1278	-0.7962
1.5075	-2.9231	0.0000	-2.0226	-0.0232	-0.0381	0.7929	-1.3047
1.4207	-0.2088	0.6650	-1.9285	0.5803	-0.3585	0.4510	-0.3479

б

Рис.3. Матрица яркости фонового блока изображения POUT.TIF (а); коэффициенты дискретного косинусного преобразования (б)

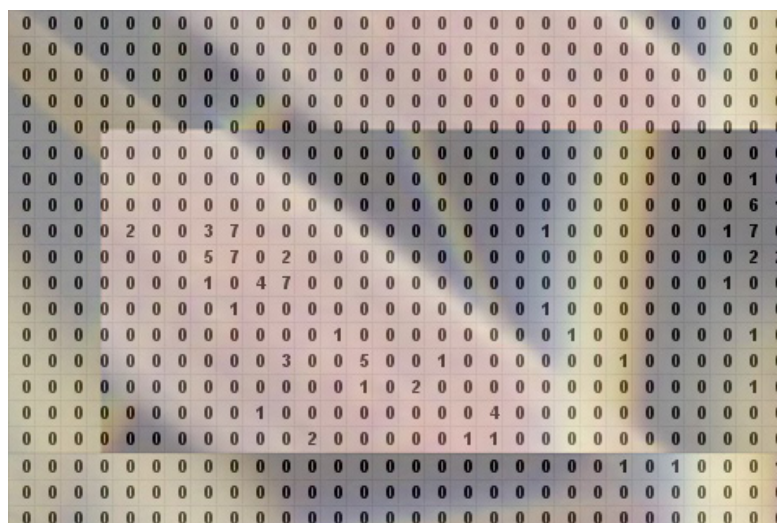


Рис.4. МНСЧБ, наложенная на соответствующее изображение фотомонтажа, сохраненного с  $Q = 10$ , в котором для ОИ  $Q = 10$ , для ЗО -  $Q = 5$

Однако, если монтаж делается неделитантом, то значение  $Q$  для вклейки и основного изображения будут мало отличаться друг от друга либо будут равными (конечно, если это возможно). В противном случае для обеспечения надежности восприятия полученного фотомонтажа необходимы будут значительно бóльшие усилия по обработке в Photoshop полученного изображения, чем для мало отличающихся значений  $Q$  ОИ и ЗО, что может привести к «отрицательным» последствиям для автора фотомонтажа в силу того, что результаты применения некоторых часто используемых инструментов Photoshop, каковым является, например, «размытие» изображения [7], «вылавливаемые» при помощи общего подхода к анализу состояния информационной системы [1], сами по себе являются дополнительными указателями на фальсификацию. В связи с этим необходимо установить минимальное отличие значений  $Q$  ОИ и ЗО - порог  $P_Q$ , при котором МНСЧБ дает возможность выявить фальсификацию за счет различий в значениях и локализации ее элементов.

Путем проведения вычислительного эксперимента со 100 различными ЦИ было установлено, что значение  $P_Q = 2$  (типичный случай для фонового изображения приведен на рис.5, аналогично будет выглядеть МНСЧБ для изображения, содержащего контуры).

Полученный результат имеет четкое объяснение. При разнице значений  $Q$  2 и более соответствующие МНСЧБ, как уже было указано выше, качественно отличаются друг от друга. Эти отличия определяются количеством нулевых СНЧ в блоках, возникших после квантования и восстановления ЦИ. Заметим, что если сохраненное в JPEG с определенным показателем  $Q_1$  изображение пересохранить с бóльшим показателем  $Q_2$ , то МНСЧБ качественно не изменится. Действительно, если коэффициенты дискретного косинусного преобразования обнулились (остались ненулевыми) при  $Q_1$ , то они останутся нулевыми (ненулевыми) при повторном квантовании с *меньшими* коэффициентами, отвечающими  $Q_2$ , что, учитывая соответствие между коэффициентами частотного спектра и сингулярными тройками матриц [6], повлечет за собой аналогичную картину для сингулярных спектров блоков. Таким образом, если фотомонтаж формируется из частей, показатели  $Q$  для которых отличаются не менее, чем на  $P_Q$ , после чего сохраняется в лучшем из двух используемых качеств, то качественное различие в областях МНСЧБ, отвечающих ОИ и ЗО, сохранятся, что и даст возможность определить присутствие фотомонтажа. Для случаев, когда разница между значениями  $Q$  ОИ и ЗО равна 1 или 0 МНСЧБ не всегда дает ответ о наличии или отсутствии фальсификации, здесь нужны более чувствительные математические инструменты.

Выявленные характерные особенности МНСЧБ для различных вариантов формата JPEG могут использоваться не только для идентификации фальсификации ЦИ. На рис.6 в качестве примера представлена часть (в силу ограниченности объема статьи) фальсифицированного кадра видеопоследовательности известного художественного фильма «Москва слезам не верит». Фальсификация становится очевидной при анализе МНСЧБ, соответствующей матрице рассматриваемой части кадра, и определяется разбросом значений от 0 до 7, локализованным в области, отвечающей изображению главной героини.



Рис.5. МНСЧБ, наложенная на соответствующее изображение фотомонтажа, сохраненного с  $Q = 10$ , в котором для ОИ  $Q = 10$ , для ЗО -  $Q = 8$





Рис.6.МНСЧБ, наложенная на фальсифицированный кадр видеопоследовательности

Таким образом, в результате проделанной работы получены следующие результаты:

- показано, что использование общего подхода к анализу состояния и технологии функционирования информационных систем, основанного на теории возмущений и матричном анализе, является перспективным для решения вопросов идентификации фальсификаций цифровых сигналов;
- установлены характерные особенности МНСЧБ, отвечающей ЦИ, хранимому в формате JPEG с потерями;
- установлено, что при значительных отличиях  $Q$  ОИ и ЗО МНСЧБ итогового ЦИ явно указывает на присутствие фальсифицированной области;
- качественные отличия подматриц МНСЧБ позволяют определить наличие фальсификации в случае, когда разность между  $Q$  ОИ и ЗО не менее двух;
- для идентификации фальсификаций ЦИ в случае, когда разность  $Q$  ОИ и ЗО менее двух, требуются дополнительные исследования, проведением которых авторы заняты в настоящий момент.

#### Список литературы

1. Кобозева А.А., Хорошко В.А. Анализ информационной безопасности. – К.: Изд. ГУИКТ, 2009. –251 с.
2. Кобозева А.А. Использование теории возмущений для обнаружения фальсификации цифрового изображения // Проблеми інформатизації та управління. Матеріали міжнародної науково-технічної конференції «Комп'ютерні системи та мережні технології». Збірник наукових праць. — 2008. — №1(23). — С.16—22.
3. Кобозева А.А., Рыбальский О.В., Трифонова Е.А. Матричный анализ – основа общего подхода к обнаружению фальсификации цифрового сигнала // Вісник Східноукр-го нац-го ун-ту ім. В.Даля. — 2008. — № 8(126), ч.1. — С.62—72.
4. Кобозева А.А., Зоріло В.В. Метод виявлення фальсифікації цифрового зображення в умовах збурних дій // Збірник наукових праць Військового інституту Київського національного університету ім.Т.Шевченка. - 2009. - Вип.№20. - С.147-154
5. Гонсалес Р., Вудс Р. Цифровая обработка изображений.- М.: Техносфера, 2005.- 1072с.
6. Кобозева А.А. Связь свойств стеганографического алгоритма и используемой им области контейнера для погружения секретной информации // Искусственный интеллект. — 2007. — №4. — С.531—538.
7. Кобозева А.А. Особенности сингулярного спектра матрицы изображения, идентифицирующие его фальсификацию в условиях обработки некоторыми программными средствами // Матеріали V Міжнародної науково-практичної конференції «Спеціальна техніка у правоохоронній діяльності» (Україна, Київ, жовтень 2009 року).

Рецензент: д.т.н., Шелест М.Є.  
Надійшла 03.02.2010 р.