

С.В. КУЛЕШОВ, А.А. ЗАЙЦЕВА, А.Ю. АКСЕНОВ  
**ИНФОРМАЦИОННАЯ ИЗБЫТОЧНОСТЬ  
В МУЛЬТИМЕДИЙНЫХ ДАННЫХ**

---

*Кулешов С.В., Зайцева А.А., Аксенов А.Ю.* **Информационная избыточность в мультимедийных данных.**

**Аннотация.** В статье рассматривается феномен возникновения информационной избыточности в мультимедийных данных на примере изображений и видеоданных. Приводятся примеры появления и изменения информационной избыточности, предлагается метод ее оценки.

**Ключевые слова:** информационная избыточность, предобработка, мультимедийные данные, компрессия.

*Kuleshov S.V., Zaytseva A.A., Aksenov A.Y.* **Information redundancy in multimedia data.**

**Abstract.** The paper considers the phenomena of information redundancy emergence in multimedia data on the example of images and video. The examples of redundancy emergence and variation are given, the method for redundancy estimation is proposed.

**Keywords:** redundancy, preprocessing, multimedia, compression, video.

---

**1. Введение.** Современные тенденции развития методов доставки и распространения цифрового информационного контента до потребителя демонстрируют постепенный переход от телекоммуникации к инфотелекоммуникации (или просто инфокоммуникации), в которой информационные потоки рассматриваются не с точки зрения последовательностей битов, обладающих некоторыми статистическими характеристиками, а как контент, наделяемый смысловым содержанием, важным для потребителя.

Это приводит к появлению в данных информационной избыточности — явления, при котором количество потенциально различных состояний передаваемого информационного объекта превышает количество состояний, различимых потребителем или требуемых для решения некоторой задачи. Другими словами информационная избыточность цифрового информационного объекта проявляется в том, что часть его битового объема может быть удалена или изменена незаметно для потребителя.

Таким образом, наличие информационной избыточности с одной стороны является предпосылкой для возможности уменьшения объема передаваемых данных [1–7], а с другой стороны открывает возможности для возникновения угроз информационной безопасности и различных видов кибермошенничества.

Следует отметить, что в данной статье не будет рассматриваться вариант искусственного повышения информационной избыточности с целью помехоустойчивого кодирования.

Негативное влияние информационной избыточности может проявиться, например, при распространении информационного контента, заявленные параметры которого не соответствуют реальным. Наиболее известными примерами такого распространения на данный момент являются: продажа из фотобанков «дорогостоящих высококачественных» изображений, полученных из изображений меньшего размера (соответственно, более дешевых) путем линейной интерполяции, распространение видеоконтента стандартного (SD) качества на HD-носителях (например, фильм DVD-качества на дисках BD-формата) и ряд других.

Таким образом, разработка теоретических подходов к описанию информационной избыточности и методов инструментального определения информационной избыточности является актуальной современной задачей.

**2. Характерные типы возникновения информационной избыточности.** Рассмотрим наиболее распространенные примеры возникновения избыточности вследствие применения преобразований, используемых в популярных методах обработки изображений и видеоданных.

1. Появление пространственной избыточности при увеличении размера изображения (upscale) с применением интерполяции. Данный вид преобразования очень распространен на практике, его особенность заключается в увеличении количества элементов (пикселей), из которых состоит объект, при этом количество различных (идентифицируемых) состояний объекта не изменяется (рис. 1).

2. Появление временной избыточности при повышении частоты кадров для видеоданных (рис. 2). По аналогии для аудиоданных временная избыточность появляется при повышении частоты сэмплирования. Подобные преобразования могут осуществляться, например, при конвертации различных форматов для обеспечения совместимости аппаратуры записи и воспроизведения.

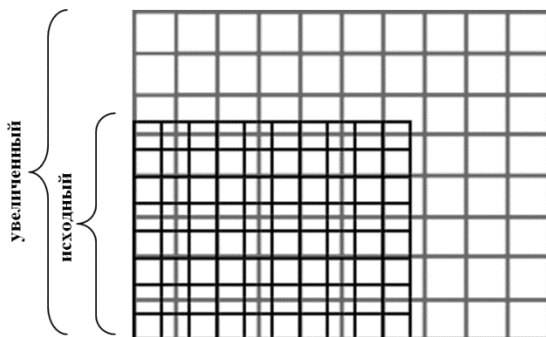


Рис. 1. Иллюстрация появления избыточности при увеличении изображения.

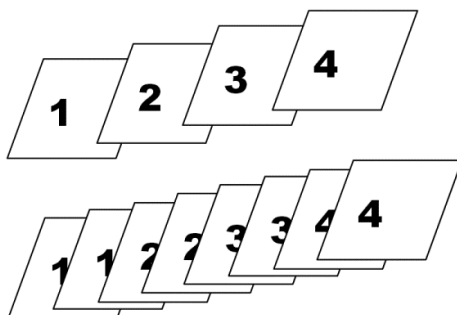


Рис. 2. Иллюстрация появления избыточности на примере повышения частоты кадров.

Перечисленные примеры представляют наиболее типичные случаи возникновения информационной избыточности и дают предпосылки для создания метода ее оценки.

Существует ряд других способов возникновения и изменения информационной избыточности. Например, рассмотрим простейшую последовательность ограничений при восприятии визуальных данных человеком–потребителем: потребитель не способен увидеть (ограничение физиологических возможностей зрения) — не способен воспринять (ограничение способностей восприятия, например скоростных) —

не способен понять (ограничение, основанное на ограниченности предзнания).

### **3. Типовые способы снижения информационной избыточности:**

1. Использование различных видов сегментации для снижения информационной избыточности при последующей автоматизированной обработке.

2. Применение сжатия с потерями.

В большинстве методов сжатия визуальных данных с потерями содержится этап квантования, заключающийся в уменьшении точности представления данных в соответствии с некоторым алгоритмом (моделью восприятия). На рис. 3 приведен пример использования квантования для значений яркости изображения.

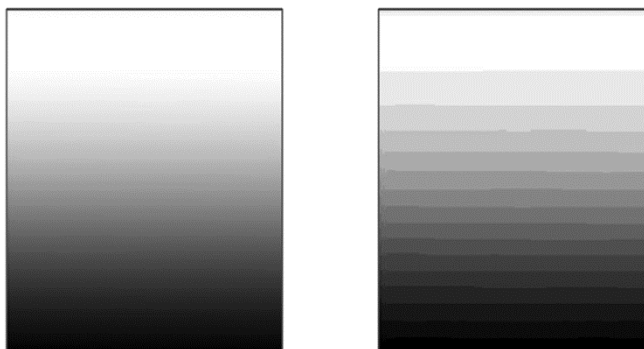


Рис. 3. Пример использования квантования: слева — исходное изображение с точностью представления 8 бит, справа — изображение после применения квантования (точность представления 4 бита).

3. Уменьшение размера изображения и уменьшение частоты кадров для видеоданных.

**4. Определение информационной избыточности визуальных данных.** Анализ различных типов возникновения и снижения информационной избыточности позволил сформулировать наиболее общее определение информационной избыточности, а также сформулировать признаки ее наличия.

Информационная избыточность — увеличение количества дискретных элементов, составляющих информационный объект, которое не приводит к увеличению семантически различимых (идентифицируемых) состояний самого информационного объекта.

Введем следующие обозначения:  $N_1$  — количество различных состояний информационного объекта (полезное),  $N_2$  — количество различных состояний информационного объекта (полное),  $l_1$  — количество элементов (полезное),  $l_2$  — количество элементов (полное).

Тогда  $\frac{l_2}{l_1} = R$  — избыточность, возникающая в результате преоб-

разования,  $l_2^{p_2} = N_2 > N_1 = l_1^{p_1}$ , где  $l_1$  определяется как эквивалентное число дискретных элементов информационного объекта, соответствующее количеству идентифицируемых состояний  $N_1$ ,  $p_1$  — число различных состояний дискретного элемента (полезное),  $p_2$  — число различных состояний дискретного элемента (полное).

Для примера рассмотрим информационную избыточность изображения размером 640x480 пикселей с глубиной цвета 24 бит. Полное количество различных состояний для такого изображения равно 307200<sup>24</sup>. При масштабировании такого изображения до размера 1024x768 полное количество идентифицируемых состояний  $N_2$  возрастает до значения 786432<sup>24</sup> при неизменном значении  $N_1$ , что означает появление информационной избыточности равной 2,56<sup>24</sup>. Подобный вид информационной избыточности дают, например, web-камеры, имеющие cmos-сенсор 0,3 МПикс (стандартное VGA-разрешение) и выдающее изображение размером 1024x768.

Для случаев, когда количество различных состояний заранее неизвестно, например, при сжатии с потерями JPEG, априорная оценка величины информационной избыточности более сложна, так как зависит от самих данных.

**5. Метод последовательной оценки информационной избыточности визуальных данных.** При оценке информационной избыточности, вносимой инфокоммуникационными системами, состоящими из последовательности независимых этапов обработки, требуется учитывать величину вносимой информационной избыточности с последующим определением общей информационной избыточности всей последовательности этапов.

Рассмотрим метод определения общей информационной избыточности всех этапов.

Предпосылки метода основаны на аналогии с физическим принципом Гюйгенса–Френеля, утверждающего, что каждая точка поверх-

ности, достигнутая световой волной, является вторичным источником световых волн. Тогда можно утверждать, что каждый предыдущий этап является источником данных для последующих этапов и может рассматриваться независимо от остальных, обладающих собственными характеристиками информационной избыточности.

Достоинством метода последовательной оценки информационной избыточности является возможность оценить общую информационную избыточность при отсутствии доступа к техническим параметрам отдельных этапов обработки и передачи.

Каждый этап обработки является источником вторичных данных, данные после него рассматриваются независимо, а изменение избыточности оценивается непосредственно изменением числа идентифицируемых элементов изображений до и после этапа дискретизации.

Избыточность этапа обработки определяется как:

$$R_D = \frac{m_1}{m_2},$$

где  $m_1$  — количество различных элементов изображения до этапа дискретизации,  $m_2$  — количество элементов изображения после этапа дискретизации.

Особенность цифровых систем заключается в том, что при передаче или сохранении визуальных данных вводимая дискретность уменьшает мощность множества значений измеряемых параметров и формирует однозначный идентификатор изображения как совокупность цветовых компонентов, находящихся в определенных дискретных позициях. Уменьшение мощности множества значений измеряемых параметров приводит к уменьшению мощности множества представимых изображений. В случае если в инфокоммуникационной системе содержится несколько этапов передискретизации ( $N_d > 1$ ), например, дискретизация при экспонировании на элементы светочувствительной матрицы, дискретизация при обработке и компрессии, дискретизация при отображении на экране монитора, то при этом обобщенная избыточность последовательности этапов равна  $R = R_1 R_2 \dots R_n$ , где  $R_i$  — избыточность  $i$ -го этапа обработки.

**6. Заключение.** Новый уровень технологического развития, связанный с использованием цифровых методов обработки данных, обусловил появление феномена информационной избыточности.

Наличие информационной избыточности с одной стороны дает возможность уменьшать битовый объем представления мультимедий-

ных данных при их передаче и хранении, а с другой стороны оставляет возможность для негативных манипуляций над контентом.

Использование особенностей избыточности способно повысить эксплуатационные характеристики методов сжатия данных, используемых при обработке мультимедийных данных, а также выявить и снизить риск мошеннических действий, производимых над цифровым контентом.

### Литература

1. Кулешов С.В. Пространственно-временное представление, обработка и компрессия видеопотока. // «Информационно-измерительные и управляющие системы». 2008. №4, т.6. С. 33–37.
2. Кулешов С.В., Зайцева А.А., Аксенов А.Ю. Ассоциативно-пирамидальное представление данных. // «Информационно-измерительные и управляющие системы» 2008. №4, т.6. С. 14–17.
3. Milic D., Brogle M., Braun T. Video Broadcasting using Overlay Multicast. // ISM. Seventh IEEE International Symposium on Multimedia (ISM'05). 2005. PP. 515–522.
4. Nguyen K., Nguyen Th., Cheung S.-Ch. Video Streaming with Network Coding. // Signal Processing Systems. 2010. 59(3). PP. 319–333.
5. Abd-Elrahman E., Boutabia M., Afifi H. Video Streaming Security: Reliable Hash Chain Mechanism Using Redundancy Codes. // MoMM2010 Proceedings 2010. PP. 69–76.
6. Hui Wang, Joyce Liang, C.-C. Jay Kuo. Overview of Robust Video Streaming with Network Coding // Journal of Information Hiding and Multimedia Signal Processing Volume 1, Number 1, January 2010. PP. 36–50.
7. Zhenjiang Li, Yao Yu, Xiaojun Hei, Danny H.K. Tsang. Towards Low-Redundancy Push-Pull P2P Live Streaming. [Электронный ресурс]. — Доступ: [http://www.cse.ust.hk/~lzjiang/2008QShine\\_PullPush.pdf](http://www.cse.ust.hk/~lzjiang/2008QShine_PullPush.pdf)

**Кулешов Сергей Викторович** — д.т.н, старший научный сотрудник лаборатории автоматизации научных исследований СПИИРАН. Область научных интересов: инфологические информационные системы, инфокоммуникационные системы, гибридные кодеки, обработка видеоданных. Число научных публикаций — 70. [kuleshov@iias.spb.su](mailto:kuleshov@iias.spb.su), [sial.iias.spb.su](mailto:sial.iias.spb.su); СПИИРАН, 14-я линия В.О., 39, Санкт-Петербург, 199178, РФ; р.т. +7(812)323-5139, факс +7(812)328-4450.

**Kuleshov Sergey Victorovich** —D.Sc. in Tech. Sc., Senior researcher, Laboratory of Research Automation SPIIRAS. Research interest: infology information systems, infocommunication systems, hybrid codecs, video data streams processing. The number of publication — 70. [kuleshov@iias.spb.su](mailto:kuleshov@iias.spb.su), [sial.iias.spb.su](mailto:sial.iias.spb.su); SPIIRAS, 14-th Line V.O., 39, St. Petersburg, 199178, Russia; office phone +7(812)323-5139, fax +7(812)328-4450.

**Зайцева Александра Алексеевна** — канд. техн. наук, научный сотрудник лаборатории автоматизации научных исследований СПИИРАН. Область научных интересов: обработка видеоданных, когнитивная 3D технология быстрого прототипирования. Число научных публикаций — 20. [cher@iias.spb.su](mailto:cher@iias.spb.su), [sial.iias.spb.su](mailto:sial.iias.spb.su); СПИИРАН, 14-я линия В.О., 39, Санкт-Петербург, 199178, РФ; р.т. +7(812)323-5139, факс +7(812)328-4450.

**Zaytseva Alexandra Alexeevna** — PhD in Tech. Sc., researcher, Laboratory of Research Automation SPIIRAS. Research interest: video data streams processing, 3D rapid prototyping technology. The number of publication — 20. [cher@iias.spb.su](mailto:cher@iias.spb.su), [sial.iias.spb.su](mailto:sial.iias.spb.su); SPIIRAS, 14-th Line V.O., 39, St. Petersburg, 199178, Russia; office phone +7(812)323-5139, fax +7(812)328-4450.

**Аксенов Алексей Юрьевич** — младший научный сотрудник лаборатории автоматизации научных исследований СПИИРАН. Область научных интересов: адаптивно динамическая сегментация сигналов. Число научных публикаций — 11. [a\\_aksenov@mail.iias.spb.su](mailto:a_aksenov@mail.iias.spb.su), [sial.iias.spb.su](mailto:sial.iias.spb.su); СПИИРАН, 14-я линия В.О., 39, Санкт-Петербург, 199178, РФ; р.т. +7(812)323-5139, факс +7(812)328-4450.

**Aksenov Alexey Yurievich** — researcher, Laboratory of Research Automation SPIIRAS. Research interest: signal processing. The number of publication — 11. [a\\_aksenov@mail.iias.spb.su](mailto:a_aksenov@mail.iias.spb.su), [sial.iias.spb.su](mailto:sial.iias.spb.su); SPIIRAS, 14-th Line V.O., 39, St. Petersburg, 199178, Russia; office phone +7(812)323-5139, fax +7(812)328-4450.

Рекомендовано лабораторией автоматизации научных исследований СПИИРАН.  
Статья поступила в редакцию 28.09.2012.



## РЕФЕРАТ

*Кулешов С.В., Зайцева А.А., Аксенов А.Ю.* **Информационная избыточность в мультимедийных данных.**

Современные тенденции развития методов доставки и распространения цифрового информационного контента до потребителя демонстрируют постепенный переход от телекоммуникации к инфокоммуникации, в которой информационные потоки рассматриваются не с точки зрения последовательностей битов, обладающих некоторыми статистическими характеристиками, а как контент, наделяемый смысловым содержанием, важным для потребителя. Это приводит к появлению информационной избыточности — явления, при котором количество потенциально различных состояний передаваемого информационного объекта превышает количество состояний, различимых потребителем или требуемых для решения конкретной задачи. Другими словами, информационная избыточность цифрового информационного объекта проявляется в том, что часть битового объема в его представлении может быть удалена или изменена незаметно для потребителя.

Таким образом, наличие информационной избыточности с одной стороны является предпосылкой для уменьшения объема передаваемых данных, а с другой стороны открывает возможности для возникновения угроз информационной безопасности и различных видов кибермошенничества.

Анализ различных типов возникновения и способов снижения информационной избыточности позволил сформулировать наиболее общее определение информационной избыточности, а также сформулировать признаки ее наличия.

Информационная избыточность — увеличение количества дискретных элементов, составляющих информационный объект, которое не приводит к увеличению семантически различимых (идентифицируемых) состояний самого информационного объекта.

Изучение особенностей информационной избыточности способно повысить эксплуатационные характеристики методов сжатия данных, используемых в кодеках мультимедийных данных, а также выявить и предупредить мошеннические действия, производимые над цифровым контентом.

## SUMMARY

*Kuleshov S.V., Zaytseva A.A., Aksenov A.Y.* **Information redundancy in multimedia data.**

Modern trends in methods of delivery and distribution of digital information content to the consumer demonstrate a gradual transition from telecommunications to information communications in which information flows are not considered in terms of the bit patterns having some statistical characteristics, but as a content with semantics important to the consumer. This leads to information redundancy — phenomena in which the number of potentially distinct states of transferred informational object is surpassing the consumers abilities to percept. In other words information redundancy allows to commit or change some parts of bit content insensibly to consumer.

Thus information redundancy from one hand is a precondition for reducing data rates and from the other it opens way for different types of cyber crime.

Analysis of different types of appearance and ways of reducing of information redundancy allowed us to formulate the most general definition of information redundancy, and to formulate the signs of its presence. Information redundancy is the increasing of the number of discrete elements that make up an information object, which does not increase the semantically distinct (identified) states of the information object.

The study the features of information redundancy can enhance the performance of data compression methods used in multimedia codecs, as well as to identify and prevent fraudulent activities performed on digital content.