

# Метод обнаружения удаления кадров на видео со стационарной камеры наблюдения

А.Ю. Баврина

Институт систем обработки  
изображений - филиал ФНИЦ  
«Кристаллография и фотоника»  
РАН

Самарский национальный  
исследовательский университет  
им. академика С.П. Королева  
Самара, Россия  
bavrina@mail.ru

**Аннотация**—Предлагается метод пассивной защиты видео с камеры наблюдения от атаки удаления фрагмента видео. Метод основан на формировании локальных признаков по отсчетам двух последовательных кадров с последующей классификацией многослойной нейронной сетью. Постобработка и вычисление статистики по получаемому в результате классификации изображению позволяет принять решение о том, является ли данная пара кадров последовательной или между ними были удаленные кадры. Эксперименты показали эффективность обнаружения факта удаления кадров даже из стационарных сцен, когда такое удаление визуально незаметно.

**Ключевые слова**— обнаружение подделки видео, локальные признаки, камеры наблюдения, нейронные сети

## 1. ВВЕДЕНИЕ

В последние годы распространение программного обеспечения для визуального редактирования наряду с быстрым развитием методов глубокого обучения привело к ситуации, когда аутентификация видеоданных стала сложной задачей [1, 2]. Видеоданные являются веским доказательством в суде, поэтому их экспертиза стала более актуальной, чем когда-либо.

Атаки на видео в современных публикациях разделяют по области атаки: пространственная, временная и пространственно-временная. Наиболее распространенной является временная атака, при которой изменения могут заключаться в удалении, добавлении, дублировании или перетасовке группы кадров. Среди пассивных методов обнаружения изменений выделяют методы, основанные на использовании схемы компрессии видео, на вычислении признаков последовательных кадров и на глубоком обучении [3].

Большинство существующих методов обнаруживают временную атаку путем выявления изменений корреляции между соседними кадрами видео. Для этого может анализироваться несогласованность корреляции кадров, закодированная локальными двоичными шаблонами [4], показатель структурного сходства [5], моменты Цернике [6], интенсивность поля скоростей [7], постоянство оптического потока [8], и другие. С увеличением вычислительных мощностей в последние годы все активнее стали применяться модели машинного обучения и, в частности, глубокие нейронные сети [9]. Однако, непрозрачность работы нейронных сетей, неоднозначность их построения и сложность настройки

не позволяет пока этой группе методов занять лидирующие позиции [1].

Предложенные в литературе методы показывают довольно хорошие показатели для динамических сцен при постоянных настройках камеры и равномерном освещении. Однако, в обзорах [3, 10] констатируется факт, что на данный момент не существует универсального инструмента для обнаружения фальсификации видео и существует потребность в эффективных и действенных методах, способных решать открытые вопросы. К таким открытым вопросам относится удаление кадров из статичного видео, вариант решения которого предлагается в настоящей работе.

## 2. МЕТОД ОБНАРУЖЕНИЯ ФАКТА УДАЛЕНИЯ КАДРОВ

Предлагаемый метод состоит из следующих этапов. Для каждого отсчета текущего кадра исследуемого видео вычисляется набор признаков, характеризующий степень изменения локальных характеристик этого отсчета по сравнению с предыдущим кадром. Далее выполняется классификация этого вектора с использованием предварительно обученной многослойной нейронной сети. В результате формируется изображение, отсчеты которого содержат степень уверенности в факте, что данная пара кадров не является последовательной. Постобработка заключается в ранговой фильтрации для удаления малых выбросов. Окончательно решение о факте вырезки в текущем положении в видео принимается на основании вычисления статистики по результатам ранговой фильтрации.

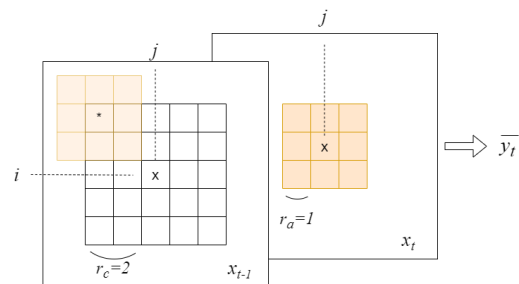


Рис. 1. Схема вычисления вектора признаков для пары последовательных кадров

На (Рис. 1) показана схема вычисления компонент вектора признаков для отсчета  $(i, j)$  кадров  $x_t$  и  $x_{t-1}$ .

$$\bar{y}_t(i, j) = \{f(x_t(i-k, j-l), x_{t-1}(i-m-k, j-n-l))\},$$

где  $f(x_i, x_{i-1}) = \sum_{c=1..3} (x_i^c - x_{i-1}^c)^2$  – евклидово расстояние между цветовыми компонентами отсчетов,  $(m, n)$  – смещение на кадре  $x_{i-1}$  от точки  $(i, j)$  положения локальной области кадра  $x_i$ ,  $(k, l)$  – индексы внутри локальной области кадра  $x_i$ ,  $m, n \in [-r_c, r_c]$ ,  $k, l \in [-r_a, r_a]$ ,  $r_c$  и  $r_a$  – параметры. То есть для каждого отсчета  $(i, j)$  формируется вектор признаков длиной  $n_f = (2r_c + 1)^2 \cdot (2r_a + 1)^2$ . Такой способ формирования признаков позволяет, во первых, учитывать "мелкие" структуры, характерные для конкретной камеры и условий наблюдения (в отличие от интегральных признаков), и, во вторых, закладывает возможность учёта движения объектов в кадре.

Для классификации вектора признаков была выбрана нейронная сеть с двумя внутренними слоями (размером  $2n_f + 1$  и  $n_f$ ) и активационной функцией *softmax* для выходного слоя.

Для экспериментов использовались 2 группы видеофайлов, параметры которых приведены в таблице 1, а примеры кадров даны на (Рис. 2).

Таблица I. ПАРАМЕТРЫ ТЕСТОВЫХ ВИДЕО

<b>Road</b>	1080x1920, fps=100, кодек H264 - MPEG AVC (part 10), GOP=24 (I,P,B-кадры). Присутствуют малые смещения камеры и изменение освещённости.
<b>SU_corridor</b>	1080x1920, fps=25, кодек H264 - MPEG 4 AVC (part 10), GOP=50 (I,P-кадры). Присутствуют сцены при естественном и искусственном освещении.

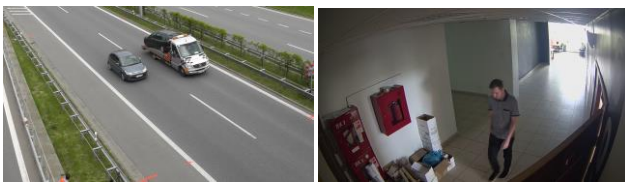


Рис. 2. Кадры тестовых видео

Пары кадров для обучения, отобранные вручную, содержали как сцены с движением объектов, так и стационарные сцены (для обоих классов). Таблица 2 показывает точность классификации (последовательные: да-нет), получаемую при обучении нейронной сети для различных значений параметров (также приведено количество признаков).

Таблица II. ВЕРОЯТНОСТЬ ПРАВИЛЬНОЙ КЛАССИФИКАЦИИ ДЛЯ РАЗЛИЧНЫХ ЗНАЧЕНИЙ ПАРАМЕТРОВ

	$r_c=1$ $r_a=0$ $n_f=9$	$r_c=1$ $r_a=1$ $n_f=81$	$r_c=2$ $r_a=0$ $n_f=25$	$r_c=2$ $r_a=1$ $n_f=225$	$r_c=3$ $r_a=1$ $n_f=441$
<b>Road</b>	0,72	0,87	0,72	<b>0,88</b>	0,87
<b>SU</b>	0,85	0,89	0,86	<b>0,89</b>	0,89

Затем для каждой пары предположительно последовательных изображений тестовой выборки было сформировано изображение, содержащее степень уверенности, что данная пара не является последовательной. После ранговой фильтрации по этому изображению вычислялась статистика (среднее значение), на основании которой можно вынести

решение о наличии вырезки фрагмента в данной точке. Сводные значения этих статистик приведены на (Рис. 3.)

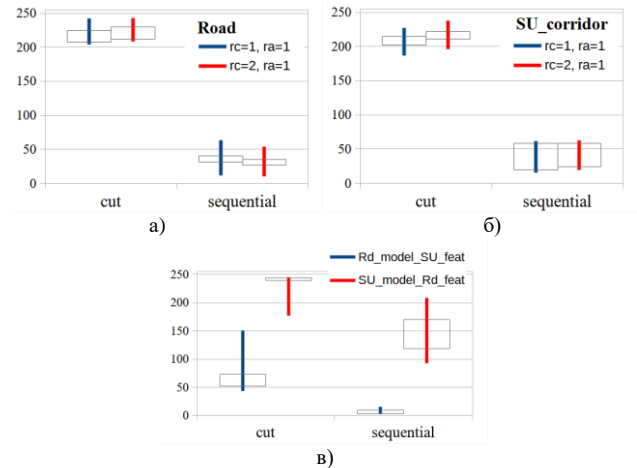


Рис. 3. Диапазоны значений статистики: а) для видео Road; б) для видео SU\_corridor; в) сеть, обученная на Road, применялась к SU\_corridor и наоборот

### 3. ЗАКЛЮЧЕНИЕ

В статье предложен метод обнаружения нарушения последовательности кадров на видео, остающийся работоспособным для статических фрагментов видео.

#### БЛАГОДАРНОСТИ

Работа выполнена при поддержке РФФИ (грант 19-29-09045 мк).

#### ЛИТЕРАТУРА

- [1] Javed, A.R. A comprehensive survey on digital video forensics: Taxonomy, challenges, and future directions / A.R. Javed, Z. Jalil, W. Zehra, T.R. Gadekallu, D.Y. Suh, M.J. Piran // Eng. Appl. Artif. Intell. – 2021. – Vol. 104456.
- [2] Денисова, А.Ю. Обнаружение атак на биометрическое предъявление на системы аутентификации лиц при помощи специальных устройств съёмки / А.Ю. Денисова, В.А. Федосеев // Компьютерная оптика. – 2022. – Том 46, №4. – С. 612-620.
- [3] Sitara, K. Digital video tampering detection: An overview of passive techniques / K. Sitara, B.M. Mehtre // Digital Investigation. – 2016. – Vol. 18. – P. 8-22.
- [4] Zhang, Z. Efficient video frame insertion and deletion detection based on inconsistency of correlations between local binary pattern coded frames / Zhang Z., Hou J., Ma Q., Li Z. // Secur Commun Netw. – 2015. – Vol. 8(2). – P. 311-320.
- [5] Li, Z. Video inter-frame forgery identification based on the consistency of quotient of MSSIM / Li Z., Zhang Z., Guo S., Wang J. // Secur Commun Netw. – 2016. – Vol. 9(17). – P. 4548-4556.
- [6] Liu, Y. Exposing video inter-frame forgery by zernike opponent chromaticity moments and coarseness analysis / Y. Liu, T. Huang // Multimed Syst. – 2015. – P. 1-16.
- [7] Wu, Y. Exposing video inter-frame forgery based on velocity field consistency / Y. Wu, X. Jiang, T. Sun, W. Wang // Acoustics, speech and signal processing (ICASSP). – 2014. – P. 2674-2678.
- [8] Chao, J. A Novel Video Inter-frame Forgery Model Detection Scheme Based on Optical Flow Consistency / J. Chao, X. Jiang, T. Sun // Lecture Notes in Computer Science (IWDW). – 2012. – Vol. 7809.
- [9] Kaur, H. Deep convolutional neural network for graphics forgery detection in video / H. Kaur, N. Jindal // Wirel. Pers. Commun. – 2020. – P. 1-19.
- [10] Bakas, J. Detection and localization of inter-frame forgeries in videos based on macroblock variation and motion vector analysis / J. Bakas, R. Naskar, S. Bakshi // Comput. Electr. Eng. – 2021. – Vol. 106929.