

The Review of Existing Digital Fingerprinting Algorithms

Key words: watermarking, digital fingerprinting, content-based copy detection.

This article is dedicated to review of technical methods of multimedia data copyrighting, particularly of video content. This survey broaches such actual methods of video copyrighting as digital watermarking and digital fingerprinting. The task of content-based copy detection is described and existing techniques of digital fingerprinting are systematized.

П.Д. Гусев

ОБЗОР СУЩЕСТВУЮЩИХ АЛГОРИТМОВ ПОСТРОЕНИЯ ЦИФРОВЫХ ОТПЕЧАТКОВ

В настоящее время широкое распространение в сети Интернет получило такое явление, как интернет-пиратство, т.е. нарушение авторских прав, в том числе нелегальное копирование и распространение мультимедийных файлов. В то же время поиск и отслеживание всех нелегально распространенных копий являются, в общем случае, нетривиальной задачей. Один из способов защиты авторских прав является – в искомый объект мультимедиа так называемого цифрового водяного знака (ЦВЗ). Он представляет собой набор бит, встроенных в цифровые данные, но которые не воспринимаются как таковые [1]. Важнейшее применение цифровые водяные знаки нашли в системах защиты от копирования, которые стремятся предотвратить несанкционированное копирование цифровых данных. Даже в случае несанкционированного копирования медиафайла ЦВЗ позволяет однозначно установить владельца авторских прав на этот файл. Цифровые водяные знаки используются для защиты авторских прав уже довольно давно, и потому это направление очень хорошо изучено. Чаще всего объектом (контейнером), в который встраивается ЦВЗ, является графическое изображение, хотя существуют алгоритмы встраивания цифрового водяного знака в видео- и аудиофайлы. Несмотря на явные преимущества, ЦВЗ обладают рядом недостатков, которые делают их недостаточно эффективным средством защиты авторских прав на видеоконтент. Во-первых, при любом перекодировании видео-файла практически все ЦВЗ уничтожаются. Во-вторых, может возникнуть ситуация, когда правообладатель хочет изъять все копии оригинального видеофайла, в который не был изначально внедрен ЦВЗ, из открытого доступа. В этом случае применение только ЦВЗ нецелесообразно, поэтому, наряду с ЦВЗ, в частности для видеофайлов, применяется также метод, называемый «цифровой отпечаток» (digital video fingerprinting).

Цифровой отпечаток (digital fingerprint) представляет собой некий набор свойств исходной видеопоследовательности, характеризующих содержимое этого видео. Иными словами, цифровой отпечаток применяется в системах обнаружения копирования на основе содержимого видеопотока (Content-Based Copy Detection, CBCD). Главной идеей обнаружения копирования на основе содержимого видеопотока является использование самих данных видео в качестве ЦВЗ. Считается, что медиафайл содержит достаточно уникальной информации, которую можно использовать в решении задачи обнаружения копий. Схему обнаружения копирования можно описать в общих чертах следующим образом. Предположим, что имеется два видеофайла: оригинальный и тестируемый. Из обоих файлов извлекаются цифровые отпечатки, затем сравниваются между собой и, согласно определенному алгоритму принятия решения (voting algorithm),

устанавливается, содержит ли тестируемый видеопоток копию оригинальной видеопоследовательности или же нет [2].

По сравнению с цифровым водяным знаком цифровой отпечаток – относительно новое понятие. Это направление до сих пор изучено недостаточно хорошо, даже несмотря на существование множества алгоритмов и техник извлечения цифровых отпечатков. Что касается практических реализаций существующих алгоритмов, то в силу специфики этой области все они – закрытые, что не дает возможность оценить эффективность работы подхода в реализации цифрового отпечатка.

Задача обнаружения копирования на основе содержимого видеопотока возникла только в XXI веке, когда крупные теле- и кинокомпании пришли к выводу, что пиратский контент необходимо каким-либо образом обнаруживать и удалять, наказывая при этом людей, нарушающих авторское право. Нелегальные копии дисков, которые распространялись повсюду, контролировать было практически невозможно, но по мере того как распространение нелегального контента переходило в Интернет, правообладатели задались вопросом можно ли создать алгоритм, который мог бы обнаруживать видеопоследовательности, даже если они были незначительно изменены? Поначалу предлагалось использовать и отслеживать цифровой водяной знак, по аналогии с тем, как это было реализовано для статических изображений. Оказалось, однако, что любое пережатие видео и добавление шумов делают цифровой водяной знак несчитываемым. Тогда и появились первые идеи для алгоритмов, идентифицирующих видео по его содержанию.

В настоящее время, цифровые отпечатки применяются в двух областях. Во-первых, их используют в задачах компьютерно-технических экспертиз, например, в задачах поиска и обнаружения видеофайлов на носителях. Второе направление, в котором применяют цифровые отпечатки, является уже упомянутая задача обнаружения фактов копирования и распространения видеоматериалов, защищенных авторским правом. Ставить такую задачу на всем пространстве сети Интернет не представляется возможным, поэтому ее стараются решать локально, например, в пределах одного видеохостинга. Пример работы в этом направлении – технология ContentID, которую использует YouTube. Каждое видео, которое загружает пользователь, сравнивается с базой данных цифровых отпечатков медиафайлов, предоставленных правообладателями. Если найдется совпадение, правообладателю будет сообщено об этом, и он сможет заблокировать, удалить или получить прибыль с этого видео за счет пользователя. Также одно из подразделений AT&T занимается исследованием цифровых отпечатков, что говорит об актуальности данной тематики.

Как уже было отмечено выше, обнаружение копирования на основе содержимого видеопотока состоит из двух алгоритмов: алгоритма извлечения цифрового отпечатка (extraction algorithm) и алгоритма принятия решения (voting algorithm). Схематичное описание алгоритмов обнаружения копирования на основе содержимого видеопотока представлено на рис. 1. Алгоритм извлечения заключается в извлечении из видео какой-либо из его характеристик: пространственной, временной или пространственно-временной. Обычно для этих целей используются такие особенности видео, как яркость и цветность кадров. Алгоритм принятия решения заключается в сравнении двух имеющихся цифровых отпечатков с последующим принятием решения о том, содержит ли одно видео внутри себя копию другого.

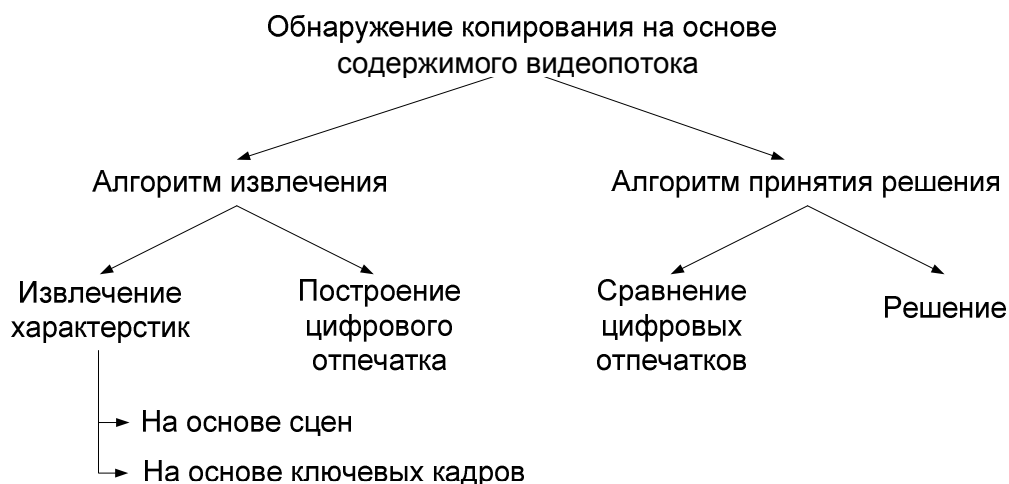


Рис. 1. Схематическое описание алгоритмов обнаружения копирования на основе содержимого видеопотока

Алгоритм извлечения цифрового отпечатка можно условно разделить на две составляющих: извлечение характеристик видео и построение цифрового отпечатка на основе этих характеристик. Существует два принципиально разных подхода к извлечению цифрового отпечатка: на основе ключевых кадров и на основе сцен. В алгоритмах на основе ключевых кадров сначала находится набор некоторых ключевых кадров, затем из каждого такого кадра извлекаются его характеристики в виде вектора (здесь под вектором понимается просто набор характеристик) и, наконец, векторы от каждого ключевого кадра либо комбинируются, формируя новый вектор, либо просто формируют последовательность векторов. В алгоритмах на основе сцен векторы извлекаются от каждого кадра сцены, и также из имеющихся векторов комбинируется новый вектор для сцены или просто формируется векторная последовательность.

Алгоритмы извлечения цифрового отпечатка можно также разделить на два класса: на основе глобальных дескрипторов и локальных дескрипторов. Цифровые отпечатки на основе глобальных дескрипторов используют характеристики всех пикселей кадра, в то время как цифровые отпечатки на основе локальных дескрипторов используют характеристики так называемых значимых точек (points of interest).

Также можно разделить существующие алгоритмы, основанные на глобальных дескрипторах, на классы по тому, какая именно характеристика извлекается из видеопотока: границы сцен, цветность, яркость и интенсивность, изменения в кадре или между кадрами, порядковые измерения и преобразования.

Границы сцен. Этот класс алгоритмов включает в себя метод, предложенный в работе «Finding pirated video sequences on the internet» [3]. В нем используются временные отпечатки границ, основанные только на границах сцен видеопоследовательности, и в качестве цифрового отпечатка используется время между границами сцен. Данная техника может быть эффективна в нахождении, скажем, кинофильма, но для коротких видео с малым числом сцен она может отрабатывать некорректно.

Цветность, яркость и интенсивность. Этот класс методов включает в себя техники, использующие в качестве цифрового отпечатка гистограммы или функции от интенсивности. Цифровой отпечаток на основе цветовой гистограммы [4] использует только цветовые свойства изображений без использования информации о пространственных свойствах. Этот метод может быть неэффективным для видеофайлов, имею-

ших несколько кадров в разных частях видео одной и той же цветовой схемы, которые невозможно отличить без использования пространственной информации. Эту проблему можно решить разделением кадра на блоки и извлечением гистограмм от каждого блока. Однако цветовые свойства кадра (например, гистограмма яркости, тона и насыщенности) восприимчивы к цветовым изменениям, вызванным кодирующими устройствами. Другие методы данного класса также имеют недостаточную чувствительность.

Изменения в кадре или между кадрами. Алгоритмы этого класса используют в качестве цифрового отпечатка связь между пикселями (пространственное распределение), изменение значений пикселей, находящихся в одной и той же области или изменение участка интенсивности в одном блоке между двумя последовательными кадрами (временное распределение). Именно связь – ключевой фактор, дающий достаточную пространственную и временную информацию. Цифровые отпечатки, основанные на движении, используют зависимость интенсивности от времени, потому эти алгоритмы работают корректно в случае глобального изменения цветовой гаммы, однако они не учитывают пространственную информацию. Временное измерение, предложенное в работе [5] также отбрасывает пространственную информацию, а изменение пикселей между кадрами не является достаточно достоверным. Центр масс направления градиента (Centroids of gradient orientation) [6] использует пространственную связь между пикселями.

Порядковые измерения. Первый метод порядковых измерений был предложен в работе [7] для вычисления соответствия изображений и позднее был адаптирован для применения к видео. Суть алгоритмов этого класса заключается в разбиении изображения на блоки, извлечении какой-то характеристики из каждого блока и нумерации блоков по возрастанию этой характеристики. В качестве цифрового отпечатка хранится последовательность порядковых номеров блоков каждого кадра согласно возрастанию этой характеристики. Эти алгоритмы работают корректно в случае глобальных изменений качества видео, которые являются неотъемлемой частью оцифровки или перекодирования сигнала. Недостаток такого рода методов в их недостаточной эффективности в случаях вставки логотипов, сдвигов и масштабирования, что является наиболее частыми изменениями в телевизионной обработке. На самом деле, эти изменения – большая проблема для всех алгоритмов, основанных на глобальных дескрипторах.

Преобразования. Техники этого класса алгоритмов используют более сложные математические формулы для генерации цифрового отпечатка, например, преобразование Фурье. Существуют алгоритмы, использующие компактное преобразование Фурье-Меллина [8], дискретное косинусное преобразование [9] и вейвлет-преобразование [10].

В противовес цифровому отпечатку, учитывающему все пиксели в кадре, отпечаток, основанный на локальных дескрипторах, использует характеристики только значимых точек. Для всех алгоритмов этого класса необходимо предварительное извлечение значимых точек. Для этих целей наиболее часто применяется детектор Харриса и Стивенса [11]. Алгоритм SIFT [12] представляет собой особый случай методов, основанных на локальных дескрипторах, поскольку цифровой отпечаток ключевого кадра представляет собой комбинацию всех локальных дескрипторов в виде вектора. Эта операция уменьшает вычислительную сложность, но также уменьшает и точность.

В дальнейшей работе на основе проведенного анализа будут выработаны требования к эффективному алгоритму построения цифрового отпечатка, а также предложены идеи по модификации существующих алгоритмов.

СПИСОК ЛИТЕРАТУРЫ:

1. Грибунин В.Г., Оков И.Н., Туринцев И.В. Цифровая стеганография. М.: СОЛОН-Пресс, 2002.
2. Chao W. Introduction to Video Fingerprinting [Электронный ресурс]. Режим доступа: <http://disp.ee.ntu.edu.tw/~pujols/Introduction%20to%20Video%20Fingerprinting.pdf> (дата обращения: 10.08.2015).
3. Indyk P. Finding pirated video sequences on the internet [Электронный ресурс]. Режим доступа: <http://csce.uark.edu/~jgauch/library/Video/Indyk.1999.pdf> (дата обращения: 10.08.2015).
4. Naphade Y. M. A novel scheme for fast and efficient video sequence matching using compact signatures [Электронный ресурс]. Режим доступа: <http://csce.uark.edu/~jgauch/library/Video/Naphade.2000.pdf> (дата обращения: 10.08.2015).
5. Law-To J. Video copy detection: a comparative study [Электронный ресурс]. Режим доступа: http://www.irisa.fr/vista/Papers/2007_civr_law-to.pdf (дата обращения: 10.08.2015).
6. Lee S. Video fingerprinting based on centroids of gradients orientations [Электронный ресурс]. Режим доступа: https://wiki.dcs.shef.ac.uk/wiki/pub/Darwin2010/VF_Proposal_MGT_Page/VFProposal1.10.pdf (дата обращения: 10.08.2015).
7. Bhat D. Ordinal measures for image correspondence [Электронный ресурс]. Режим доступа: http://www1.cs.columbia.edu/CAVE/publications/pdfs/Bhat_PAMI98.pdf (дата обращения: 10.08.2015).
8. Derrode S. Robust and efficient Fourier-Mellin transform approximations for gray-level image reconstruction and complete invariant description [Электронный ресурс]. Режим доступа: <http://www.fresnel.fr/perso/derrode/publi/Cviu01.pdf> (дата обращения: 10.08.2015).
9. Coskun B. Spatio-temporal transform-based video hashing [Электронный ресурс]. Режим доступа: http://www.busim.ee.boun.edu.tr/~sankur/SankurFolder/SpatioTemporal_Hashing.pdf (дата обращения: 11.08.2015).
10. Chang E. Rime – a replicated image detector for the world-wide web [Электронный ресурс]. Режим доступа: <ftp://db.stanford.edu/pub/dbpubs/1998/21/21.pdf.gz> (дата обращения: 11.08.2015).
11. Harris C. A combined corner and edge detector [Электронный ресурс]. Режим доступа: <http://csce.uark.edu/~jgauch/library/Features/Harris.1988.pdf> (дата обращения: 11.08.2015).
12. Lowe D. Distinctive image features from scale-invariant keypoints [Электронный ресурс]. Режим доступа: <http://www.cs.ubc.ca/~lowe/papers/ijcv04.pdf> (дата обращения: 11.08.2015).

REFERENCES:

1. Gribunin V.G., Okov I.N., Turintsev I.V. Cifrovaya steganographiya. M.: SOLON-Press, 2002.
2. Chao W. Introduction to Video Fingerprinting [Electronic resource]. Access policy: <http://disp.ee.ntu.edu.tw/~pujols/Introduction%20to%20Video%20Fingerprinting.pdf> (date of the access: 10.08.2015).
3. Indyk P. Finding pirated video sequences on the internet [Electronic resource]. Access policy: <http://csce.uark.edu/~jgauch/library/Video/Indyk.1999.pdf> (date of the access: 10.08.2015).
4. Naphade Y. M. A novel scheme for fast and efficient video sequence matching using compact signatures [Electronic resource]. – Access policy: <http://csce.uark.edu/~jgauch/library/Video/Naphade.2000.pdf> (date of the access: 10.08.2015).
5. Law-To J. Video copy detection: a comparative study [Electronic resource]. Access policy: http://www.irisa.fr/vista/Papers/2007_civr_law-to.pdf (date of the access: 10.08.2015).
6. Lee S. Video fingerprinting based on centroids of gradients orientations [Electronic resource]. Access policy: https://wiki.dcs.shef.ac.uk/wiki/pub/Darwin2010/VF_Proposal_MGT_Page/VFProposal1.10.pdf (date of the access: 10.08.2015).
7. Bhat D. Ordinal measures for image correspondence [Electronic resource]. Access policy: http://www1.cs.columbia.edu/CAVE/publications/pdfs/Bhat_PAMI98.pdf (date of the access: 10.08.2015).
8. Derrode S. Robust and efficient Fourier-Mellin transform approximations for gray-level image reconstruction and complete invariant description [Electronic resource]. Access policy: <http://www.fresnel.fr/perso/derrode/publi/Cviu01.pdf> (date of the access: 10.08.2015).
9. Coskun B. Spatio-temporal transform-based video hashing [Electronic resource]. Access policy: http://www.busim.ee.boun.edu.tr/~sankur/SankurFolder/SpatioTemporal_Hashing.pdf (date of the access: 11.08.2015).
10. Chang E. Rime – a replicated image detector for the world-wide web [Electronic resource]. Access policy: <ftp://db.stanford.edu/pub/dbpubs/1998/21/21.pdf.gz> (date of the access: 11.08.2015).
11. Harris C. A combined corner and edge detector [Electronic resource]. Access policy: <http://csce.uark.edu/~jgauch/library/Features/Harris.1988.pdf> (date of the access: 11.08.2015).
12. Lowe D. Distinctive image features from scale-invariant keypoints [Electronic resource]. Access policy: <http://www.cs.ubc.ca/~lowe/papers/ijcv04.pdf> (date of the access: 11.08.2015).