

**Министерство образования Республики Беларусь
Учреждение образования
Гомельский государственный университет им. Ф. Скорины**

Физический факультет

«Инженерная и компьютерная графика»

Лекция – Графические форматы и методы их сжатия

Лектор – ст. преподаватель Грищенко В.В.

Форматы цифровых изображений

В компьютерной графике применяют по меньшей мере **три десятка форматов** файлов для хранения изображений. Но лишь часть из них стала стандартом “де-факто” и применяется в подавляющем большинстве программ. Как правило, несовместимые форматы имеют файлы растровых, векторных, трехмерных изображений, хотя существуют форматы, позволяющие хранить данные разных классов. Многие приложения ориентированы на собственные “**специфические**” **форматы**, перенос их файлов в другие программы вынуждает использовать специальные фильтры или экспортировать изображения в “**стандартный**” **формат**.

Форматы цифровых изображений

TIFF (*Tagged Image File Format*). Формат предназначен для хранения растровых изображений высокого качества (расширение имени файла .TIF). Относится к числу широко распространенных, отличается переносимостью между платформами (IBM PC и Apple Macintosh), обеспечен поддержкой со стороны большинства графических, верстальных и дизайнерских программ. Предусматривает широкий диапазон цветового охвата – от монохромного черно-белого до 32-разрядной модели цветоделения CMYK. Начиная с версии 6.0 в формате TIFF можно хранить сведения о масках (контурах) изображений. Для уменьшения размера файла применяется встроенный алгоритм сжатия.

Форматы цифровых изображений

PSD (*PhotoShop Document*). Собственный формат программы Adobe Photoshop (расширение имени файла .PSD), один из наиболее мощных по возможностям хранения растровой графической информации. Позволяет запоминать параметры слоев, каналов, степени прозрачности, множества масок. Поддерживаются 48-разрядное кодирование цвета, цветоделение и различные цветовые модели. Основной недостаток выражен в том, что отсутствие эффективного алгоритма сжатия информации приводит к большому объему файлов.

Форматы цифровых изображений

PCX. Формат появился как формат хранения растровых данных программы PC PaintBrush фирмы Z-Soft и является одним из наиболее распространенных (расширение имени файла .PCX). Отсутствие возможности хранить цвето-деленные изображения, недостаточность цветовых моделей и другие ограничения привели к утрате популярности формата. В настоящее время считается устаревшим.

JPEG (*Joint Photographic Experts Group*). Формат предназначен для хранения растровых изображений (расширение имени файла .JPG). Позволяет регулировать соотношение между степенью сжатия файла и качеством изображения. Применяемые методы сжатия основаны на удалении “избыточной” информации, поэтому формат рекомендуют использовать только для электронных публикаций.

Форматы цифровых изображений

GIF (*Graphics Interchange Format*). Стандартизирован в 1987 году как средство хранения сжатых изображений с фиксированным (256) количеством цветов (расширение имени файла .GIF). Получил популярность в Интернете благодаря высокой степени сжатия. Последняя версия формата GIF89a позволяет выполнять чересстрочную загрузку изображений и создавать рисунки с прозрачным фоном. Ограниченные возможности по количеству цветов обуславливают его применение исключительно в электронных публикациях.

PNG (*Portable Network Graphics*). Сравнительно новый (1995 год) формат хранения изображений для их публикации в Интернете (расширение имени файла .PNG). Поддерживаются три типа изображений – цветные с глубиной 8 или 24 бита и черно-белое с градацией 256 оттенков серого. Сжатие информации происходит практически без потерь, предусмотрены 254 уровня альфа-канала, чересстрочная развертка.

Форматы цифровых изображений

WMF (*Windows MetaFile*). Формат хранения векторных изображений операционной системы Windows (расширение имени файла .WMF). По определению поддерживается всеми приложениями этой системы. Однако отсутствие средств для работы со стандартизированными цветовыми палитрами, принятыми в полиграфии, и другие недостатки ограничивают его применение.

PDF (*Portable Document Format*). Формат описания документов, разработанный фирмой Adobe (расширение имени файла .PDF). Хотя этот формат в основном предназначен для хранения документа целиком, его впечатляющие возможности позволяют обеспечить эффективное представление изображений. Формат является аппаратно-независимым, поэтому вывод изображений допустим на любых устройствах – от экрана монитора до фото-экспонирующего устройства. Мощный алгоритм сжатия со средствами управления итоговым разрешением изображения обеспечивает компактность файлов при высоком качестве иллюстраций

Форматы цифровых изображений

EPS (*Encapsulated PostScript*). Формат описания как векторных, так и растровых изображений на языке PostScript фирмы Adobe, фактическом стандарте в области допечатных процессов и полиграфии (расширение имени файла .EPS). Так как язык PostScript является универсальным, в файле могут одновременно храниться векторная и растровая графика, шрифты, контуры обтравки (маски), параметры калибровки оборудования, цветовые профили. Для отображения на экране векторного содержимого используется формат WMF, а растрового – TIFF. Но экранная копия лишь в общих чертах отображает реальное изображение, что является существенным недостатком EPS. Действительное изображение можно увидеть лишь на выходе выводного устройства, с помощью специальных программ просмотра или после преобразования файла в формат PDF в приложениях Acrobat Reader, Acrobat Exchange.

Форматы цифровых изображений

CDR - основной внутренний формат программы Corel DRAW (внутренний - это означает, что данный формат используется только программой Corel DRAW и никакой больше);

PCT - векторный формат, используемый на компьютерах Apple Macintosh в операционных системах Mac;

AI - внутренний формат файла для программы Adobe Illustrator;

PNG - еще один популярнейший растровый формат, используемый в основном в сфере дизайна. В этом формате отсутствует сжатие, изображение имеет высокое качество, но в то же время имеет большой размер. Формат поддерживает зоны прозрачности;

BMP (bit map) - карта битов. Один из простейших растровых форматов, который не использует сжатие. Максимально качественно передает все цветовые оттенки рисунка. Изображение в данном формате занимает на компьютере гораздо больше места, по сравнению с таким же изображением в другом растровом формате.

PSD - внутренний формат рисунка для программы Adobe Photoshop.

- **Характерной особенностью большинства «классических» типов данных, с которыми традиционно работают люди, является определенная избыточность.**
- **Степень избыточности зависит от типа данных.** Например, у видеоданных степень избыточности обычно в несколько раз больше, чем у графических данных, а степень избыточности графических данных в несколько раз больше, чем текстовых.
- **Кроме того, степень избыточности данных зависит от принятой системы кодирования.** Так, например, можно сказать, что кодирование текстовой информации средствами русского языка (с использованием русской азбуки) дает в среднем избыточность на 20-30% больше, чем кодирование адекватной информации средствами английского языка.

- **Избыточность можно уменьшить с помощью эффекта сжатия данных.**
- **В современной информационной науке важно выполнение резервного копирования информации. Резервной копией называется копия файла или диска, с помощью которой можно восстановить оригинал при его случайном искажении или уничтожении.**
- **При резервном копировании, возникает задача разработки специальных способов записи файлов, при которых файлы бы занимали на диске как можно меньше места, - это задача сжатия файлов.**

Способы сжатия

1. *Уплотнение файлов* применяют для уменьшения их размеров при подготовке к передаче по каналам электронных сетей или к транспортировке на внешнем носителе малой емкости, например на гибком диске.
2. *Уплотнение папок* используют как средство архивации данных перед длительным хранением, в частности, при резервном копировании.
3. *Уплотнение дисков* служит целям повышения эффективности использования их рабочего пространства и, как правило, применяется к дискам, имеющим недостаточную емкость.

Обратимость сжатия

Несмотря на изобилие алгоритмов сжатия данных, теоретически есть только **три способа** уменьшения их избыточности. Это либо изменение содержания данных, либо изменение их структуры, либо и то и другое вместе.

- Если при сжатии данных происходит изменение их содержания, метод сжатия необратим и при восстановлении данных из сжатого файла не происходит полного восстановления исходной последовательности. Такие методы называют также **методами сжатия с регулируемой потерей информации**.
- Они применимы только для тех типов данных, для которых формальная утрата части содержания не приводит к значительному снижению потребительских свойств. В первую очередь, это относится к мультимедийным данным: видеорядам, музыкальным записям, звукозаписям и рисункам.
- Методы сжатия с потерей информации обычно обеспечивают гораздо более высокую степень сжатия, чем обратимые методы, но их нельзя применять к текстовым документам, базам данных и, тем более, к программному коду.

- Понятно, что эти методы можно применять только для таких типов данных, для которых потеря части содержимого не приводит к существенному искажению информации. К таким типам данных относятся видео- и аудиоданные, а также графические данные. Методы сжатия с регулирующими потерями информации обеспечивают значительно большую степень сжатия, но их нельзя применять к текстовым данным. Примерами форматов сжатия с потерями информации могут быть:

- JPEG - для графических данных;
- MPG - для для видеоданных;
- MP3 - для аудиоданных.

- Если при сжатии данных происходит только изменение структуры данных, то метод сжатия называется обратимым. В этом случае, из архива можно восстановить информацию полностью. Обратимые методы сжатия можно применять к любым типам данных, но они дают меньшую степень сжатия по сравнению с необратимыми методами сжатия. Примеры форматов сжатия без потери информации:

- GIF, TIFF - для графических данных;

- AVI - для видеоданных;

- ZIP, ARJ, RAR, CAB, LH - для произвольных типов данных.

Алгоритмы обратимых методов

- При исследовании методов сжатия данных следует иметь в виду **существование следующих теорем.**
 - 1. Для любой последовательности данных существует теоретический предел сжатия, который не может быть превышен без потери части информации.
 - 2. Для любого алгоритма сжатия можно указать такую последовательность данных, для которой он обеспечит лучшую степень сжатия, чем другие методы.
 - 3. Для любого алгоритма сжатия можно указать такую последовательность данных, для которой данный алгоритм вообще не позволит получить сжатия.

Таким образом, обсуждая различные методы сжатия, следует иметь в виду, что наивысшую эффективность они демонстрируют для данных разных типов и разных объемов.

Алгоритм RLE

- В основу алгоритмов **RLE** (Run-Length Encoding) положен принцип выявления повторяющихся последовательностей данных и замены их простой структурой, в которой указывается код данных и коэффициент повтора.
- Например, для последовательности: **0; 0; 0; 127; 127; 0; 255; 255; 255; 255** (всего 10 байтов) образуется следующий вектор, при записи в строку который имеет вид:
0; 3; 127; 2; 0; 1; 255; 4 (всего 8 байтов).
- В данном примере коэффициент сжатия равен $8/10$ (80 %).

- Программные реализации алгоритмов **RLE** отличаются простотой, высокой скоростью работы, но в среднем обеспечивают недостаточное сжатие.
- Наилучшими объектами для данного алгоритма являются графические файлы, в которых большие одноцветные участки изображения кодируются длинными последовательностями одинаковых байтов.
- Этот метод также может давать заметный выигрыш на некоторых типах файлов баз данных, имеющих таблицы с фиксированной длиной полей. Для текстовых данных методы RLE, как правило, неэффективны.

Алгоритм KWE

- В основу алгоритмов кодирования по ключевым словам Keyword Encoding положено кодирование лексических единиц исходного документа группами байтов фиксированной длины. Примером лексической единицы может служить слово. Результат кодирования сводится в таблицу, которая прикладывается к результирующему коду и представляет собой словарь. Обычно для англоязычных текстов принято использовать двухбайтную кодировку слов. Образующиеся при этом пары байтов называют *токенами*.
- Эффективность данного метода существенно зависит от длины документа, поскольку из-за необходимости прикладывать к архиву словарь длина кратких документов не только не уменьшается, но даже возрастает.
- Данный алгоритм наиболее эффективен для англоязычных текстовых документов и файлов баз данных. Для русскоязычных документов, отличающихся увеличенной длиной слов и большим количеством приставок, суффиксов и окончаний, не всегда удастся ограничиться двухбайтными токенами, и эффективность метода заметно снижается.

Алгоритм Хаффмана

- В основе этого алгоритма лежит кодирование не байтами, а битовыми группами.
- Перед началом кодирования производится частотный анализ кода документа и выявляется частота повтора каждого из встречающихся символов.
- Чем чаще встречается тот или иной символ, тем меньшим количеством битов он кодируется (соответственно, чем реже встречается символ, тем длиннее его кодовая битовая последовательность).
- Образующаяся в результате кодирования иерархическая структура прикладывается к сжатому документу в качестве таблицы соответствия.
- Например, 1 бит - буква А;
2 бита – буква О;
4 бита – буква Е и т.д.

- Используя 16 бит, можно закодировать до 256 различных символов; 20 бит — можно закодировать до 1024 лексических единиц (это могут быть не символы, а группы символов, слоги и даже слова).
- В связи с тем, что к сжатию архиву необходимо прикладывать таблицу соответствия, на файлах малых размеров алгоритм Хаффмана малоэффективен.
- Практика также показывает, что его эффективность зависит и от заданной предельной длины кода (размера словаря). В среднем, наиболее эффективными оказываются архивы с размером словаря от 512 до 1024 единиц (длина кода до 18-20 бит).

ГРАФИЧЕСКИЕ РЕДАКТОРЫ

Adobe Photoshop требует для своей работы профессиональных навыков, поскольку эта программа создана для профессионалов. Имеет большое количество функций и возможностей. Она является базовой программой, которую изучают все дизайнеры и художники.

Программа **Adobe Illustrator** - конкурент программы **Corel DRAW**. Это менее мощный графический редактор, нежели Corel. Его визуальное оформление и программный интерфейс делает ее похожей на Adobe Photoshop, только под векторную графику. С ней работают новички-любители. Она особо нигде не изучается.

Macromedia Flash - мощный графический редактор, предназначенный для создания векторных анимаций. Принцип создания анимации построен так: на каждом кадре анимации векторный объект имеет свои свойства, в этом и заключается основной принцип векторной анимации. Данная программа довольно широко используется при создании анимации для вэб-страничек, однако ее еще используют не только для дизайна, но и для создания целых роликов и даже мультфильмов.

3d Studio MAX - это мощный трехмерный графический редактор, один из самых лучших в мире. Он очень многофункционален. Если в векторной графике описание координат точек производится по двухмерной системе (x и y), то в трехмерной графике идет описание по трехмерной системе: x, y и z. Данная программа нашла широкое применение в дизайне и в игростроении, то есть при создании компьютерных игр.

ГРАФИЧЕСКИЕ РЕДАКТОРЫ

Maya - трехмерный редактор, основной конкурент 3d Studio MAX, однако он менее функционален. Программа рассчитана на любителя.

Corel Draw - изначально был задуман как универсальный векторный редактор, применяемый для решения абсолютно всех задач векторной графики. Продукт достаточно старый: первая версия вышла еще в 1989 году. Но широкую популярность приобрел Corel Draw 4, вышедший в начале 1994 года. В то время это был единственный пакет, который поддерживал ввод русского текста напрямую, безо всяких ухищрений. **Важное достоинство - кроссплатформенность** Corel Draw. Есть версии для Windows, OS/2, MAC, различных вариантов UNIX. Главный **недостаток** Corel Draw - это искажение цветов при экспорте в растровые СМΥК изображения. Экспорт в PostScript слишком часто производится некорректно, и в результате с вероятностью 90% такой файл завесит программу верстки полос или устройство вывода.

Macromedia FreeHand - является главным конкурентом Adobe Illustrator в области художественного рисования как на PC, так и на MAC. Причем во многих странах FreeHand отодвинул Illustrator на вторые роли. Он быстрее и удобнее в работе, к тому же поддерживает многое из того, что в Illustrator не предусмотрено.