

# Улучшенный алгоритм встраивания ЦВЗ в видео с возможностью извлечения с помощью камеры мобильного устройства

О.О. Евсютин  
Национальный исследовательский университет «Высшая школа экономики»  
Москва, Россия  
oevsyutin@hse.ru

Д.И. Подболотов  
Национальный исследовательский университет «Высшая школа экономики»  
Москва, Россия  
dipodbolotov@edu.hse.ru

А.С. Мельман  
Национальный исследовательский университет «Высшая школа экономики»  
Москва, Россия  
amelman@hse.ru

А.Г. Станкевич  
Национальный исследовательский университет «Высшая школа экономики»  
Москва, Россия  
agstankevich@edu.hse.ru

**Аннотация**—Использование цифровых водяных знаков (ЦВЗ), устойчивых к съёмке с экрана, является перспективным способом незаметного хранения информации в видеопотоке для последующего её извлечения пользователем с помощью камеры смартфона. Однако разработка алгоритмов, реализующих данный сценарий, связана с проблемой соблюдения баланса между незаметностью встраивания и робастностью. В данном исследовании предлагается подход к повышению эффективности алгоритма встраивания ЦВЗ в видеоданные на основе прямоугольных паттернов, обеспечивающего устойчивость к съёмке с экрана.

**Ключевые слова**— цифровой водяной знак, видео, нейронная сеть, съёмка с экрана

## 1. ВВЕДЕНИЕ

Увеличение числа устройств, подключённых к сети, приводит к постоянному росту объёма генерируемых цифровых данных, в том числе мультимедиа-данных. При этом активное развитие иммерсивных приложений, объединяющих цифровой и физический миры, приводит к возникновению новых путей взаимодействия человека с цифровым контентом. Необходимость перехода от объектов цифрового мира к объектам реального мира и наоборот появляется в разных областях, таких как промышленность, медицина, образование, торговля. Примерами могут служить технологии дополненной реальности для моделирования окружающей среды, извлечение информации из видеопотоков и графических объектов в реальном времени.

Перспективным способом организации взаимосвязи физического и цифрового мира являются цифровые водяные знаки (ЦВЗ). ЦВЗ – это некая дополнительная информация, которая встраивается в объект-контейнер [1, 2]. Контейнерами для внедрения ЦВЗ могут быть разные цифровые объекты, например, изображения, видеопоследовательности. В отличие от QR-кодов, получивших широкое распространение в последние годы, ЦВЗ в большинстве случаев не вносят видимых изменений в контент и хранятся незаметным для пользователя образом. Основными показателями эффективности алгоритмов ЦВЗ являются их незаметность и робастность. Эти показатели взаимно обратны, поэтому основной задачей разработки новых алгоритмов является соблюдение баланса между ними.

ЦВЗ могут использоваться как для защиты прав на цифровой объект, так и для хранения дополнительной информации, например, ссылок на веб-страницы товаров

в рекламе. В том случае, если ЦВЗ используется для незаметного хранения дополнительной информации в видеопотоке, ожидаемый сценарий получения такой информации выглядит следующим образом:

- конечный пользователь наводит камеру смартфона на экран устройства, демонстрирующего видео, содержащее ЦВЗ;
- происходит захват некоторого количества кадров видео;
- программное обеспечение анализирует захваченные кадры, извлекает водяной знак и предоставляет пользователю полученную информацию.

Данный сценарий предъявляет дополнительные требования к робастности ЦВЗ, поскольку водяной знак должен отличаться устойчивостью к атаке съёмки с экрана. В большинстве случаев такая съёмка искажает цветопередачу захваченного кадра, сопровождается поворотом изображения на произвольный угол и изменением размера кадра на разных расстояниях от камеры смартфона до устройства демонстрации видео.

В работе [3] был представлен алгоритм встраивания ЦВЗ в видеоданные на основе прямоугольных паттернов, обеспечивающий повышенную устойчивость к съёмке с экрана. В настоящем исследовании предлагается модификация данного алгоритма. Предлагаемая модификация отличается следующими особенностями:

- уменьшение уровня искажений, вносимых на этапе встраивания ЦВЗ, благодаря выбору цветовой компоненты для встраивания и применения размывки к водяному знаку;
- повышение точности определения границ экрана, демонстрирующего видео, содержащее ЦВЗ, на снимке с камеры мобильного устройства за счёт применения нейронной сети.

## 2. ПОЛУЧЕННЫЕ РЕЗУЛЬТАТЫ

Авторы алгоритма [3] встраивают ЦВЗ в пространственную область кадра посредством наложения паттерна, состоящего из прямоугольников, значения яркостей которых кодируют передаваемую информацию. Каждый прямоугольник кодирует один бит данных. Сила встраивания определяется параметром  $S$ , увеличение которого повышает робастность и уменьшает незаметность встраивания. Процедуре

извлечения водяного знака предшествует нахождение области, содержащей ЦВЗ, на снимке с камеры смартфона. Определение такой области основано на применении паттернов, последовательности морфологических операций и вращения.

В рамках настоящей работы выполнено исследование алгоритма [3] и предложен ряд модификаций, устраняющих обнаруженные недостатки.

#### *А. Подход к уменьшению уровня искажений кадра при встраивании ЦВЗ*

Авторы алгоритма [3] используют для встраивания ЦВЗ синюю компоненту модели RGB. Однако исследование показало, что эффективность такого подхода существенным образом зависит от содержания маркируемого видео. Преобладание какого-либо одного цвета на кадрах видео приводит к появлению артефактов встраивания, заметных невооружённым глазом, при изменении одной цветовой компоненты. Для устранения данного недостатка предложен подход к выбору подходящей компоненты RGB-модели для встраивания в зависимости от того, какой цвет является преобладающим в оригинальном видеоряде.

Использование прямоугольных паттернов для встраивания данных может приводить к возникновению характерных искажений кадра, имеющих вид прямоугольных участков с изменённым оттенком цвета. Для борьбы с данным видом искажений в настоящей работе предлагается применять к ЦВЗ размытие по Гауссу. Результаты экспериментов показывают, что при средних значениях параметра силы встраивания ошибки при извлечении данных не возникают, до достижения значения радиуса размытия, равного 30.

Реализация предложенного подхода позволила существенно уменьшить уровень искажений, вносимых при встраивании, и повысить незаметность ЦВЗ. При размере ЦВЗ равном 16 битов и значении силы встраивания  $S = 9$  прирост значений метрик PSNR и SSIM, характеризующих незаметность встраивания, составил 17,18% и 7,90% соответственно.

#### *Б. Подход к повышению точности определения границ экрана на снимке*

Для повышения точности обнаружения границ экрана, демонстрирующего видео с ЦВЗ, была применена нейронная сеть архитектуры U-net [4]. Для целей данного исследования классическая архитектура U-net была модифицирована следующим образом:

- разрешение входного изображения установлено равным  $240 \times 160$  пикселей;
- число признаков на каждом шаге уменьшено с 128, 256, 512, 1024 единиц до 16, 32, 48, 64 соответственно.

Указанное разрешение является стандартным разрешением для устройств на ОС Android. Таким образом, после получения изображения с камеры не требуется изменение его размера перед передачей на вход нейронной сети. Такая архитектура позволяет

обрабатывать в среднем 43 кадра в минуту (до 175) при запуске на мобильном устройстве на 8 потоках.

Для обучения и тестирования нейронной сети был сформирован новый датасет. Общий размер датасета составил 9205 изображений, из которых 6903 использовались для обучения, 1381 для настройки и 921 для тестирования (15:3:2).

Исследование влияния конструкции функции потерь на качество решения задачи показало большую эффективность метрики MSE в качестве функции потерь, по сравнению с метриками SSIM и MAE.

Результаты экспериментов показали, что новый подход к обнаружению границ экрана на снимке с камеры смартфона позволяет повысить надёжность передачи данных. Значение метрики BER, характеризующей долю ошибок при извлечении, достигает 0,03. При этом сохраняется приемлемая незаметность встраивания данных: среднее значение метрик PSNR и SSIM для разных паттернов и значений силы встраивания  $S$  составляет 34,26–37,87 дБ и 0,95–0,97 соответственно. Стоит отметить, что, в отличие от оригинального алгоритма [3], модифицированный алгоритм обеспечивает высокую точность извлечения ЦВЗ без использования штатива для съёмки экрана.

### 3. ЗАКЛЮЧЕНИЕ

В данном исследовании была предложена модификация алгоритма встраивания ЦВЗ в видеоданные на основе прямоугольных паттернов, обладающего устойчивостью к съёмке с экрана [3]. Подход к уменьшению уровня искажений кадра при встраивании ЦВЗ путём выбора подходящей цветовой компоненты и размытия ЦВЗ по Гауссу улучшил незаметность встраивания данных. Подход к повышению точности определения границ экрана, демонстрирующего видео с ЦВЗ, за счёт применения нейронной сети повысил надёжность передачи встроенных данных. Таким образом, предложенная модификация повысила практическую применимость алгоритма для основного сценария использования.

### БЛАГОДАРНОСТИ

Работа выполнена при финансовой поддержке Российского научного фонда, проект № 21-71-10113.

### ЛИТЕРАТУРА

- [1] Егорова, А.А. Классификация и сравнительное исследование систем аутентификации JPEG-изображений, основанных на встраивании полухрупких водяных знаков / А.А. Егорова, В.А. Федосеев // Компьютерная оптика. – 2019. – Т. 43, № 3. – С. 419433. – DOI: 10.18287/2412-6179-2019-43-3-419-433
- [2] Выборнова, Ю.Д. Новый метод встраивания цифровых водяных знаков в векторные картографические данные / Ю.Д. Выборнова, В.В. Сергеев // Компьютерная оптика. – 2017. – Т. 41, № 6. – С. 913–919. – DOI: 10.18287/2412-6179-2017-41-6-913-919
- [3] Iwata, M. Practical Watermarking Method Estimating Watermarked Region from Recaptured Videos on Smartphone / M. Iwata, N. Mizushima, K. Kise // IEICE Transactions on Information and Systems. – 2017. – Vol. E100.D(1). – P. 24–32.
- [4] Ronneberger, O. U-Net: Convolutional Networks for Biomedical Image Segmentation / O. Ronneberger, P. Fischer, T. Brox // LNCS. – Vol. 9351. – P. 234–241.