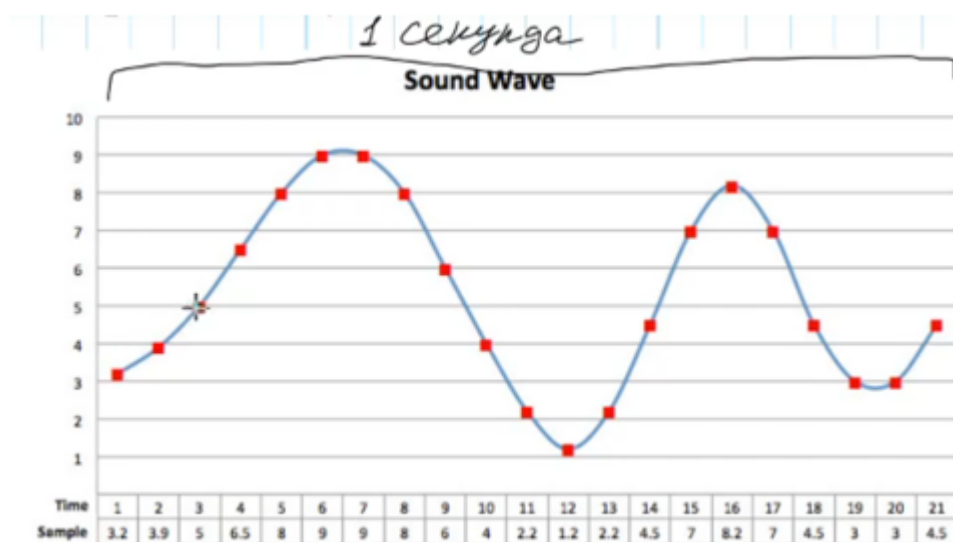


Конспект занятия

Сегодня посмотрим, как кодируется звук в компьютере. И это гораздо интереснее, чем текст или картинки. Если текст состоит из символов, то мы кодируем символы. Если растровые картинки состоят из пикселей, то мы кодируем пиксели.

Из чего же состоит звук?



На картинке представлена звуковая волна, изображенная условно. Цифры слева по вертикали - это сила сигнала, полученная от микрофона, которая идет в виде волны.

Представим, что вся волна на картинке - это ровно одна секунда звука. Точки - это определенные временные моменты, расположенные через равные промежутки времени. В данном случае временных моментов 21 штука. В реальном кодировании этих интервалов будет несколько десятков тысяч.

Каждый этот момент, каждая точка (значение сигнала в этот момент) записывается в память в виде двоичного кода, в виде одинакового количества бит.

То есть вместо непрерывной волны в компьютере сохраняются только эти точки. Но эти точки расположены так близко, их так много, что нам кажется, что сигнал непрерывный, хотя на самом деле это не так. На самом деле он состоит из маленьких точек, которые очень плотно стоят друг к другу. Как пиксели на картинках.

Итак, одна секунда разбивается на моменты через равные промежутки времени. Каждый отдельный момент записывается в виде двоичного кода какой-то одинаковой длины.

Примечание Джобса: звуковой файл содержит замеры уровня сигнала, сделанные с определенным интервалом. Временные интервалы обычно одинаковые для всего файла и показывают, как часто совершаются замеры. Тайм-моментом, очевидно, Алексей называет конкретный момент времени, в который сделан замер.

Далее моменты я буду называть уровнем звукового сигнала. «Запись момента звука» по тексту выглядит странно. На мой взгляд «Запись уровня звукового сигнала» читается проще и выглядит понятнее.

Частота дискретизации — это количество временных интервалов в одной секунде. Мы с вами будем её обозначать буквой v (ню). Как в физике, частота буквой v записывается. Частота дискретизации измеряется в герцах (Гц) и показывает, на какое количество одинаковых временных интервалов разбивается одна секунда.

Представим, что у нас одна конкретная секунда звука. Эту секунду разбиваем на 10 равных частей и берем через каждую десятую долю секунды значение уровня звукового сигнала в данный момент времени. Получается, что одна секунда превращается в 10 точек, которые располагались на оригинальной звуковой волне. Это и есть частота дискретизации - 10 раз в секунду. Мы секунду разбили на 10 частей и сделали 10 измерений – значений уровней звукового сигнала в данных точках.

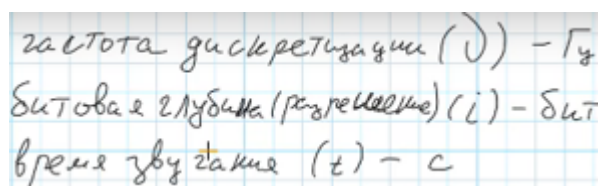
Мы можем взять другую частоту, например, 20 раз в секунду. Тогда одна секунда превратится в 20 точек. График будет точнее. Возьмем частоту 1000 Гц, будет 1000 точек. Одна секунда станет еще детальнее. То есть, чем больше мы берем частоту, тем более точной будет копия оригинальной звуковой волны.

То есть частота дискретизации – это то количество измерений уровня звукового сигнала, которые мы делаем за одну секунду. На картинке, например, 21 Гц, 21 измерение.

Каждый уровень звукового записывается каким-то одинаковым количеством бит. Это количество называется битовой глубиной. Глубина кодирования, разрешение, разрешение кодирования – синонимичные понятия. В задачах вы можете найти все эти названия, но обозначают они одно – длину каждого измерения в битах. Измеряется размер момента в битах. Мы будем обозначать его буквой i .

Если нужно найти размер одной секунды (сколько занимает ровно одна секунда звука), мы должны частоту дискретизации (количество измерений) умножить на размер каждого измерения. То есть количество бит для одной точки перемножаем с количеством точек и получаем размер одной секунды в битах.

Но это одна секунда. Звук-то у нас обычно длится дольше. Поэтому третий параметр, который нам понадобится – это *время звучания*. Собственно, мы будем записывать его буквой t . Измеряется оно в секундах.



частота дискретизации (f) – Гц
битовая глубина (разрешение) (i) – бит
время звучания (t) – с

Чтобы найти размер звукозаписи (размер звукового канала), нужно: размер одной секунды, умножить на количество секунд.

Звуковой канал – это очень интересная штука. Если мы слушаем музыку в наушниках или в колонках, то они всегда разделены на две части – левая и правая. И что самое интересное, звук в них отличается. Звук в левом наушнике и в правом наушнике может быть разный. За счёт этого у нас формируется иллюзия, допустим, объёма.

У нас есть положение источника звука, как он передвигается в пространстве. В принципе, мы с вами ориентируемся на слух за счёт разницы звуковых сигналов, которые попадают в левое и правое ухо.

Вы понимаете, где нахожусь я, где находится источник звука от меня, благодаря тому, что в левое и правое ухо звук доходит чуть с разной скоростью, чуть с разным временем. И мы это используем также при звукозаписи. Когда записывают звук, записывают его с двух разных устройств, с двух микрофонов. То есть записываются два звуковых потока параллельно. Есть поток левый, есть поток правый. И да, каждый из них будет занимать какое-то одинаковое количество места.

Отсюда получается, что четвертый параметр, который влияет на размер звука, это то, сколько в нем каналов (*число каналов звука*). Сколько одновременно параллельно потоков звуков записано. Будем обозначать буквой K .

В рамках ЕГЭ есть такие числа каналов:

- моно – это один канал.
- стерео – это два канала.
- квадро – это четыре канала.

Какое-то другое количество может быть просто записано числом.

Соответственно, если у нас стереозвук, значит размер звукового файла умножается на 2, потому что 2 параллельных звуковых потока записано. Если квадро, значит мы умножаем на 4, потому что у нас 4 звуковых потока параллельно.

Из этих четырёх множителей и складывается размер звукового файла. Итоговая формула у нас получается такая: число каналов мы умножаем на частоту дискретизации, умножаем на глубину кодирования, умножаем на время звучания. И это получается итоговый размер звука.

$K \cdot \Delta \cdot i \cdot t$ – итоговый размер

Задание 1.

Музыкальный фрагмент длительностью 5 минут был записан в формате квадро, частотой дискретизации 40 кГц и 16-битным разрешением. Сжатие данных не производилось. Определите объем получившегося файла в МБайтах. В качестве ответа укажите целую часть полученного числа.

Ссылка на видео-разбор: https://vk.com/video-205546952_456241168?t=13m50s

Решение

Нам нужно посчитать какой объем памяти будет занимать звук в МБайтах.

Мы знаем, что в звуке 4 канала. Частота дискретизации 40 кГц. Кило в данном случае тысяч раз в секунду, то есть 40 тысяч измерений в секунду. 16-битное разрешение или глубина кодирования. И длительность 5 минут, которую мы переводим в секунды ($5 \cdot 60 = 300$ секунд). Перемножив все 4 эти числа, получаем размер звука в битах.

$$\text{Размер звука} = 4 \cdot 40000 \cdot 16 \cdot 300 \text{ бит}$$

Так как нам нужно найти объем памяти в МБайтах, нам нужно размер звука в битах поделить на 2^{23} . Так как $1 \text{ Мбайт} = 2^{23} \text{ бит}$.

$$\text{Размер звука} = \frac{4 \cdot 40000 \cdot 16 \cdot 300}{2^{23}} = 91,55 \text{ Мбайт}$$

```
>>> 4*40000*16*300/2**23
91.552734375
```

В ответ надо записать целую часть полученного числа. Что такое целая часть? Собственно, 91 - это целая часть числа, а 55 (после запятой) - дробная.

Дробную часть отбрасываем, остается целая. Это называется округление вниз.

Поэтому в ответ записываем: 91.

Почему мы не округляем? Потому что просят найти целую часть полученного числа. Не ближайшее целое число, а целую часть полученного числа.

Как мне понять, когда округлять, когда отбрасывать? Читать формулировку. Если написано, укажите целую часть полученного числа, значит, отбрасываем дробную часть. Написано, укажите ближайшее целое число, значит, надо округлить. Написано, укажите минимальный целый объем, в который поместится данный звук. Тогда надо будет округлить в большую сторону.

Ответ: 91

Музыкальный фрагмент длительностью 5 минут был записан в формате квадр (четырёхканальная запись), частотой дискретизации 40 кГц и 16-ти битным разрешением. Сжатие данных не производилось. Определите объем получившегося файла в МБайтах. В качестве ответа укажите целую часть полученного числа.

5 мин. $5 \cdot 60 \text{ с.}$

$4 \cdot 40000 \cdot 16 \cdot 300 \text{ бит}$

$\frac{4 \cdot 40000 \cdot 16 \cdot 300}{2^{23}} = 91,55 \text{ Мбайт}$

целая часть + числа

Округление

1 Мбайт = 2^{20} бит

1 Мс = $1024 \text{ Кс} = 1024^2 \text{ Б} = 1024^2 \cdot 8 \text{ бит}$

Ответ: 91

Задание 2 (4005).

Ссылка на видео-разбор: https://vk.com/video-205546952_456241168?t=22m10s

Производится двухканальная (стерео) звукозапись с частотой дискретизации 44 кГц. Запись длится 5 минут 25 секунд, ее результаты записываются в файл без сжатия данных, причем каждый сигнал кодируется минимально возможным и

одинаковым количеством бит. Информационный объем полученного файла без учета заголовка не превышает 82 МБайт. Определите максимальную битовую глубину кодирования звука, которая могла быть использована в этой записи. В ответе запишите только число.

Решение.

Что нам дано? Нам дано всё, кроме битовой глубины, кроме размера каждого измерения звука. Давайте его найдем.

У нас два канала на стерео запись, частота дискретизации 44 кГц, то есть 44 тысячи раз в секунду происходит измерение. Мы не знаем битовую глубину, это i . Но знаем длительность 5 минут 25 секунд. То есть переводим это в секунды, получается 325 секунд ($5 \cdot 60 + 25$). Это размер в битах.

$$2 \cdot 44000 \cdot i \cdot 325 \text{ бит}$$

По условию задачи этот размер не превышает 82 МБайт. То есть, 82 — это верхняя граница размера звука. Сразу переведем в биты ($82 \cdot 2^{23}$).

$$2 \cdot 44000 \cdot i \cdot 325 \leq 82 \cdot 2^{23}$$

Получаем такое простенькое неравенство.

$$i \leq \frac{82 \cdot 2^{23}}{2 \cdot 44000 \cdot 325}$$

```
>>> 82*2**23/(2*44000*325)  
24.051253706293707
```

$i \leq 24,05$ бит - верхняя граница битовой глубины.

То есть размер каждого измерения не должен превышать величины 24,05 бит.

Максимальная битовая глубина - 24 бита. Это максимальное целое значение, которое подходит нам под это неравенство. Поэтому ответ получается 24.

Ответ: 24.

$$2 \cdot 44000 \cdot i \cdot 325 \text{ бит} \leq 82 \cdot 2^{23} \text{ бит}$$

$$i \leq \frac{82 \cdot 2^{23}}{2 \cdot 44000 \cdot 325}$$

$$i \leq 24,05 \text{ бит}$$

Ответ: 24

Задание 3.

Ссылка на видео – разбор: https://vk.com/video-205546952_456241168?t=26m25s

Производится запись музыкального фрагмента в формате стерео с частотой дискретизации 44 кГц, длительностью 2 минуты. Размер полученного файла 20625 Кбайт. Определите количество уровней громкости цифрового звука. В ответе укажите целое число.

Решение

Тут не написано, но пояснить надо, что нужно найти максимальное количество уровней громкости. Давайте подумаем, что это такое, уровни громкости. С чем связана эта характеристика? Подскажу, что это связано с тем, чего в задаче нет.

В задаче нет одной характеристики - битовой глубины. А количество уровней громкости связано именно с битовой глубиной.

Получается, это точно та же самая история, что с кодировкой текста.

Количество бит определяет, какое количество символов мы можем использовать в тексте. Количество бит определяет, какое количество цветов можно использовать в картинках.

И битовая глубина в звуке определяет максимальное количество уровней громкости. Мы можем записать, соответственно, чем их больше, тем получается точнее запись звука, тем она лучше.

Итак, нам надо найти битовую глубину, так же, как в прошлой задаче. У нас получается два канала с вами, потому что формат стерео, на 44000, умножаем на i , умножаем на длительность 2 минуты (120 секунд) и это равно 20625 Кбайт.

Переводим сразу в бит, то есть умножаем на 2^{13} бит.

$$2 \cdot 44000 \cdot i \cdot 120 = 20625 \cdot 2^{13}$$

$$i = \frac{20625 \cdot 2^{13}}{2 \cdot 44000 \cdot 120} = 16$$

```
>>> 20625*2**13/(2*44000*120)
16.0
>>>
```

Мы получили битовую глубину. Но вопрос про количество уровней громкости, которые можно в неё записать. Соответственно, тут работает то же самое правило, что с кодированием текста и цветов, что i бит кодируют 2^i (два в степени i) разных значений.

Отсюда получается в 16 бит можно записать 2^{16} бит уровней громкости.

2^{16} бит уровней громкости = 65536 – количество уровней громкости

```
>>> 2**16
65536
```

Обратите внимание, есть разница битовая глубина - это количество бит на каждое измерение. И есть количество уровней громкости - это что с помощью них кодируется. Количество различных значений, которое может принимать эта громкость, это 2 в степени 16.

Ответ: 65536

Производится запись музыкального фрагмента в формате стерео, с частотой дискретизации 44кГц, длительностью 2 минуты. Размер полученного файла – 20625 Кбайт. Определите количество уровней громкости цифрового звука. В ответе укажите целое число.

$1 \text{ Кбайт} = 2^{13} \text{ байт}$
 $i \text{ байт} \rightarrow 2^i \text{ значений}$

$$2 \cdot 44000 \cdot i \cdot 120 = 20625 \cdot 2^{13} \text{ байт}$$

$$i = \frac{20625 \cdot 2^{13}}{2 \cdot 44000 \cdot 120}$$

$$i = 16 \text{ байт} \rightarrow 2^{16} \text{ уровней громкости}$$

65536

Задание 4 (6715, ЕГЭ-2023)

Производилась четырёхканальная (квадро) звукозапись с частотой дискретизации 192 кГц и 16-битным разрешением. В результате был получен файл размером 967 Мбайт, без учёта размера заголовка и без сжатия данных. Определите длительность звукозаписи (в минутах). В качестве ответа укажите ближайшее к полученному времени записи целое число.

Ссылка на видео-разбор: https://vk.com/video-205546952_456241168?t=34m35s

Примечание

Опять же, я повторяю, что приписки про заголовок и сжатие не влияют на задачу. То есть это условно вам написали просто для красного словца. По факту здесь нет ни заголовков, ни сжатия. Раз их по факту в числах нигде нет, значит и учитывать не надо.

Решение

Значит ответ нужно округлить по правилам математики.

Посчитаем длительность. У нас получается 4 канала на 192 000 частоту килогерца, на 16-битное разрешение, на время t , которое мы не знаем. И файл получился размером 967 мегабайт, то есть умножить на 2^{23} .

$$4 \cdot 192000 \cdot 16 \cdot t = 967 \cdot 2^{23}$$

$$t = \frac{967 \cdot 2^{23}}{4 \cdot 192000 \cdot 16}$$

```
>>> 967*2**23/(4*192000*16)  
660.1386666666667
```

Сколько минут? Дополнительно разделим еще на 60. Получится 11 с очень маленьким хвостиком.

Собственно понятно, что ближайшее целое число это 11 минут.

$$4 \cdot 192000 \cdot 16 \cdot t = 967 \cdot 2^{23} \text{ снт}$$
$$t = \frac{967 \cdot 2^{23}}{4 \cdot 192000 \cdot 16} = \frac{660.138}{60} = 11 \text{ мин}$$

Ответ: 11

Задание 5

Голосовое сообщение длительностью 3 минуты было закодировано в формате стерео с разрешением 24 бита и частотой дискретизации 54 000 измерений в секунду и передано по каналу связи. Сжатия данных не производилось.

Пропускная способность канала связи равна 3200 бит/с. Определите, сколько

минут необходимо для передачи голосового сообщения. В ответе запишите только целое число.

Ссылка на видео-разбор с таймингом: https://vk.com/video-205546952_456241168?t=41m

Решение

Начнем с нахождения объема голосового сообщения. У нас получается два канала, потому что стереозвук, частота 54000 Гц, 24 битовая глубина разрешения и длительность 3 минуты, то есть длительность голосового сообщения 180 секунд.

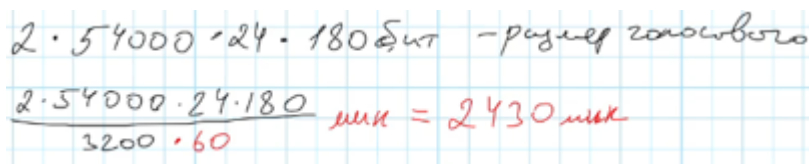
$$2 \cdot 54000 \cdot 24 \cdot 180$$

Скорость передачи 3200 бит/сек. Ой, как медленно. Ну, значит, время передачи этого сообщения с такой скоростью равна, соответственно, размер изображения поделить на 3200. Время передачи получится в секундах, ответ надо дать в минутах, поэтому я сразу еще поделю на 60.

$$\frac{2 \cdot 54000 \cdot 24 \cdot 180}{3200 \cdot 60}$$

```
>>> 2*54000*24*180/(3200*60)
2430.0
```

Собственно, получается 2430 минут.


$$\frac{2 \cdot 54000 \cdot 24 \cdot 180}{3200 \cdot 60} \text{ мин} = 2430 \text{ мин}$$

Задача №6

Музыкальный альбом записан в формате стерео с частотой дискретизации 48 кГц и разрешением 34 бит без использования сжатия. В альбоме 13 треков общей длительностью 42 минуты 20 секунд. Каждый трек содержит заголовок размером 110 Кбайт. Сколько секунд потребуется для скачивания альбома по каналу со скоростью передачи данных 314572800 бит/с? В ответе укажите целую часть числа.

Ссылка на видео-разбор с таймингом: https://vk.com/video-205546952_456241168?t=48m45s

Решение

У нас два канала. Частота 48000 Гц. Глубина 34 бита. Длительность 42 минуты 20 секунд = $42 \cdot 60 + 20 = 2540$ секунд.

Но кроме звука есть заголовки. Каждый заголовок каждому треку добавляет по 110 Кбайт.

То есть мы еще прибавляем 13 раз по 110 Кбайт.

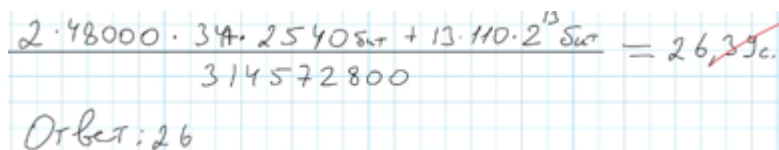
$$2 \cdot 48000 \cdot 34 \cdot 2540 + 13 \cdot 110 \cdot 2^{13}$$

И это всё передаётся по каналу связи со скоростью 314572800 бит/сек. И надо поделить размер альбома на эту скорость.

$$\frac{2 \cdot 48000 \cdot 34 \cdot 2540 + 13 \cdot 110 \cdot 2^{13}}{314572800}$$

```
>>> (2*48000*34*2540 + 13*110*2**13)/314_572_800  
26.392220052083335
```

В ответ запишите целую часть числа, то есть мы отбрасываем дробную часть и получаем ответ 26.



2 · 48000 · 34 · 2540 + 13 · 110 · 2¹³
314572800 = 26.39
Ответ: 26

Задача №7

Музыкальный фрагмент был записан в формате моно, оцифрован и сохранён в виде файла без использования сжатия данных. Размер полученного файла – 72 Мбайт. Затем тот же музыкальный фрагмент был записан повторно в формате стерео (двухканальная запись) и оцифрован с разрешением в 3 раза выше и частотой дискретизации в 4,5 раза меньше, чем в первый раз. Сжатие данных не производилось. Укажите размер файла в Мбайт, полученного при повторной записи.

Ссылка на видео-разбор с таймингом: https://vk.com/video-205546952_456241168?t=56m20s

Решение

Было	Стало	Изменение
1 канал (моно)	2 канала (стерео)	*2
i	3i	*3
ν	$\frac{\nu}{4,5}$:4,5
72Мбайт	$\frac{72 \cdot 2 \cdot 3}{4,5} = 96$ Мбайт	

Задача №8

Музыкальный фрагмент был оцифрован и записан в виде файла без использования сжатия данных. Получившийся файл был передан в город А по каналу связи за 120 секунд. Затем тот же музыкальный фрагмент был оцифрован повторно с разрешением в 3 раза выше и частотой дискретизации в 1,5 раза меньше, чем в первый раз. Сжатие данных не производилось. Полученный файл был передан в город Б; пропускная способность канала связи с городом Б в 4 раза выше, чем канала связи с городом А. Сколько секунд длилась передача файла в город Б?

Ссылка на видео-разбор с таймкодом: https://vk.com/video-205546952_456241168?t=1h3m

Решение

Было	Стало
Разрешение: i	Разрешение: $3i$
Частота дискретизации: ν	Частота дискретизации: $\frac{\nu}{1,5}$
Размер: X Мбайт	Размер: $\frac{X \cdot 3}{1,5} = 2X$

Значит после изменения характеристик файл станет в 2 раза больше.

Передача в город А.

Оригинальный файл: 120 сек, измененный файл: $120 \cdot 2 = 240$ сек

Скорость и время зависят обратно пропорционально. Чем больше скорость, тем меньше время. Если скорость стала в 4 раза выше, то время стало в 4 раза

Отсюда получается, что измененный файл будет передаваться $\frac{240}{4} = 60$ сек

Задача №9

Музыкальный фрагмент был записан в формате стерео (двухканальная запись), оцифрован и сохранён в виде файла без использования сжатия данных. Размер полученного файла – 45 Мбайт. Затем тот же музыкальный фрагмент был записан повторно в формате моно и оцифрован с разрешением в 5 раз выше и частотой дискретизации в 4,5 раз меньше, чем в первый раз. При этом производилось сжатие данных, объем сжатого фрагмента стал равен 60% от исходного. Укажите размер файла в Мбайт, полученного при повторной записи. В ответе запишите только целое число, единицу измерения писать не нужно.

Ссылка на видео-разбор с таймингом: https://vk.com/video-205546952_456241168?t=1h10m30s

Решение

Было	Стало	Изменение размера файла
2 канала	1 канал	:2
i	5i	*5
ν	$\frac{\nu}{4,5}$:4.5
45 Мбайт	$45 \cdot \frac{5}{2 \cdot 4,5} = 25$ Мбайт	

Следовательно, до сжатия преобразованный звуковой файл занимал 25 Мбайт.

Сжали запись и она стала занимать 60% от исходного.

$$25 \cdot 0.6 = 15 \text{ Мбайт}$$

Задача №10

Для мультипликационного фильма видеоряд с частотой 60 кадров в секунду и звуковая восьмиканальная дорожка записываются отдельно. Для хранения на сервере видео преобразуют так, что частота уменьшается до 30 кадров в секунду, а количество пикселей уменьшается в 4 раза. Звук перезаписывается в формате стерео с уменьшением частоты дискретизации и глубины кодирования в 2 раза. Другие преобразования и иные методы сжатия не используются. После преобразования 1 минута видеоряда в среднем занимает 1,5 Мегабайта, а 1 минута звуковой дорожки – 512 Килобайт. Сколько Мбайт в среднем занимают 10 минут исходного видеоряда и звуковой дорожки вместе?

Ссылка на видео-разбор с таймингом: https://vk.com/video-205546952_456241168?t=1h15m30s

Решение

Что мы знаем про видео? Что было 60 кадров в секунду, стало 30 кадров в секунду. За счет уменьшения частоты кадров видео в 2 раза стало меньше.

Что произошло с пикселями? Пикселей стало в 4 раза меньше. То есть, отсюда и размер тоже в 4 раза уменьшается. Итого размер видео стал в 8 раз меньше.

Звук давайте посмотрим. У было 8 каналов, стало 2 канала. В 4 раза меньше. Частота дискретизации в 2 раза меньше. Битовая глубина уменьшилась в 2 раза. Соответственно, отсюда, получается, он размер звуковой записи уменьшился в 16 раз по сравнению с оригинальной записью.

Видео		Звук	
Было	Стало	Было	Стало
60 к/с	30 к/с : 2	8 каналов	2 канала : 4
мксели	: 4	частота дискретизации	: 2
	<hr/>		<hr/>
	+ : 8		: 16

Одна минута сжатого видеоряда - это 1,5 мегабайта. Размер в 8 раз меньше исходного. Значит размер минуты оригинального видеоряда $8 \cdot 1.5 = 12$ Мбайт

1 минута сжатой звуковой дорожки 512 Кбайт. Было в 16 раз больше. Значит размер исходной звуковой дорожки $512 \cdot 16$ Кбайт = 8 Мбайт

А нас спрашивают, сколько мегабайт в среднем занимают 10 минут исходного видеоряда и звуковой дорожки вместе.

$$10 \cdot (12 + 8) = 200 \text{ Мбайт}$$

12 МБ	1,5 МБ : 8	8 МБ	512 КБ : 16
$10 \cdot (12 \text{ МБ} + 8 \text{ МБ}) = 200 \text{ МБ}$			