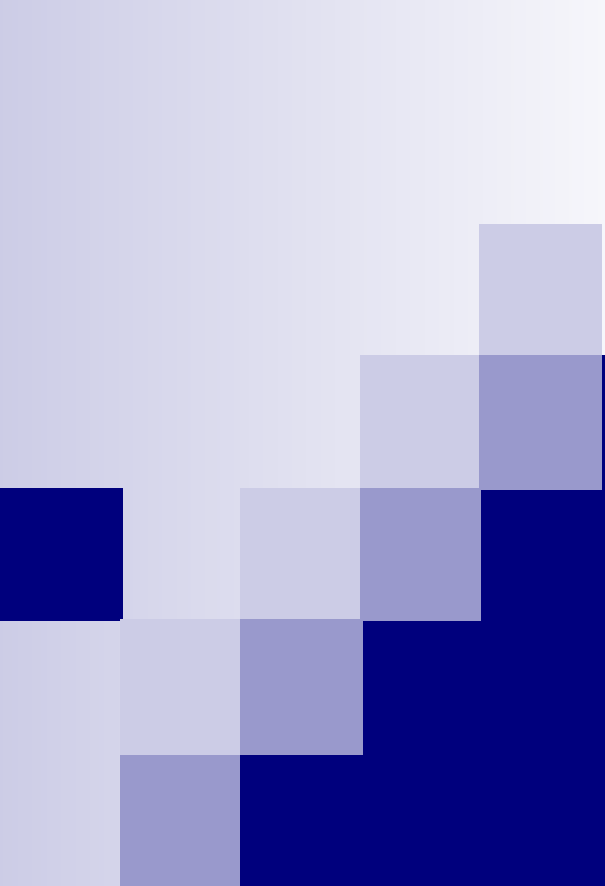




Алгоритмы сжатия видеоизображений



Введение

Основной сложностью при работе с видео являются большие объемы дискового пространства, необходимого для хранения даже небольших фрагментов.

COMPRESSION

COMPRESSION

COMPRESSION

COMPRESSION

COMPRESSION

COMPRESSION

COMPRESSION



10 ч музыки

несколько тыс.
фотографий

полчаса видео

Видео «телевизионного» формата 720×576 пикселей 25 кадров в секунду в системе RGB требует потока данных примерно в 240 Мбит/с (т. е. 1.8 Гбит/мин).

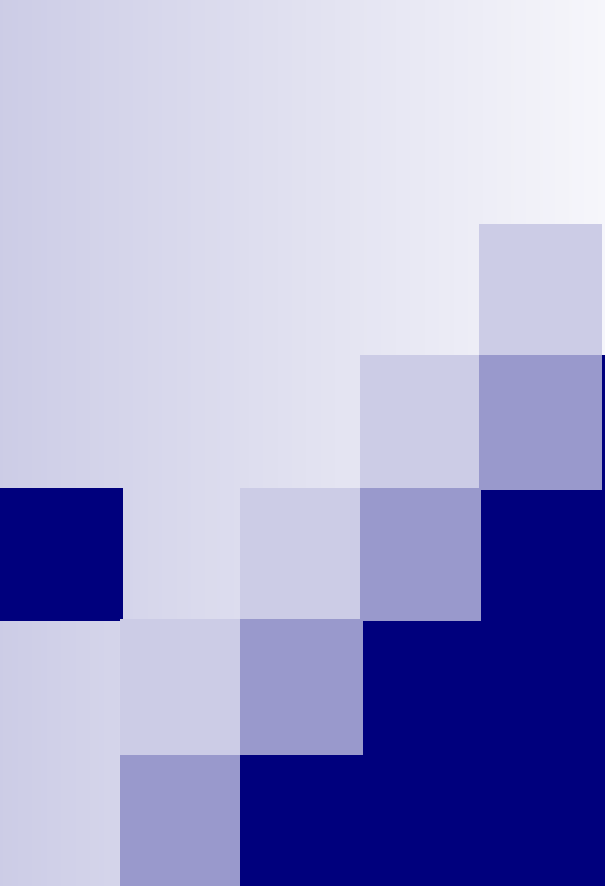
Правила сжатия видеоданных вырабатываются на
основе модели восприятия человеком
видеоизображений

HVS - Human Visual Sense



Избыточность изображения согласно HVS
определяется по **трем основным критериям**:

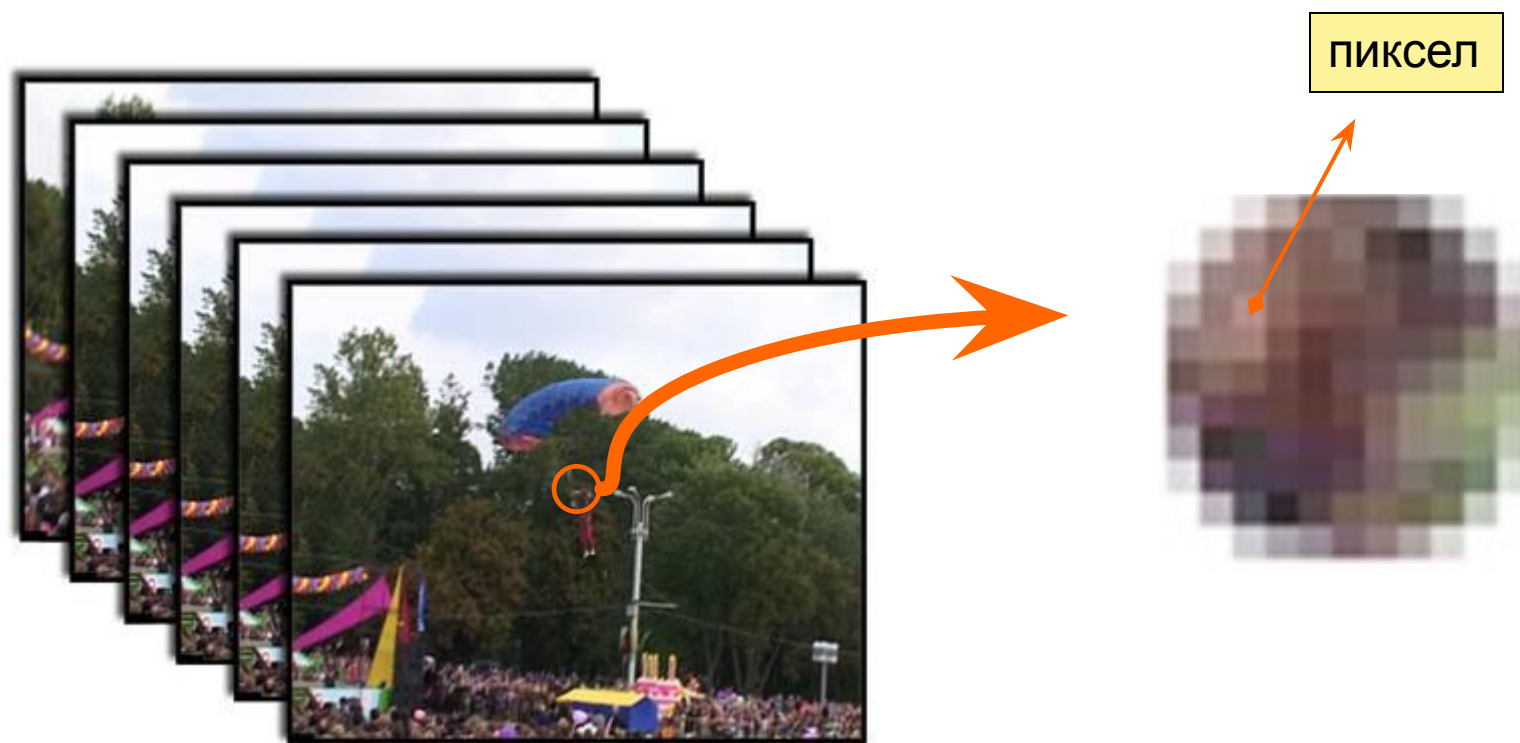
- 
1. **Невидимые человеческим глазом детали изображения** – места гашения по вертикали и горизонтали. Удаление этой информации вообще никак не сказывается на изображении.
 2. **Статистическая избыточность**. Подразделяется на пространственную и временную. Под *пространственной* избыточностью понимаются участки изображения, на которые смежные пикселы практически одинаковы. Под *временной* – не изменяемые во времени фрагменты изображения.
 3. **Избыточность по цвету и яркости** - рассчитывается исходя из ограниченной чувствительности человека к небольшим изменениям цветов и яркости деталей изображения.



Основные понятия

Цифровое видео - это последовательность кадров, в которой каждый кадр рассматривается как набор отсчетов аналогового изображения.

Отдельные отсчеты называются элементами изображения или **пикселями**.

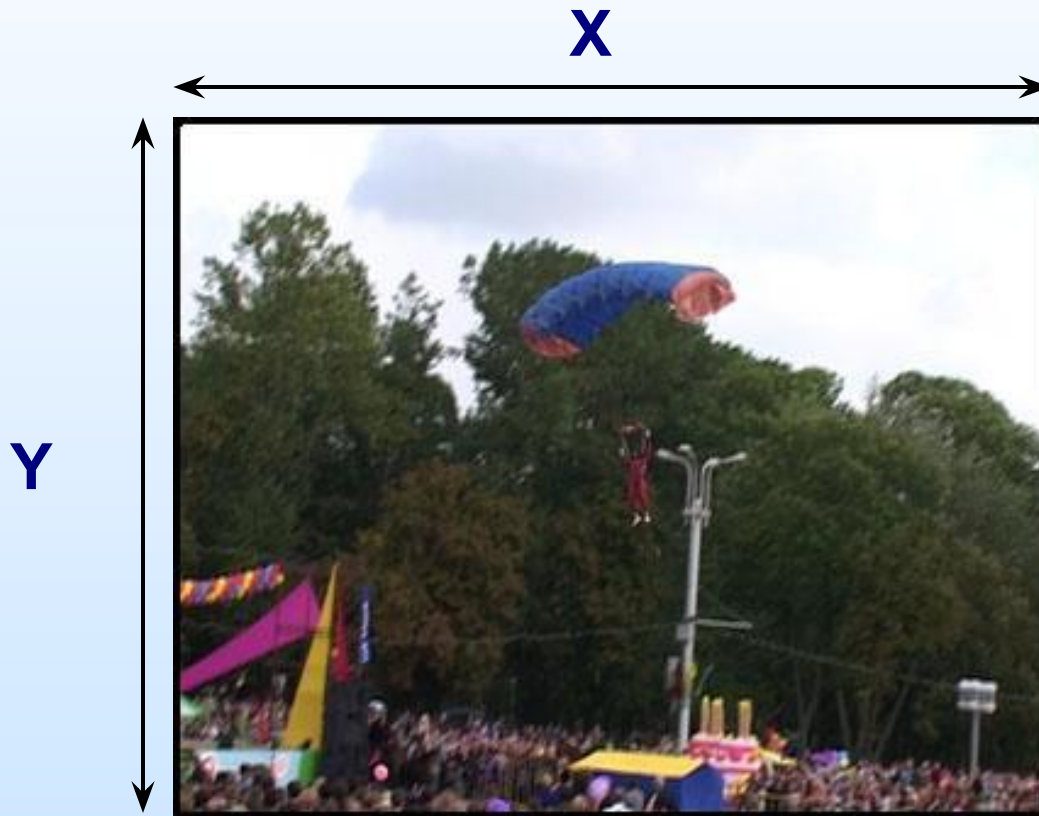


Видеопоток характеризуется *тремя основными факторами*:

- разрешением
- частотой кадров
- системой представления цветов



Разрешение определяется горизонтальным и вертикальным размерами кадра в пикселах



$X \times Y$

Из телевизионных стандартов пришли разрешения:

720×576



640×480



Частоты: ► 25 (PAL или SECAM), ► 30 (NTSC) кадров в секунду.

Существуют также стандарты с низкими разрешениями:

► CIF – Common Interchange Format ► QCIF – Quartered Common Interchange Format



352×288

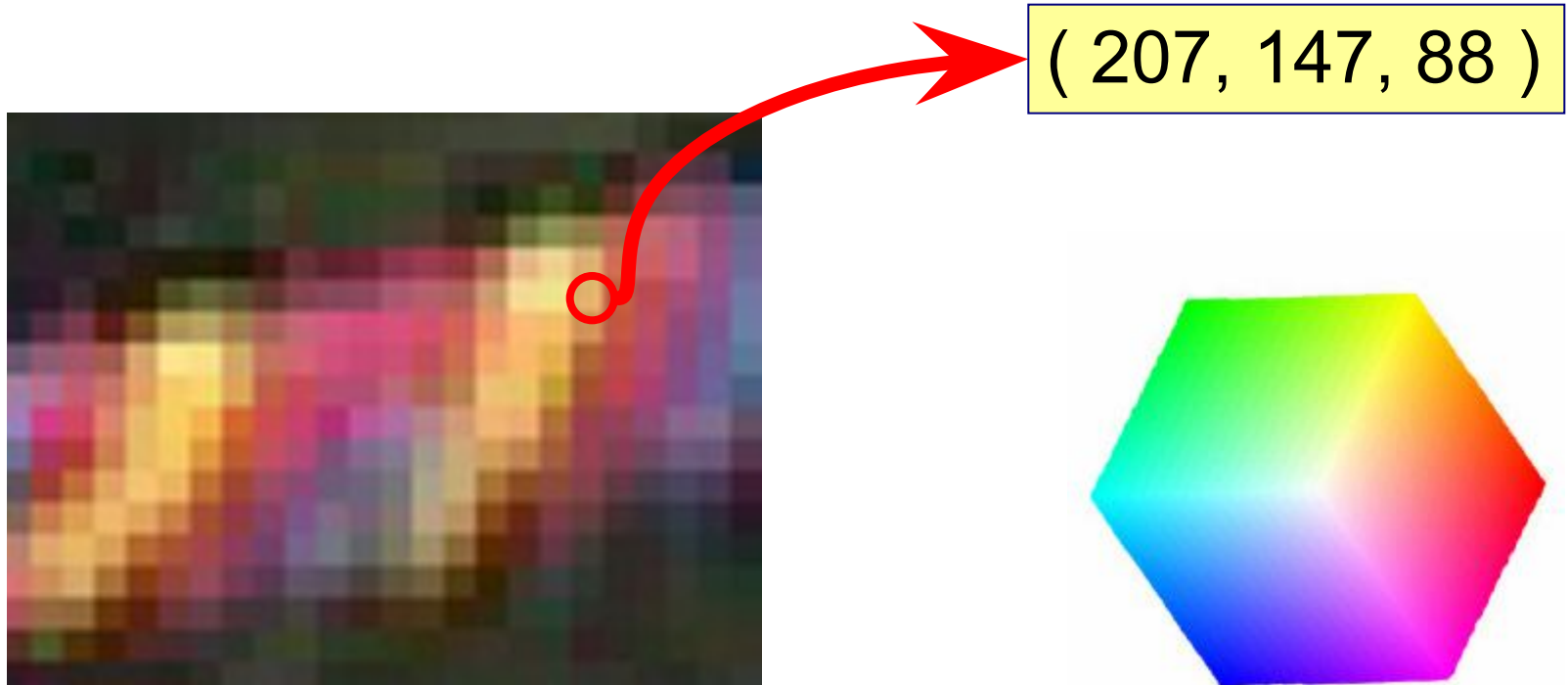


176×144

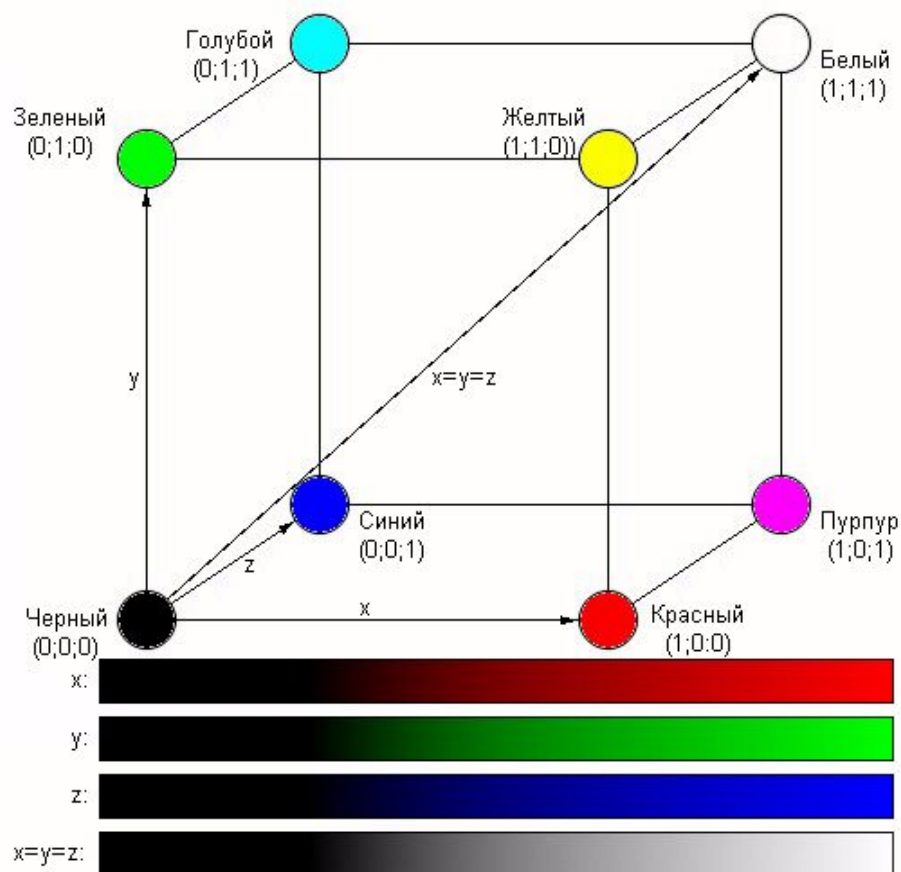
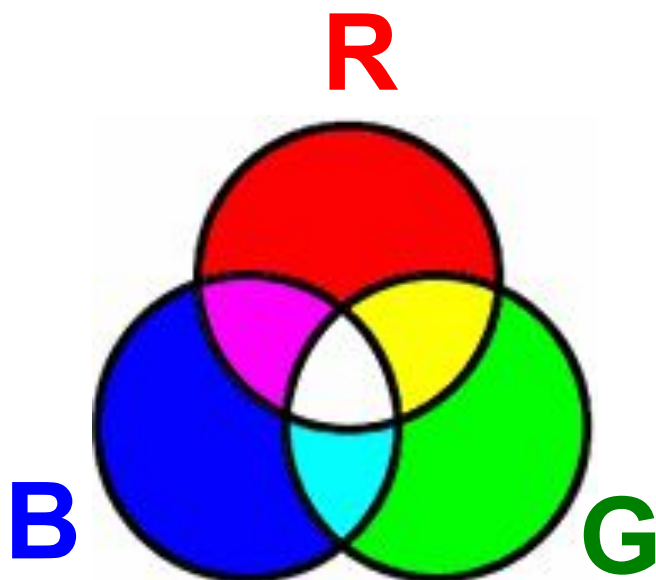
Поскольку CIF и QCIF ориентированы на крайне небольшие потоки,

с ними работают на частотах от 5 до 30 кадров в секунду.

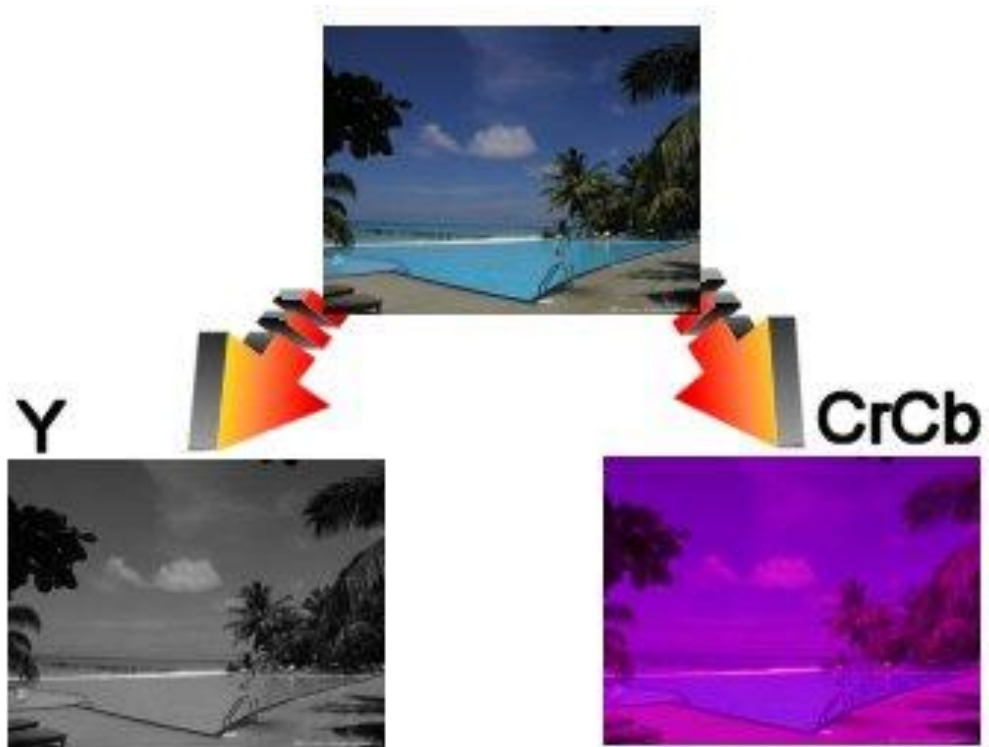
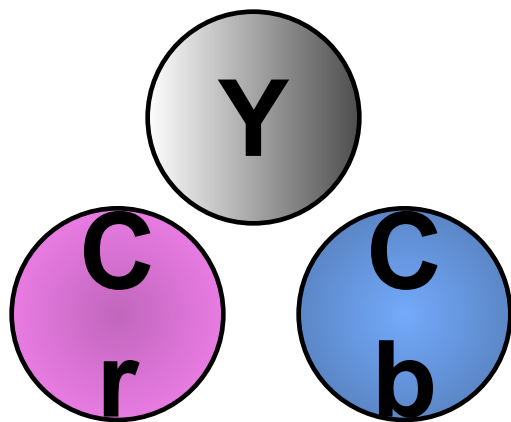
При использовании некой системы цветопредставления каждый пиксел представляет собой запись (структуру), полями которой являются компоненты цвета.



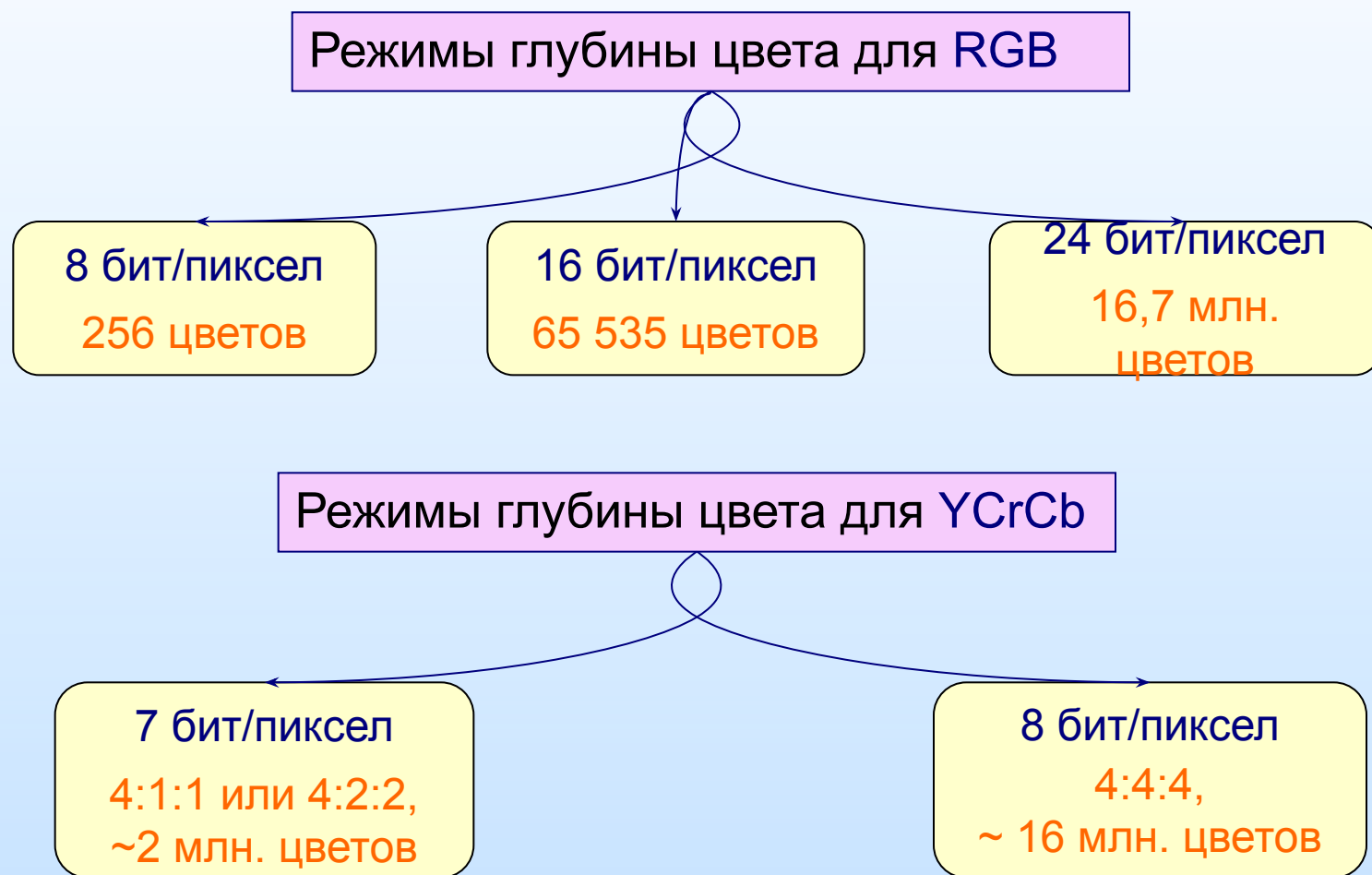
Самой распространенной является система **RGB**, в которой цвет представлен значениями интенсивности красной (**R**), зеленой (**G**) и синей (**B**) компонент.



Особенностям человеческого зрения наиболее соответствует другая цветовая система – **YCrCb**.
В ней **Y** — яркостная составляющая, а **Cr**, **Cb** — компоненты, отвечающие за цвет (хроматический красный и хроматический синий).



Каждая из моделей RGB и YCrCb может быть представлена разными уровнями *глубины цвета* (максимального количества цветов).



Требования, предъявляемые к алгоритму сжатия:

- высокая степень компрессии
- высокое качество видеоизображений
- высокая скорость компрессии
- высокая скорость декомпрессии
- произвольный доступ
- быстрый поиск вперед/назад
- показ кадров фильма в обратном направлении
- аудиовизуальная синхронизация
- устойчивость к ошибкам
- редактируемость
- масштабируемость
- небольшая стоимость аппаратной реализации

Носители информации, на которые ориентированы алгоритмы сжатия:

DVD-ROM



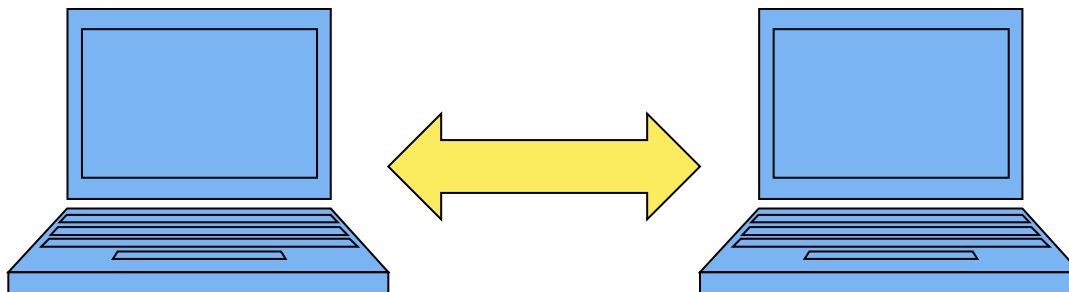
CD-ROM



жесткий диск



компьютерные сети



```
graph TD; A[Программное обеспечение, использующее видеокомпрессию] --> B[Симметричное]; A --> C[Асимметричное];
```

Программное обеспечение,
использующее видеокомпрессию

Симметричное

Предъявляет одинаково жесткие требования на время, память и другие ресурсы как при кодировании, так и при декодировании.

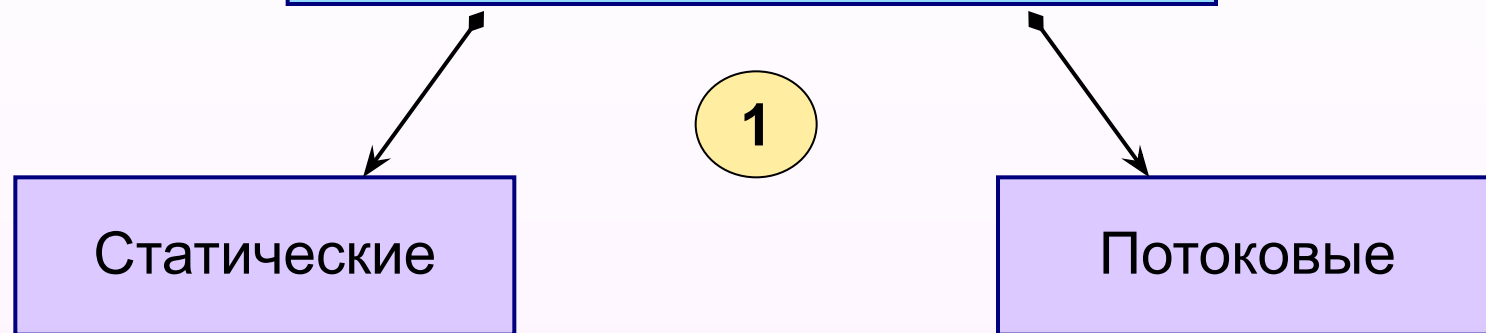
(Видеопочта, видеотелефон, видеоконференции, редактирование и подготовка видеоматериалов)

Асимметричное

Предъявляет серьезные требования к декодеру (по времени и памяти), но для него безразличны затраты ресурсов при кодировании.

(Мультимедиа-энциклопедии, путеводители, справочники, игры и просто фильмы)

Алгоритмы сжатия



Работают с каждым из кадров отдельно.

Сжатие достигается с помощью методов обработки изображений, при этом может обрабатываться либо все изображение целиком, либо разбиваться на отдельные блоки.

(Алгоритмы групп Wavelet и Jpeg.)

Работают с последовательностями кадров (поток).

Учитывается тот факт, что, близкорасположенные кадры не сильно отличаются друг от друга и сжатие достигается за счет кодирования лишь разницы между кадрами.

(Алгоритмы MPEG и MJPEG.)

Алгоритмы сжатия

2

Сжатие без потерь
данных

Полученное после декомпрессии изображение в точности (побитно) совпадает с оригиналом.

Сжатие с потерей
данных

Полученное после декомпрессии изображение и оригинал побитно не совпадают.

Большинство алгоритмов, распространенных в настоящее время, используют сжатие с потерей данных.

Сжатие с потерей данных

```
graph TD; A[Сжатие с потерей данных] --> B[Без заметных потерь с точки зрения восприятия]; A --> C[Сжатие с естественной потерей качества]; A --> D[Сжатие с неестественными потерями качества];
```

Без заметных потерь с точки зрения восприятия

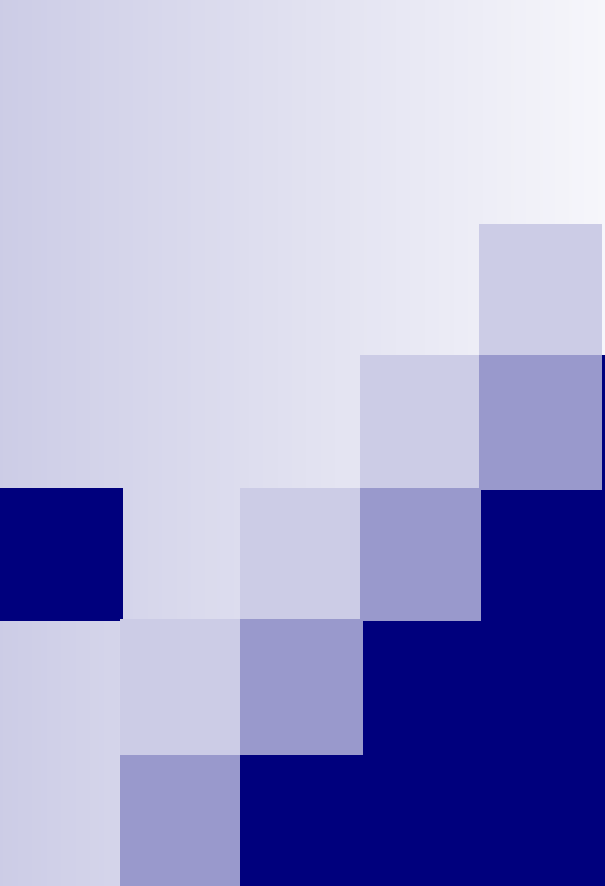
Данные после декомпрессии побитно не совпадают с исходными, однако из-за особенностей восприятия человек не способен отличить разницу на статическом и тем более на «живом» изображении.

Сжатие с естественной потерей качества

Характеризуется появлением воспринимаемых глазом, но незначительных искажений изображения: уменьшение детализации сцены, размытость вблизи резких границ и т. п. Однако они мало влияют на процесс зрительного восприятия картинки.

Сжатие с неестественными потерями качества

Низкое качество сжатия, в значительной степени искажающее изображение и вносящее в него искусственные (не существующие в оригинале) детали сцены. Нарушение самых важных характеристик изображения - контуров.



Обзор стандартов MPEG



В 1988 г. в рамках **Международной организации по стандартизации (ISO)** начала работу группа **MPEG (Moving Pictures Experts Group)** – группа экспертов в области цифрового видео (ISO-IEC/JTC1/SC2/WG11/MPEG)

Группа работала в трёх направлениях:

- **MPEG-Video** - сжатие видеосигнала в поток со скоростью до 1.5 Мбит/с,
- **MPEG-Audio** - сжатие звука до 64, 128 или 192 Кбит/с на канал,
- **MPEG-System** - синхронизация видео- и аудиопотоков.

Характеристики MPEG-1

Поток, разрешение: 1.5 Мбит/с, 352x240x30, 352x288x25.

Плюсы: сравнительно прост в аппаратной реализации, содержит преобразования, поддерживаемые на аппаратном уровне большим количеством видеокарт.

Минусы: невысокая степень сжатия, малая гибкость формата.

Характеристики MPEG-2

Поток, разрешение: 3 – 15 Мбит/с, универсальный.

Плюсы: поддержка звуковых стандартов Dolby Digital5.1, DTS, высокая универсальность, сравнительная простота аппаратной реализации.

Минусы: недостаточная на сегодня степень сжатия, недостаточная гибкость формата.

Характеристики MPEG-4

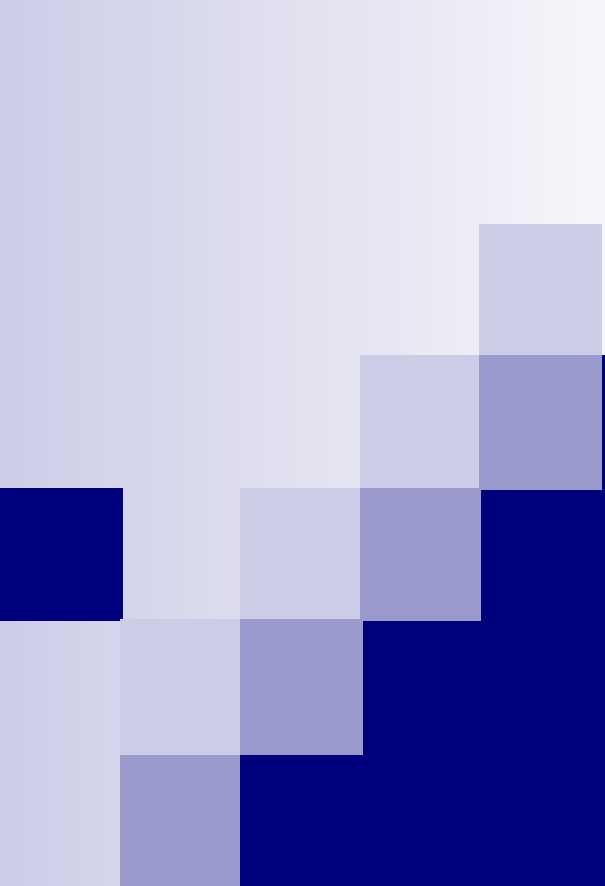
Поток, разрешение: 0,0048 – 20 Мбит/с, поддерживаются все основные стандарты видеопотоков.

Плюсы: поддержка прогрессивных звуковых стандартов, высокая степень универсальности, поддержка новых технологий (различные виды синтеза звука и изображения).

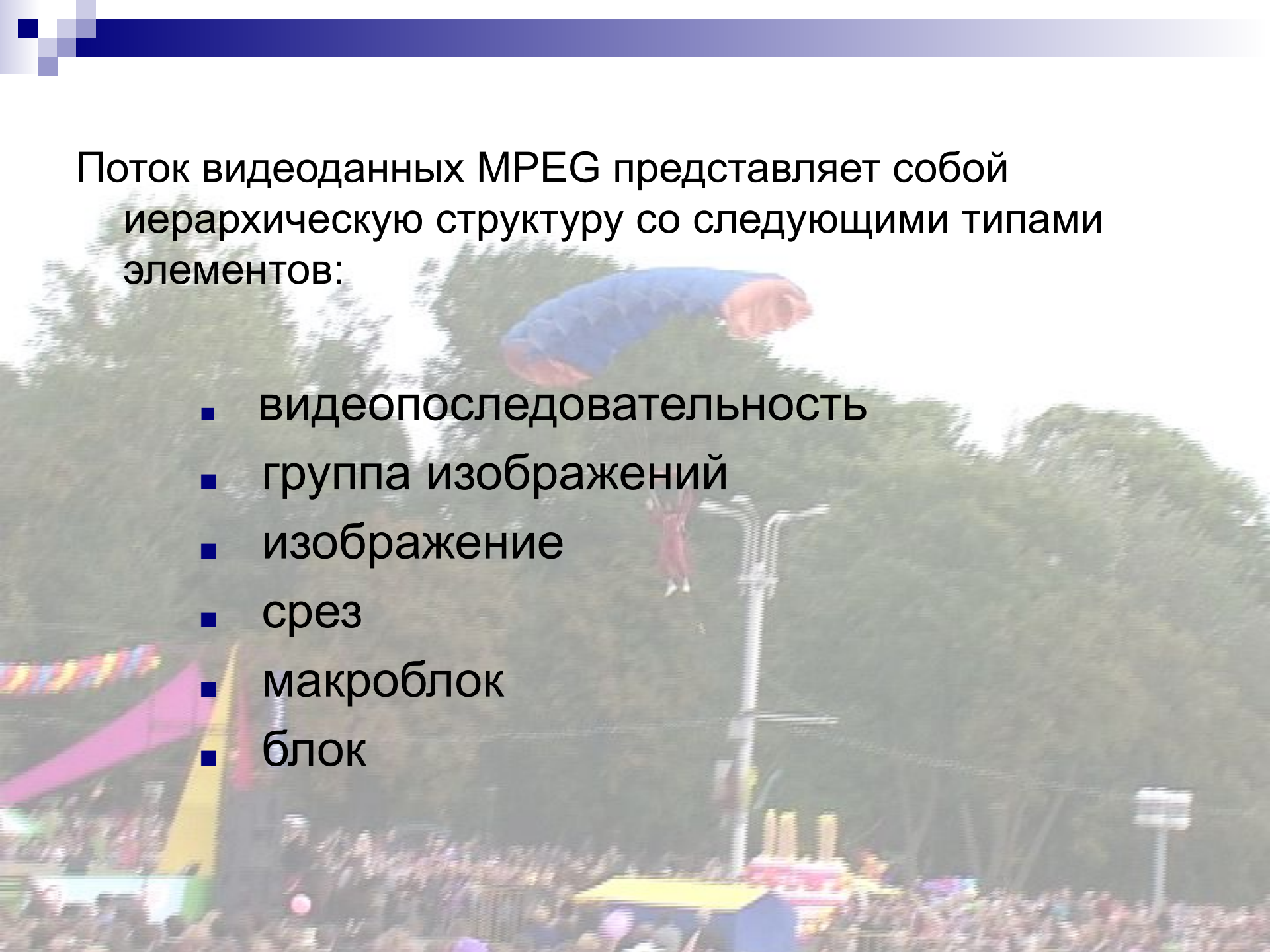
Минусы: высокая сложность реализации.

Сравнение стандартов

Название	Годы	Разрешения и поток	Аудио	Применение
MPEG-1	1992	352x240x30, 352x288x25, 1.5 Мбит/с	MPEG-1Layer II	Video CD первого поколения
MPEG-2	1995	Универсальный, 3 – 15 Мбит/с	MPEG-1Layer II, Dolby Digital5.1, DTS	DVD
MPEG-3 не принят	1993- 1995	Телевидение высокой четкости, 20 – 40 Мбит/с		HDVD
MPEG-4	1999	Универсальный, 0,0048 - 20 Мбит/с	MPEG-1Layer II, MPEG-1Layer III, Dolby Digital5.1, DTS	Video CD второго поколения



Базовые технологии сжатия видеоданных

A background image showing a person skydiving with a blue and red parachute. The skydiver is in the center, suspended in the air. Below them is a large crowd of people gathered on the ground, and a tall metal structure, possibly a tower or part of the event's infrastructure, is visible. The scene is outdoors with trees in the background.

Поток видеоданных MPEG представляет собой иерархическую структуру со следующими типами элементов:

- видеопоследовательность
- группа изображений
- изображение
- срез
- макроблок
- блок

Блок

Блок - наименьший синтаксический элемент структуры видеоданных MPEG. Блоки имеют размер 8x8 пикселей.

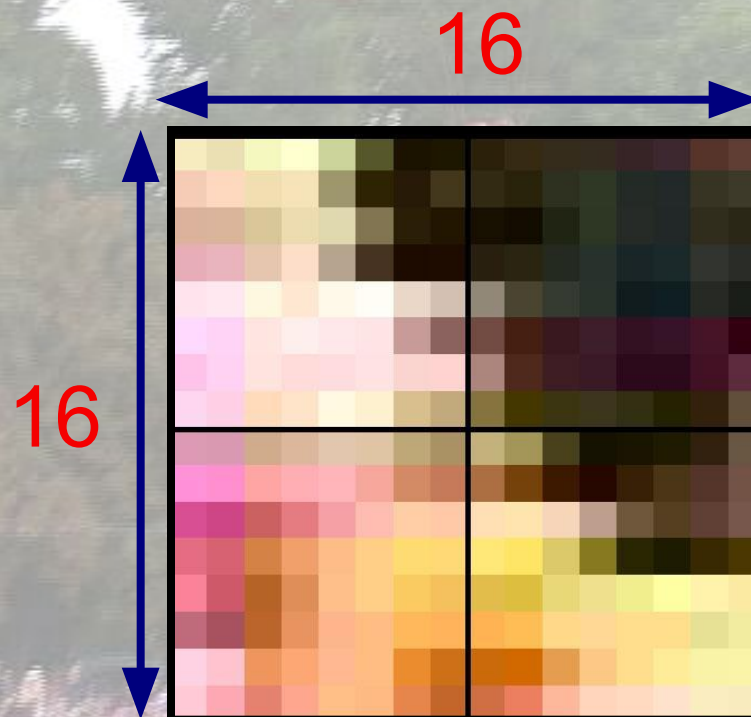
Они являются основными элементами для DCT (Discrete Cosine Transform - дискретное косинусное преобразование) кодирования.



Макроблок

Макроблок – это область, имеющая размер 16x16 пикселей.

Макроблок складывается из блоков размером 8x8 элементов изображения.

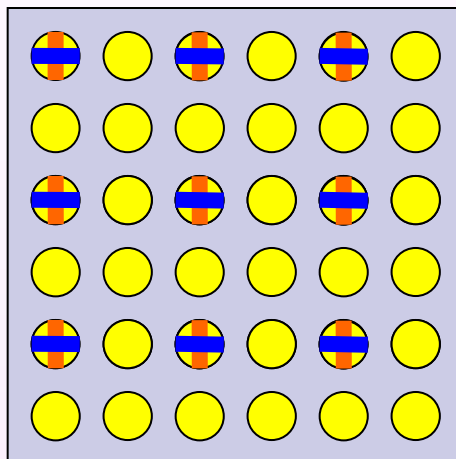


MPEG использует цветовую схему YCbCr. Плоскости Y, Cr, Cb кодируются с разным разрешением. Соотношение между количеством отсчётов яркости и цветности определяется форматом дискретизации.

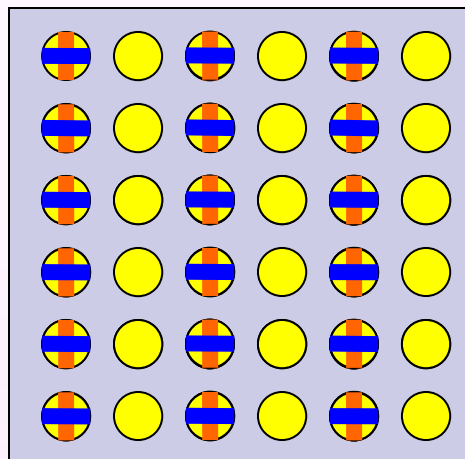
Формат дискретизации	Отношение разрешений Cb/Y (Cr/Y) по горизонтали	Отношение разрешений Cb/Y (Cr/Y) по вертикали
4:4:4	1:1	1:1
4:2:2	1:2	1:1
4:2:0	1:2	1:2
4:1:1	1:4	1:1
4:1:0	1:4	1:4

Форматы дискретизации

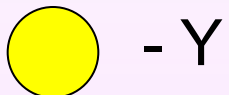
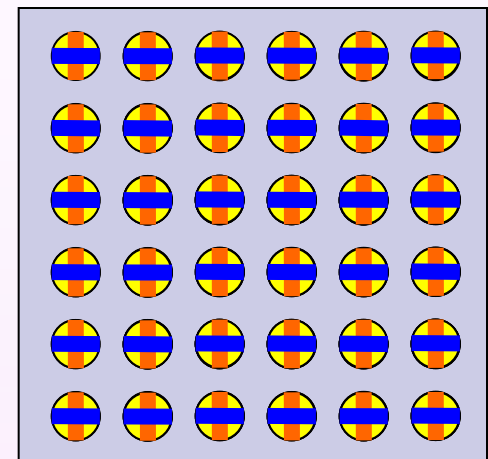
4:2:0



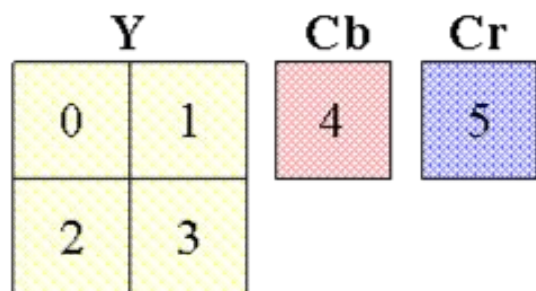
4:2:2



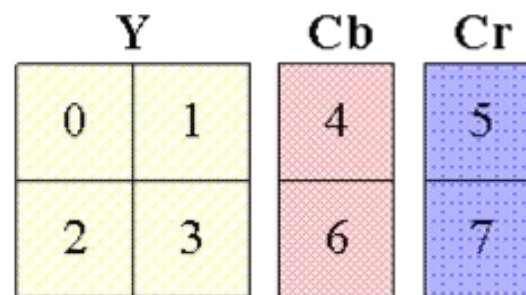
4:4:4



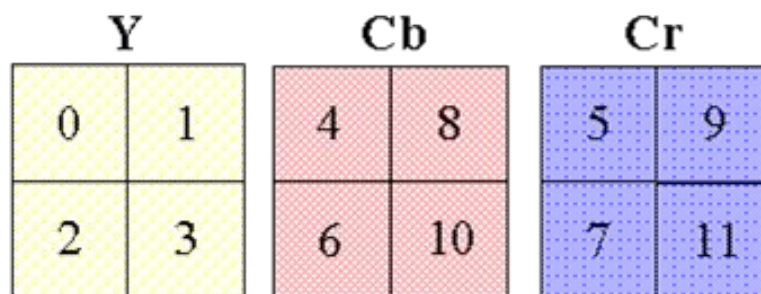
Число блоков для различных форматов



4 : 2 : 0



4 : 2 : 2



4 : 4 : 4

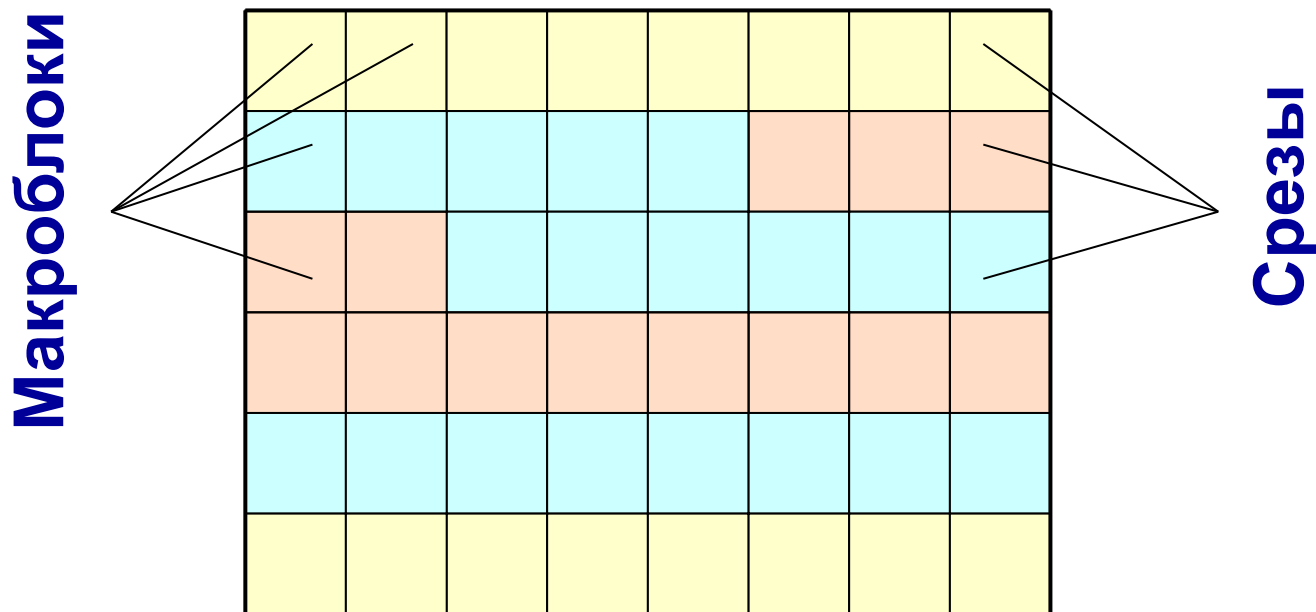
Срез

Срез - это ряд или неполный ряд макроблоков.

Макроблоки



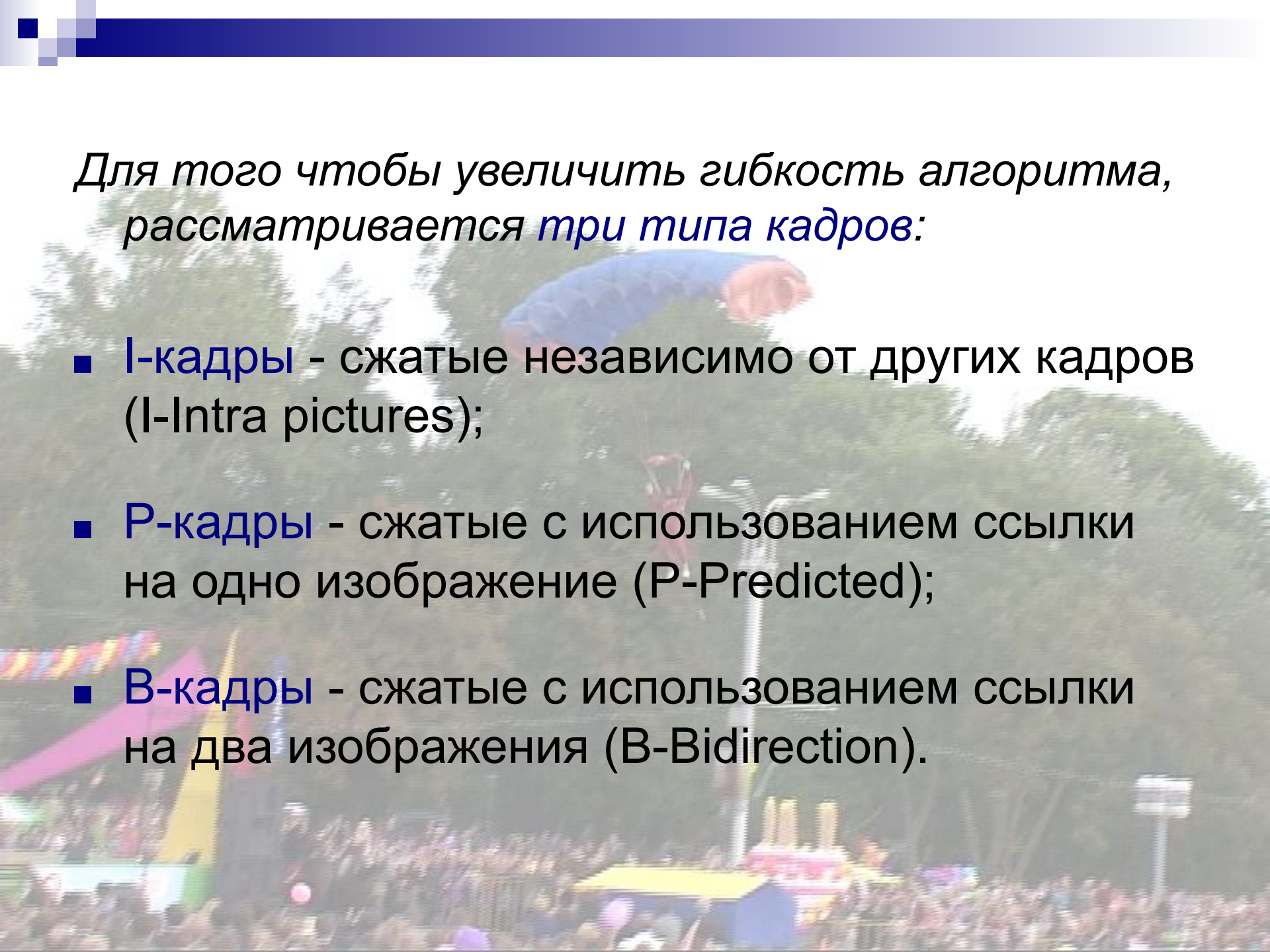
Деление изображения на срезы и макроблоки



Изображение (кадр)

Изображение — основная единица, которая принимается как входящая и выходящая для отображения.



A background image showing a person in a blue and red jumpsuit skydiving over a large crowd of people at an outdoor event. The skydiver is in the upper center, with their arms and legs spread. Below them, a dense crowd of people is visible, some holding up phones or cameras. The scene is outdoors with trees in the background.

Для того чтобы увеличить гибкость алгоритма,
рассматривается *три типа кадров*:

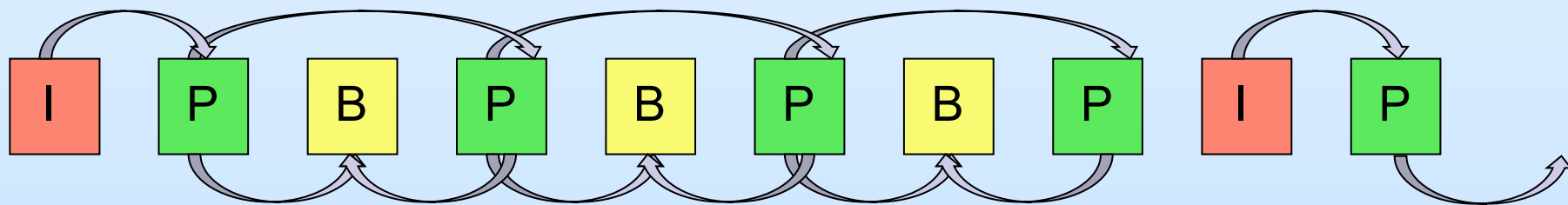
- **I-кадры** - сжатые независимо от других кадров (I-Intra pictures);
- **P-кадры** - сжатые с использованием ссылки на одно изображение (P-Predicted);
- **B-кадры** - сжатые с использованием ссылки на два изображения (B-Bidirection).

Группа изображений (Group of Pictures, GOP)

GOP — это набор изображений, который включает в себя I-кадры, P-кадры и B-кадры.

GOP может включать до 15 кадров и должна обязательно начинаться с I-кадра.

Ее структуру описывают как M/N , где M - общее число кадров в группе, а N - интервал между P-кадрами.

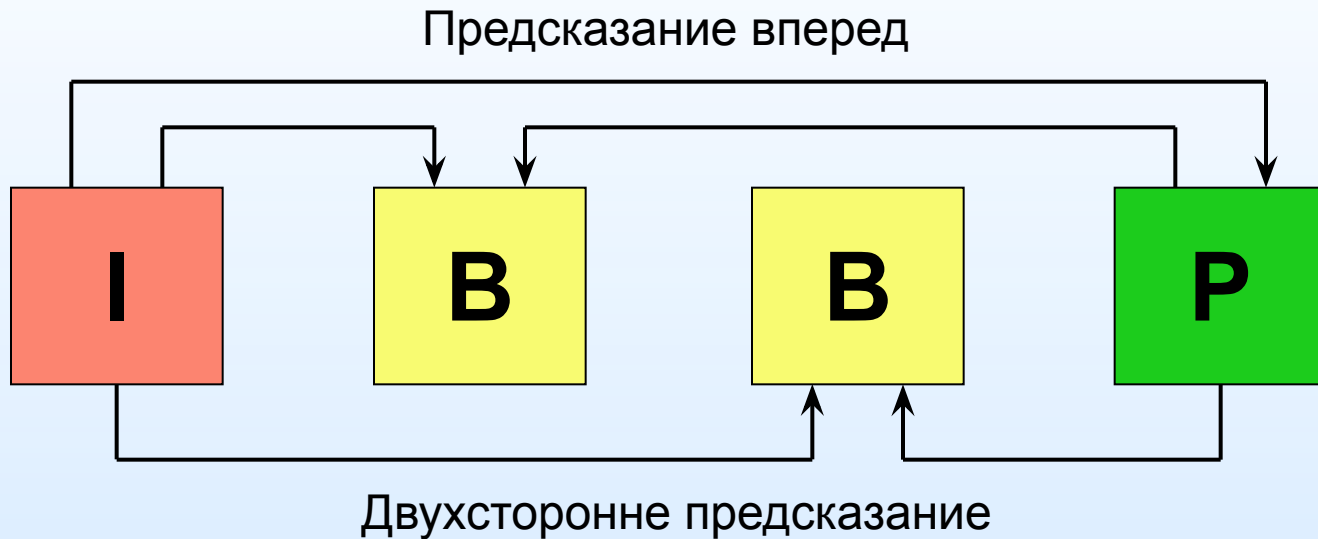


GOP структуры 15/3



Так как В-кадры должны быть получены из I- и Р-кадров, то последовательность кадров в соответствии с их номерами будет воспроизводиться в следующем порядке: 1423765...

Последовательность кадров GOP



- I** – Видеокадр с внутрикадровым кодированием
- B** – Предсказанный вперед кадр. Кодер передает разницу между I и P, декодер добавляет разницу и получает P-кадр
- P** – Двухстороннепредсказанный кадр можно получить из предыдущего I- или P-кадра или последующего I- или P-кадра. B-кадры друг из друга не получают

Видеопоследовательность

Видеопоследовательность – структура самого высокого уровня в данной иерархии. Она должна содержать минимум одну GOP, а также заголовок в начале последовательности (`sequence_header_code`) и код конца последовательности (`sequence_end_code`).

`Sequence_header_code`

`Sequence_end_code`



Видеопоследовательность

Группа изображений
(I-, P-, и B-кадры)



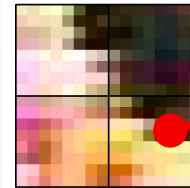
Изображение



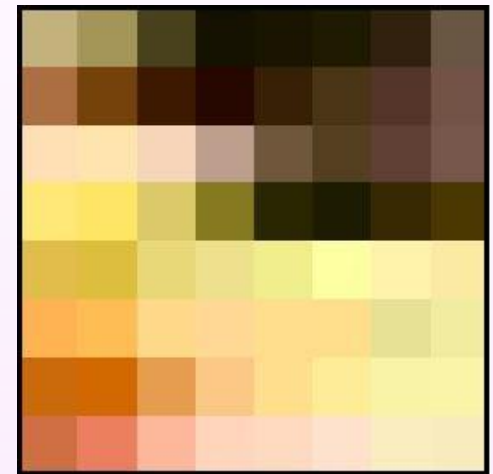
Срез



Макроблок



Блок



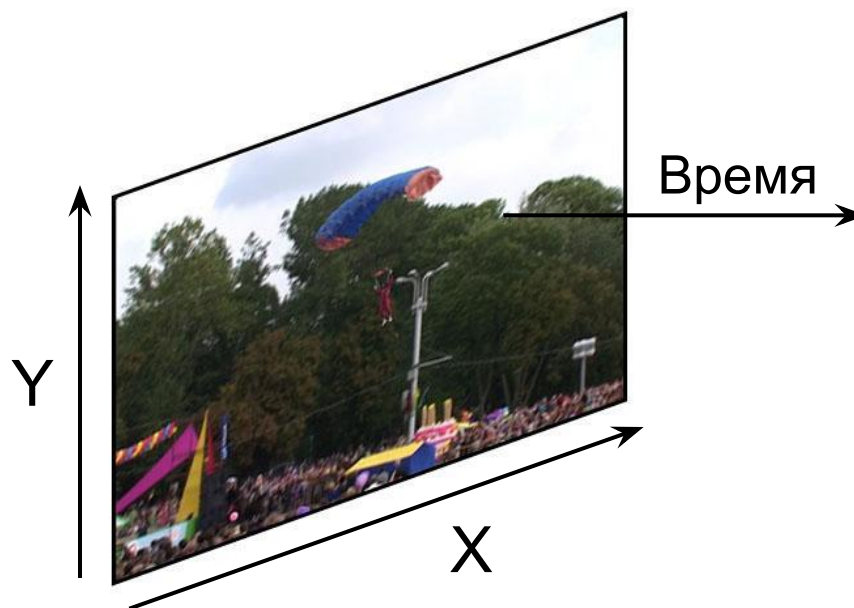
Технология сжатия видео в MPEG

Временное сжатие

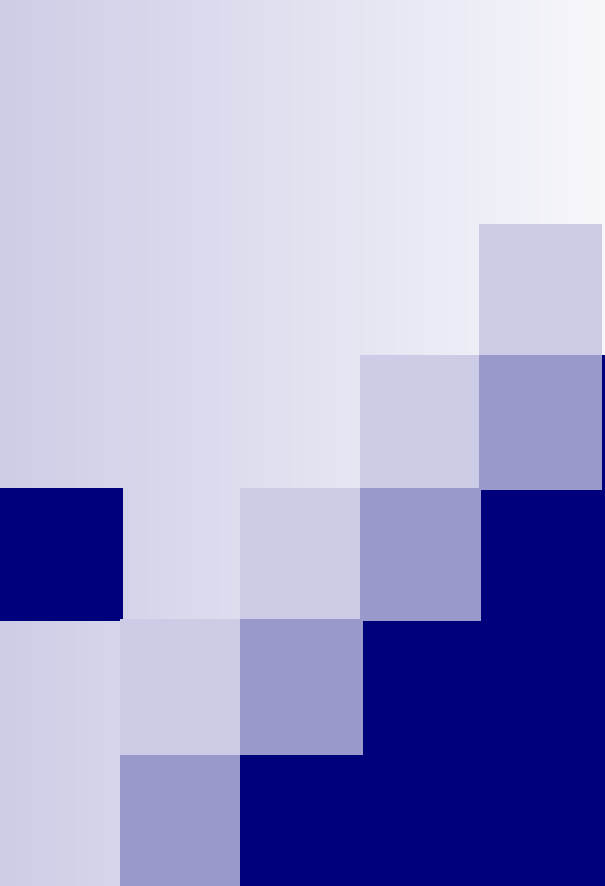
В- и Р-кадры

Пространственное сжатие

I-кадры



Удаление избыточности по трём осям

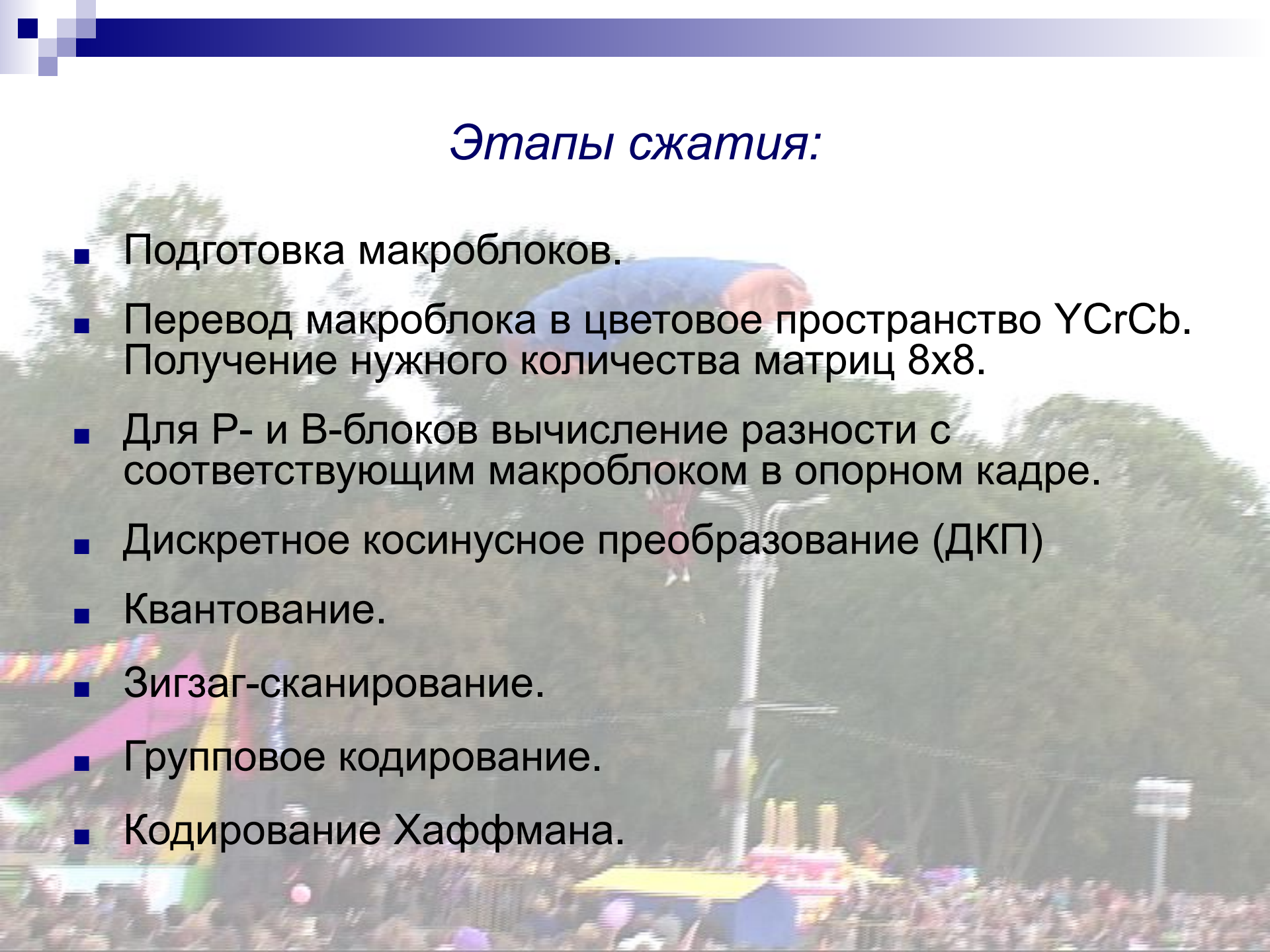


Общая схема алгоритма сжатия



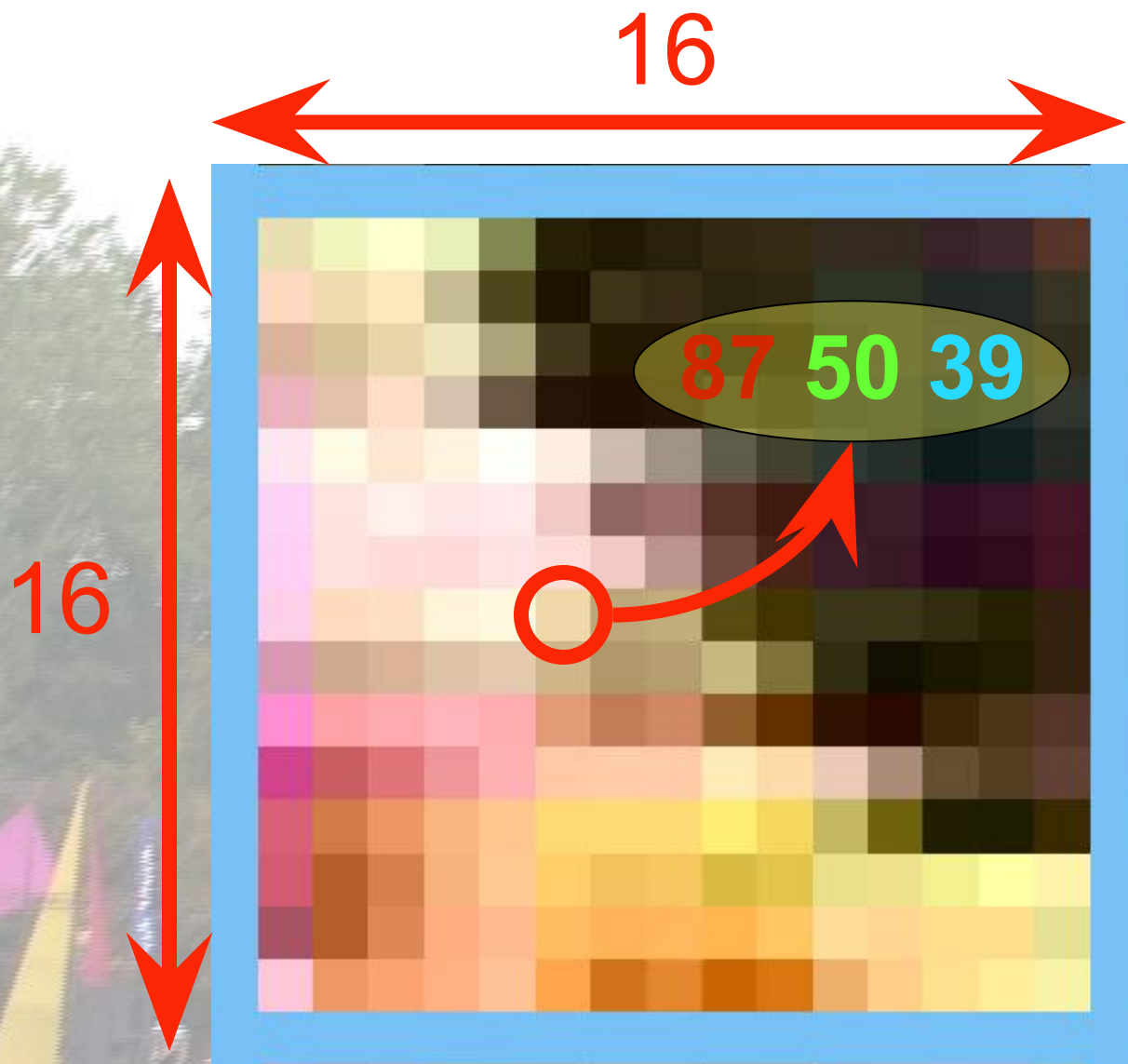
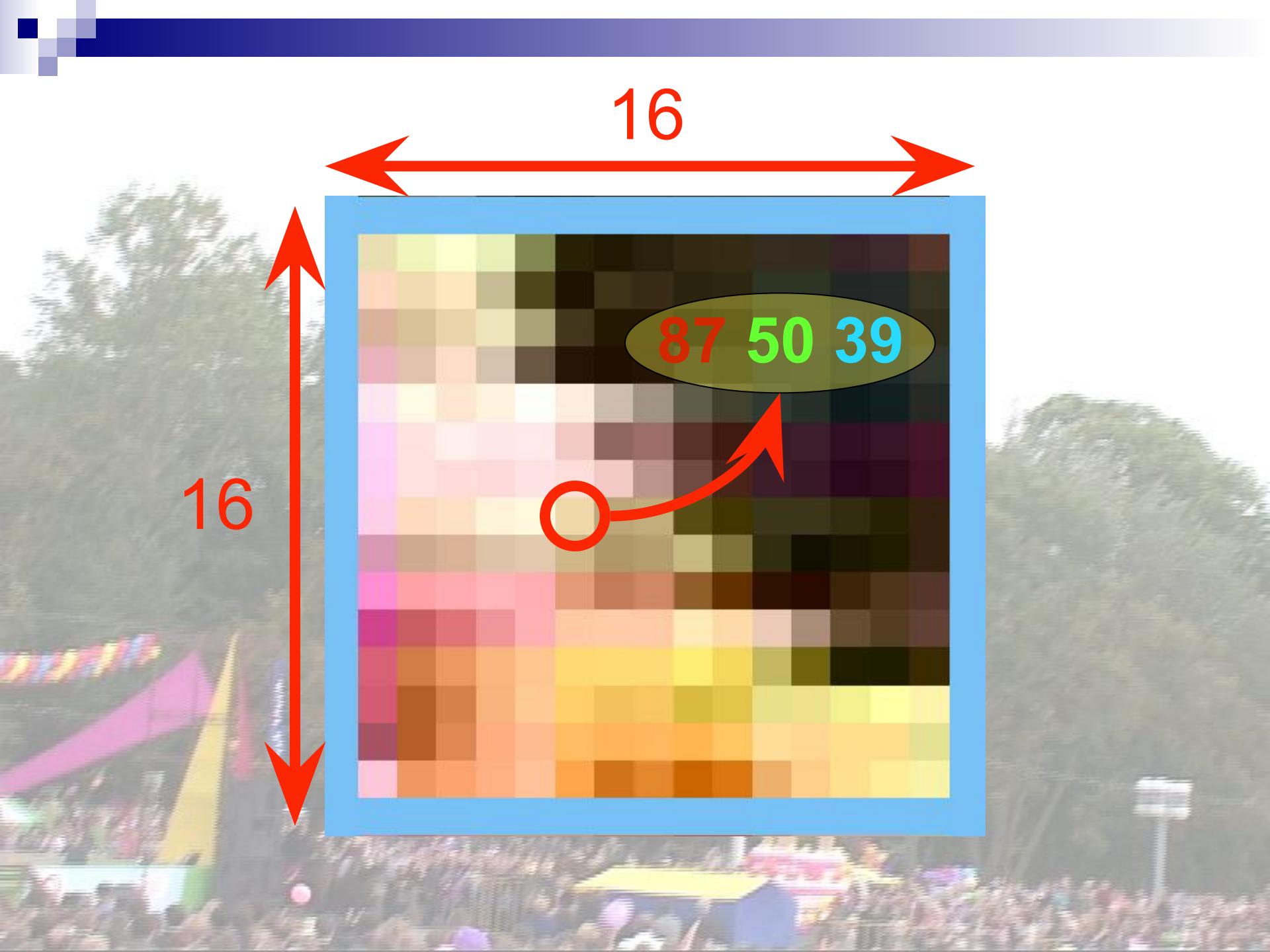


Этапы сжатия:

- Подготовка макроблоков.
 - Перевод макроблока в цветовое пространство YCrCb. Получение нужного количества матриц 8x8.
 - Для Р- и В-блоков вычисление разности с соответствующим макроблоком в опорном кадре.
 - Дискретное косинусное преобразование (ДКП)
 - Квантование.
 - Зигзаг-сканирование.
 - Групповое кодирование.
 - Кодирование Хаффмана.
- 

1. Изображение разбивается на макроблоки размером 16x16.





292 279 247 267 184 182 39 29 35 52 59 55 32 40 46 42
282 292 270 282 198 79 30 25 65 40 40 45 45 38 40 35
156 156 154 170 189 120 69 36 34 12 24 27 37 46 38 28
184 189 172 200 174 167 36 33 34 46 37 30 39 39 44 47
238 240 248 230 243 245 238 184 183 73 45 51 37 33 37 46
254 247 225 236 231 234 197 98 98 89 66 59 33 30 39 69
292 294 224 218 218 226 237 207 207 64 26 62 56 43 29 37
240 233 299 297 247 247 242 169 169 169 66 66 269 269 49 37 15
184 189 147 159 169 174 137 165 165 128 157 48 119 227 392 35
246 243 166 176 184 176 126 88 88 47 09 48 41 60 26 42
148 139 58 124 153 176 166 169 169 184 168 183 120 49 33 53
188 209 124 89 187 189 215 219 219 216 24 69 94 149 36 94 03
153 113 43 125 125 104 120 61 92 192 65 75 7 138 126 144 164 171
124 1100 43 285 2129 143 99 293 297 80 98 153 147 140 140 149
227 216 162 143 132 180 178 126 160 161 178 108 136 142 154 168
244 180 111 142 135 132 76 43 68 96 156 188 192 203 191 190
256 199 152 164 176 194 165 115 115 102 118 177 206 225 237 243
254 169 130 156 155 256 203 202 201 129 184 213 215 226 247
235 253 228 255 255 255 230 195 208 235 253 255 255 255 250
248

R

B

G

2. Перевод макроблоков в цветовое пространство YCrCb.

Матрица перехода:

$$\begin{bmatrix} Y \\ Cr \\ Cb \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.334 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{bmatrix} * \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix}$$

Обратное преобразование (на выходе декодера):

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.000 & 0 & 1.404 \\ 1.000 & -0.3434 & -0.712 \\ 1.000 & 1.773 & 0 \end{bmatrix} * \begin{bmatrix} Y \\ Cr \\ Cb \end{bmatrix} - \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix}$$

247 234 242 255 232 130 39 32 31 52 54 53 52 56 62 87
245 255 240 252 197 79 30 62 61 41 45 46 46 35 36 56
219 219 215 233 238 172 66 36 34 17 24 36 47 36 35 49
231 235 226 254 214 107 38 33 34 46 37 36 39 23 29 50
255 255 255 255 255 255 255 206 205 96 68 47 37 13 17 38
255 255 255 255 255 255 243 144 143 87 63 58 61 50 55 69
255 255 255 255 255 255 243 243 111 73 60 56 43 46 67
255 255 255 255 255 255 240 190 191 88 66 59 59 49 37 52
220 218 208 218 223 228 207 175 175 199 128 49 19 27 32 52
255 255 255 255 255 255 227 195 197 144 99 48 41 60 76 86
217 208 201 224 239 255 255 255 255 255 255 234 171 101 84 97
228 216 209 237 250 255 255 255 255 255 246 196 111 33 31 58
250
211 180 213 248 255 255 243 243 217 227 234 238 243 255 255
192 168 181 223 255 255 255 255 255 255 255
255 255 255 230
255
236 233 245 255 238 136 37 25 25 40 42 44 43 35 41 53
236 239 249 255 238 238 209 209 209 209 209 209 209 209 209
255 255 255 255 255 255 255 255 255 255 255 255 255 255
255 255 255 255 255 255 255 255 255 255 255 255 255 255
179 179 194 213 229 163 54 24 24 12 19 44 54 41 40 45
174 179 193 221 195 87 20 15 15 42 34 46 49 36 42 53
228 230 249 231 243 255 238 189 189 92 64 57 47 26 31 41
217 210 227 239 231 233 201 102 102 50 26 26 30 15 20 21
194 207 226 220 218 226 221 201 201 74 36 29 24 8 11 18
214 207 219 224 244 249 216 166 166 75 53 54 54 45 33 34
154 152 170 179 195 201 183 152 152 185 114 45 15 23 28 34
146 142 162 171 185 173 155 123 123 93 49 17 10 37 54 53
79 169 93 116 150 176 201 204 204 235 219 202 140 78 62 65
108 97 124 152 181 199 217 219 219 232 216 184 99 30 29 42
131 91 95 129 179 201 206 194 194 187 197 223 226 241 255
243
107 83 94 134 173 187 189 180 180 181 198 221 215 223 223
225
210 199 152 164 176 194 165 115 115 102 118 177 206 225 237
192 179 187 207 184 82 5 0 0 17 19 31 30 40 46 42
182 182 137 182 146 27 0 25 25 13 16 37 37 38 40 35
156 156 151 170 181 120 28 0 0 0 0 27 38 40 38 28
184 189 172 200 174 67 5 0 0 23 14 40 43 39 44 47
238 240 228 210 223 249 224 174 174 73 45 51 41 28 33 36
254 247 225 236 231 234 197 98 98 39 16 39 43 34 39 39
232 244 224 218 218 227 217 197 197 64 25 42 37 27 29 37
240 233 190 197 217 222 169 119 119 17 0 26 26 15 2 15
181 179 147 151 169 174 137 105 105 128 57 17 0 0 0 15
216 211 166 176 184 176 120 88 88 47 0 0 0 8 25 42
148 139 58 121 153 178 166 169 169 184 168 183 120 49 33 53
131 119 71 99 127 143 115 117 117 116 94 99 14 0 0 0
153 113 41 75 125 147 106 92 92 65 75 138 141 144 164 171
124 100 43 85 129 143 99 91 91 80 98 153 147 140 140 149
227 216 102 113 132 150 76 26 26 1 17 108 138 142 154 168
214 180 111 142 135 132 76 43 68 96 156 188 192 203 191 190

234 221 237 250 231 128 34 24 24 41 43 45 44 42 48 62
214 225 221 232 186 68 20 51 51 35 38 49 49 38 39 51
188 188 195 214 226 161 55 25 24 12 18 40 50 39 38 44
192 197 200 228 198 91 24 19 19 41 33 42 45 32 38 51
237 239 248 236 244 254 241 192 192 91 63 53 43 22 27 40
233 228 235 243 238 240 213 114 114 60 36 37 41 28 33 37
217 226 234 230 229 235 231 213 213 84 46 40 35 21 24 35
229 224 226 230 244 248 218 168 168 72 51 52 52 43 31 37
177 175 179 187 200 206 185 154 154 183 112 43 14 22 26
217
187 184 190 197 206 198 173 141 141 103 58 24 18 41 57 62
128 177 121 149 177 200 213 215 215 235 224 209 147 82 65 73
147 135 143 171 195 209 217 218 218 226 211 178 93 27 26 42
169 129 114 148 193 211 209 197 197 182 192 217 220 231 245
238
134 110 114 155 193 202 198 192 192 192 204 223 219 223 223
218
187 183 181 180 180 146 125 125 125 122 132 132 138 138 146 148
146
192 195 184 185 172 132 127 124 134 134 128 131 123 126 122 138
138
184 184 175 182 183 163 135 126 126 127 126 133 137 140 140
135
201 203 188 198 186 151 130 129 129 133 130 139 140 139 141
142
214 215 203 199 202 211 204 187 187 150 141 143 139 136 137
139
226 225 208 210 210 211 201 168 167 146 138 149 150 148 150
152
223 224 208 207 208 210 206 201 201 154 141 150 148 146 147
152
220 219 193 195 199 200 181 165 165 126 122 133 133 129 124
133
205 204 180 181 186 187 171 160 160 164 140 130 126 125 124 32
231 229 200 202 202 202 175 164 164 145 128 130 132 138 138
146
250 149 160 188 195 202 187 188 188 185 182 192 176 144 138
138
168 152 158 167 174 177 137 137 137 135 132 125 126 126 126
134
256 155 149 155 139 148 138 138 139 135 134 135 126 126 126
176
187 179 129 164 178 188 138 137 137 134 134 174 173 167 167
167
244 243 169 175 149 189 158 138 138 123 134 163 142 164 146
254
255 209 149 187 143 143 165 149 169 144 188 194 193 198 187
167
146 150 148 146 146 133 144 144 144 139 139 133 133 132 133
139
159 159 149 150 144 144 144 143 143 140 140 132 131 132 132 10
177 179 174 170 163 169 167 167 168 157 157 145 144 142 141

Макроблоки

```
graph TD; A[Макроблоки] --> B[I-типа]; A --> C[P-типа]; A --> D[B-типа];
```

I-типа

P-типа

B-типа

Для каждого макроблока определяется, каким образом он будет сжат. Отдельные макроблоки сжимаются независимо, т. е. в В-кадрах можно сжать конкретный макроблок как I-блок, Р-блок со ссылкой на предыдущий кадр, Р-блок со ссылкой на последующий кадр и, наконец, как В-блок.

3. Для Р- и В-макроблоков вычисление разности с соответствующим макроблоком в опорном кадре.

Р- или В-макроблок
в Р- или В-кадре

Соответствующий
макроблок в опорном кадре

Макроблок разности

38	46	48	38	20	8	4	5	4	4	4	3	2	3	4	3
33	40	44	43	38	32	20	9	2	3	4	4	4	4	4	4
34	38	39	34	29	26	18	6	6	4	4	4	4	4	4	3
24	38	44	37	44	33	33	25	9	7	4	3	3	4	4	3
40	39	38	42	48	49	36	17	7	6	5	4	4	4	4	4
25	45	53	37	58	56	50	41	23	12	4	2	2	3	5	6
44	45	42	46	52	51	42	28	10	6	5	8	7	2	1	2
31	47	36	54	56	38	56	49	36	19	4	2	4	6	5	3
42	36	48	45	46	42	36	24	7	3	3	5	4	2	2	3
33	41	46	50	35	53	53	45	31	15	4	3	5	4	4	4
42	36	38	45	34	18	10	3	5	2	6	10	8	2	2	6
42	44	46	47	45	11	3	2	1	3	5	4	1	3	5	4
39	31	31	45	39	18	6	5	15	37	48	31	7	2	5	4
34	29	36	29	6	2	2	5	11	31	26	5	1	5	4	3
40	46	67	64	57	32	41	44	53	60	58	34	7	7	15	12
44	49	60	55	16	19	20	44	67	52	22	1	4	10	7	3

31	29	27	31	45	57	47	21	5	4	5	7	6	2	2	5
27	33	34	39	47	44	22	5	5	8	7	6	5	4	4	4
33	30	25	21	20	20	9	5	8	10	13	4	1	3	3	4
29	34	35	29	24	21	9	5	6	6	6	4	6	9	7	5
35	34	31	28	27	25	18	21	33	24	9	6	4	6	13	13
33	35	33	40	33	39	29	25	20	9	3	2	2	7	10	4
36	40	28	26	26	30	37	50	57	51	23	13	13	25	34	28
34	32	35	36	38	40	34	29	20	5	2	7	18	22	13	1
36	31	27	29	36	44	49	58	66	38	47	51	60	68	62	27
35	30	32	36	38	43	45	38	29	42	44	43	57	39	17	12
38	36	35	51	50	65	54	70	71	67	65	58	63	66	63	62
35	33	43	37	40	47	53	65	66	63	62	66	67	55	47	53
57	63	54	73	58	71	67	65	64	63	56	62	60	58	61	59
39	43	46	51	57	65	73	75	66	60	61	66	61	56	54	57
52	76	78	76	72	73	75	74	71	70	71	77	73	72	68	65
49	56	74	73	71	68	67	65	62	61	61	58	59	64	62	60

7	17	21	7	-25	-49	-43	-16	-1	0	-1	-4	-4	1	2	-2
6	7	10	4	9	-16	-2	4	-3	-5	-3	-2	-1	0	0	0
1	8	14	13	9	6	9	1	-2	-6	-9	0	3	1	1	-1
-5	4	9	8	20	12	24	20	3	1	-2	-1	-3	-5	-3	-2
5	5	7	14	21	24	18	-4	-26	-18	-4	-2	0	-2	-9	-9
-8	10	20	-3	15	17	21	16	3	3	1	0	0	-4	-5	2
8	5	14	20	26	21	5	22	-47	-45	-18	-5	-6	-23	-33	-26
3	15	1	18	18	-2	22	20	16	14	1	-5	-14	-16	-8	2
6	5	21	16	10	-2	-13	-34	-59	-35	-44	-46	-56	-66	-60	-24
-2	11	14	14	-3	10	8	17	2	-27	-40	-40	-42	-35	-13	-8
4	0	3	-6	-16	-47	-44	-67	-66	-65	-59	-48	-55	-64	-61	-56
7	11	3	10	5	-36	-50	-63	-65	-60	-57	-62	-66	-52	-42	-49
-18	-32	-23	-28	-19	-53	-61	-60	-49	-26	-8	-31	-53	-56	-56	-55
-5	-14	-10	-22	-51	-63	-72	-70	-55	-29	-35	-61	-60	-51	-50	-54
-12	-30	-11	-12	-15	-41	-34	-30	-18	-10	-13	-43	-66	-65	-53	-53
-5	-7	14	-18	-55	-49	-47	-21	5	-9	-39	-57	-55	-54	-55	-57

4. Получение блоков 8x8 в соответствии с выбранным форматом дискретизации.

234	221	237	250	231	128	34	24	24	41	43	45	44	42	48	62
214	225	221	232	186	68	20	51	51	35	38	49	49	38	39	51
188	188	195	214	226	161	55	25	24	12	18	40	50	39	38	44
192	197	200	228	198	91	24	19	19	41	33	42	45	32	38	51
237	239	248	236	244	254	241	192	92	91	63	53	43	22	27	40
233	228	235	243	238	240	213	114	14	60	36	37	41	28	33	37
217	226	234	230	229	235	231	213	13	84	46	40	35	21	24	35
229	224	226	230	244	248	218	168	68	72	51	52	52	43	31	37
177	175	179	187	200	206	185	154	154	183	112	43	14	22	26	217
187	184	190	197	206	198	173	141	141	103	58	24	18	41	57	62
128	177	121	149	177	200	213	215	215	235	224	209	147	82	65	73
147	135	143	171	195	209	217	218	218	226	211	178	93	27	26	42
169	129	114	148	193	211	209	197	197	182	192	217	220	231	245	238
134	110	114	155	193	202	198	192	192	192	204	223	219	223	223	218
225	218	172	184	195	207	182	133	133	120	136	187	213	225	233	236
219	195	157	190	204	200	156	123	135	156	201	223	226	232	236	234

187 183 181 190 180 146 125 125 125 132 133 138 138 146 148 146
192 195 184 185 172 132 127 134 134 129 131 137 137 140 140 138
184 184 175 182 183 163 135 126 126 127 126 133 137 140 140 135
201 203 188 198 186 151 130 129 129 133 130 139 140 139 141 142
214 215 203 199 202 211 204 187 187 150 141 143 139 136 137 139
226 225 208 210 210 211 201 168 167 146 138 149 150 148 150 152
223 224 208 207 208 210 206 201 201 154 141 150 148 146 147 152
220 219 193 195 199 200 181 165 165 126 122 133 133 129 124 133
205 204 180 181 186 187 171 160 160 164 140 130 126 125 124 32
231 229 200 202 202 202 175 164 164 145 128 130 132 130 135 146
212 177 160 188 195 202 187 188 188 185 182 192 170 144 138 149
196 192 158 167 174 177 157 157 157 152 145 150 121 124 124 124
203 190 147 159 173 178 156 151 151 135 139 163 164 161 169 176
187 179 149 164 178 181 158 157 157 151 154 174 167 167 167 167
215 213 169 173 179 182 155 138 168 129 134 163 170 167 170 174
211 205 179 187 175 175 161 149 161 174 188 194 187 198 187 187

137 137 131 132 129 129 132 134 133 136 136 134 134 138 138 146
150 149 141 142 136 136 135 136 135 132 133 126 126 126 126 131
150 150 142 141 136 136 136 136 135 131 132 125 126 126 125 132
156 155 146 146 139 140 138 138 139 132 131 123 123 121 127 127
141 140 133 142 136 128 138 138 137 132 132 123 123 122 123 129
144 147 142 136 140 139 149 149 149 147 143 143 142 144 144 151
155 149 143 146 147 142 145 149 149 147 145 142 143 146 147 151
146 150 148 146 146 133 144 144 144 139 133 132 133 132 133 139
159 159 149 150 144 144 144 143 143 140 141 130 131 131 132 10
177 179 174 170 163 169 167 167 168 157 157 145 144 142 141 145
191 150 185 182 172 167 158 156 156 142 150 146 145 144 138 145
186 186 175 175 167 161 155 154 154 149 153 141 141 132 131 139
186 186 175 174 167 159 161 161 161 153 153 140 141 137 135 140
169 169 176 176 173 166 168 173 173 173 165 151 154 151 151 137
149 155 176 176 171 162 181 182 182 186 186 164 158 150 144 137
154 169 179 174 165 167 181 179 180 185 165 151 149 144 138 138

5. Кодирование

Состоит из пяти последовательных этапов:

- Дискретное косинусное преобразование (ДКП)
- Квантование
- Зигзагообразное сканирование
- Групповое кодирование
- Кодирование Хаффмана

Рассмотрим кодирование второго блока из макроблока
компоненты яркости Y:

$y =$

24	41	43	45	44	42	48	62
51	35	38	49	49	38	39	51
24	12	18	40	50	39	38	44
19	41	33	42	45	32	38	51
192	91	63	53	43	22	27	40
114	60	36	37	41	28	33	37
213	84	46	40	35	21	24	35
168	72	51	52	52	43	31	37

■ Дискретное косинусное преобразование (ДКП)

ДКП для блоков 8х8:

$$Y(u, v) = \frac{1}{4} \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} C(i, u) \times C(j, v) \times y(i, j)$$

$$C(i, u) = A(u) \times \cos\left(\frac{(2i+1) \times u \times \pi}{2n}\right)$$

где $A(u) = \begin{cases} 1/\sqrt{2} & \text{при } u = 0 \\ 1 & \text{при } u \neq 0 \end{cases}$

Матрица ДКП имеет вид:

.353553	.353553	.353553	.353553	.353553	.353553	.353553	.353553
.490393	.415818	.277992	.097887	-.097106	-.277329	-.415375	-.490246
.461978	.191618	-.190882	-.461673	-.462282	-.192353	.190145	.461366
.414818	-.097106	-.490246	-.278653	.276667	.490710	.099448	-.414486
.353694	-.353131	-.354256	.352567	.354819	-.352001	-.355378	.351435
.277992	-.490246	.096324	.416700	-.414486	-.100228	.491013	-.274673
.191618	-.462282	.461366	-.189409	-.193822	.463187	-.460440	.187195
.097887	-.278653	.416700	-.490862	.489771	-.413593	.274008	-.092414

После применения ДКП получили матрицу:

$Y =$

398.250	107.051	80.863	49.575	71.500	18.631	10.342	9.366
-68.790	-135.992	-81.257	-63.074	-39.575	-25.715	-16.241	-8.510
17.083	1.144	4.273	8.570	0.991	4.863	8.510	-1.385
25.595	42.235	33.589	1.920	1.705	3.913	0.357	3.027
23.500	7.070	-1.864	-13.183	-8.750	-5.946	-0.581	-2.295
-27.738	-30.414	-7.139	-12.356	-11.097	-8.798	-7.442	-3.185
31.213	53.721	34.526	24.136	22.129	16.312	11.158	5.370

■ Квантование

Для каждой компоненты Y, Cr, Cb своя матрица квантования $Q(u,v)$:

$$Q(u,v) = 1 + (1 + u + v) \times q$$

Где q - это коэффициент качества, от него зависит степень потери качества сжатого изображения.

$$Yq(u,v) = \text{IntegerRound} \left[\frac{Y(u,v)}{Q(u,v)} \right]$$

Матрица квантования для $q=2$:

$Q =$

3	5	7	9	11	13	15	17
5	7	9	11	13	15	17	19
7	9	11	13	15	17	19	21
9	11	13	15	17	19	21	23
11	13	15	17	19	21	23	25
13	15	17	19	21	23	25	27
15	17	19	21	23	25	27	29
17	19	21	23	25	27	29	31

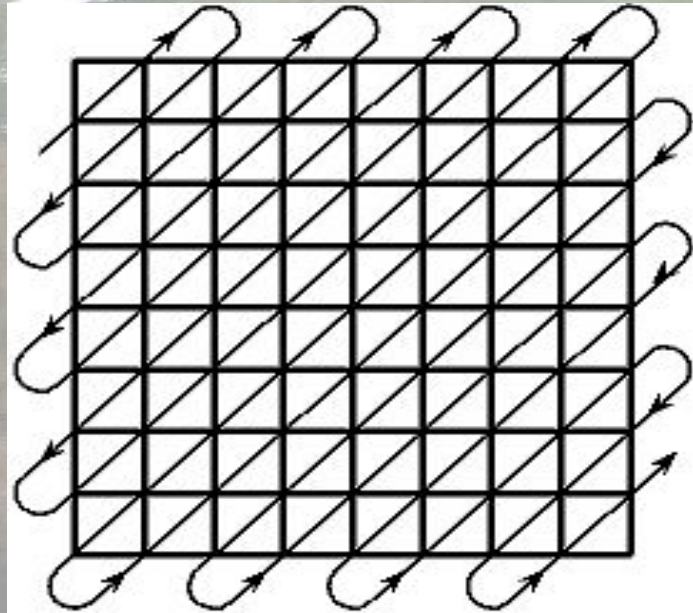
В результате квантования получили матрицу:

$Y_q =$

132	21	11	5	6	1	0	0
-13	-19	-9	-5	-3	-1	0	0
2	0	0	0	0	0	0	0
2	3	2	0	0	0	0	0
20	0	0	0	0	0	0	0
-2	-2	0	0	0	0	0	0
-2	-2	-1	0	0	0	0	0
0	1	2	1	1	0	0	0

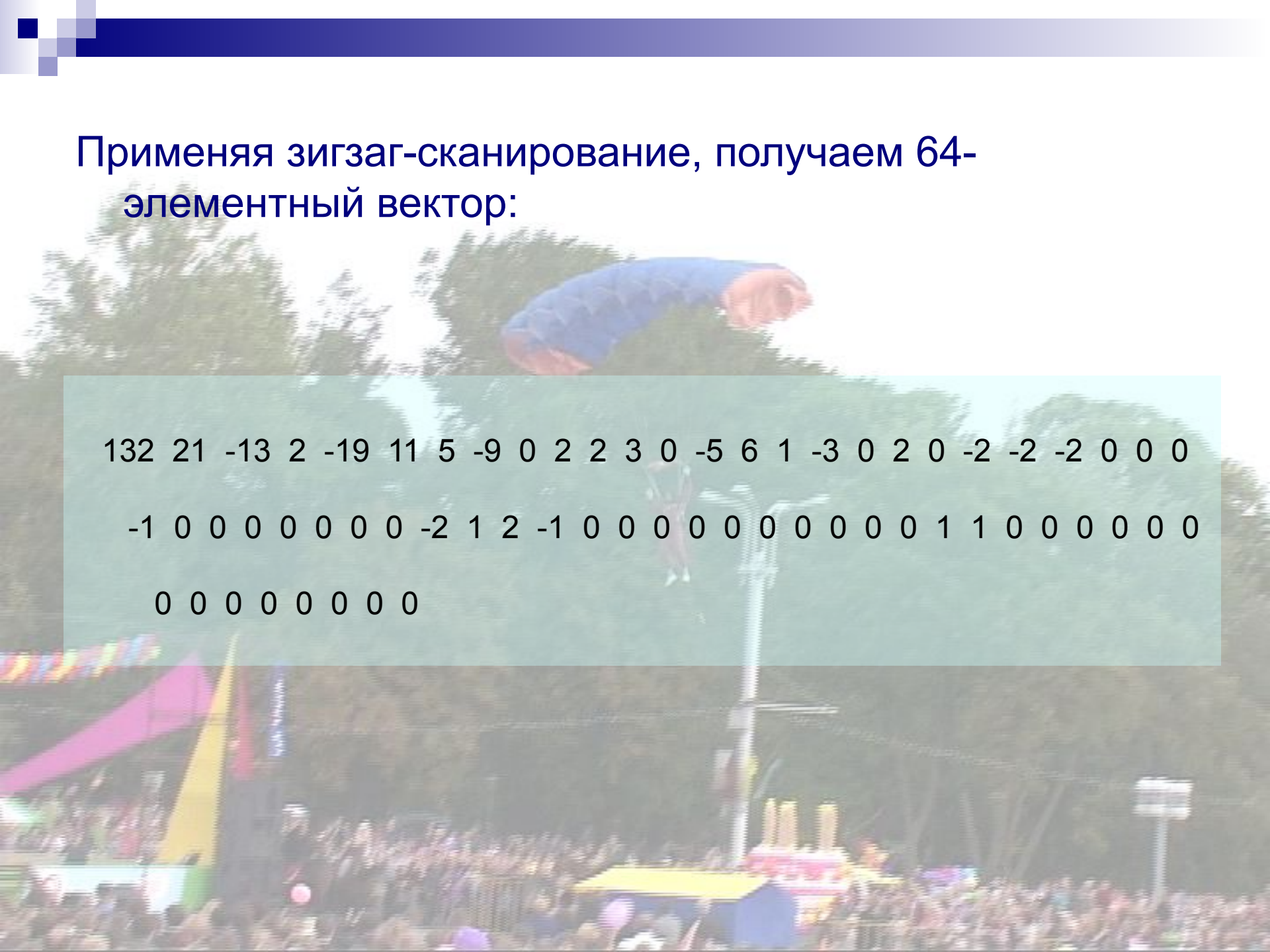
■ Зигзагообразное сканирование

Коэффициентов полученной на предыдущем этапе матрицы 8×8 $Yq(u,v)$ размещаются в одномерный 64-элементный вектор при помощи зигзагообразного сканирования матрицы, начиная с левого верхнего угла:



В начале вектора мы получаем коэффициенты матрицы, соответствующие низким частотам, а в конце — высоким.

Применяя зигзаг-сканирование, получаем 64-элементный вектор:



```
132 21 -13 2 -19 11 5 -9 0 2 2 3 0 -5 6 1 -3 0 2 0 -2 -2 -2 0 0 0  
-1 0 0 0 0 0 0 0 -2 1 2 -1 0 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0  
0 0 0 0 0 0 0 0
```

■ Групповое кодирование

Вектор, полученный на предыдущем этапе, свертывается с помощью алгоритма группового кодирования. При этом получаются пары типа **[пропустить, число]**, где “**пропустить**” является счетчиком пропускаемых нулей, а “**число**” — значение, которое необходимо поставить в следующую ячейку.

Так, полученный вектор будет свернут в пары:

```
[0, 132] [0, 21] [0, -13] [0, 2] [0, -19] [0, 11] [0, 5] [0, -9] [1, 2] [0, 2] [0, 3]  
[1, -5] [0, 6] [0, 1] [0, -3] [1, 2] [1, -2] [0, -2] [0, -2] [3, -1] [7, -2] [0, 1]  
[0, 2] [0, -1] [10, 1] [0, 1] [13, 0]
```

■ Кодирование Хаффмана

Полученные пары чисел кодируются с помощью алгоритма Хаффмана с фиксированной таблицей.

Несколько первых значений из таблицы В-14 (документ ISO 13818 – 2):

Код переменной длины ^[1]	Расстояние	Уровень
10	конец блока	
10s	0	1
11s	0	1
011s	1	1
0100s	0	2
0101s	2	1
0010 1s	0	3
0011 1s	3	1
0011 0s	4	1
0001 10s	1	2

^[1] “s” стоящее в конце каждого кода показывает знаковый разряд уровня.

При декодировании весь конвейер преобразований повторяется в обратном порядке.

Декодированная матрица

27	31	35	42	48	47	49	56
39	43	49	53	48	39	40	51
28	18	21	37	45	39	36	40
44	30	34	48	43	32	40	57
159	99	57	51	42	28	29	38
122	65	37	44	36	20	26	41
185	95	38	40	36	24	26	35
167	83	41	55	54	36	32	36

Исходная матрица изображения

24	41	43	45	44	42	48	62
51	35	38	49	49	38	39	51
24	12	18	40	50	39	38	44
19	41	33	42	45	32	38	51
192	91	63	53	43	22	27	40
114	60	36	37	41	28	33	37
213	84	46	40	35	21	24	35
168	72	51	52	52	43	31	37

$q = 2$

Потери ~ 13 %


$$q = 3$$

Декодированная
матрица

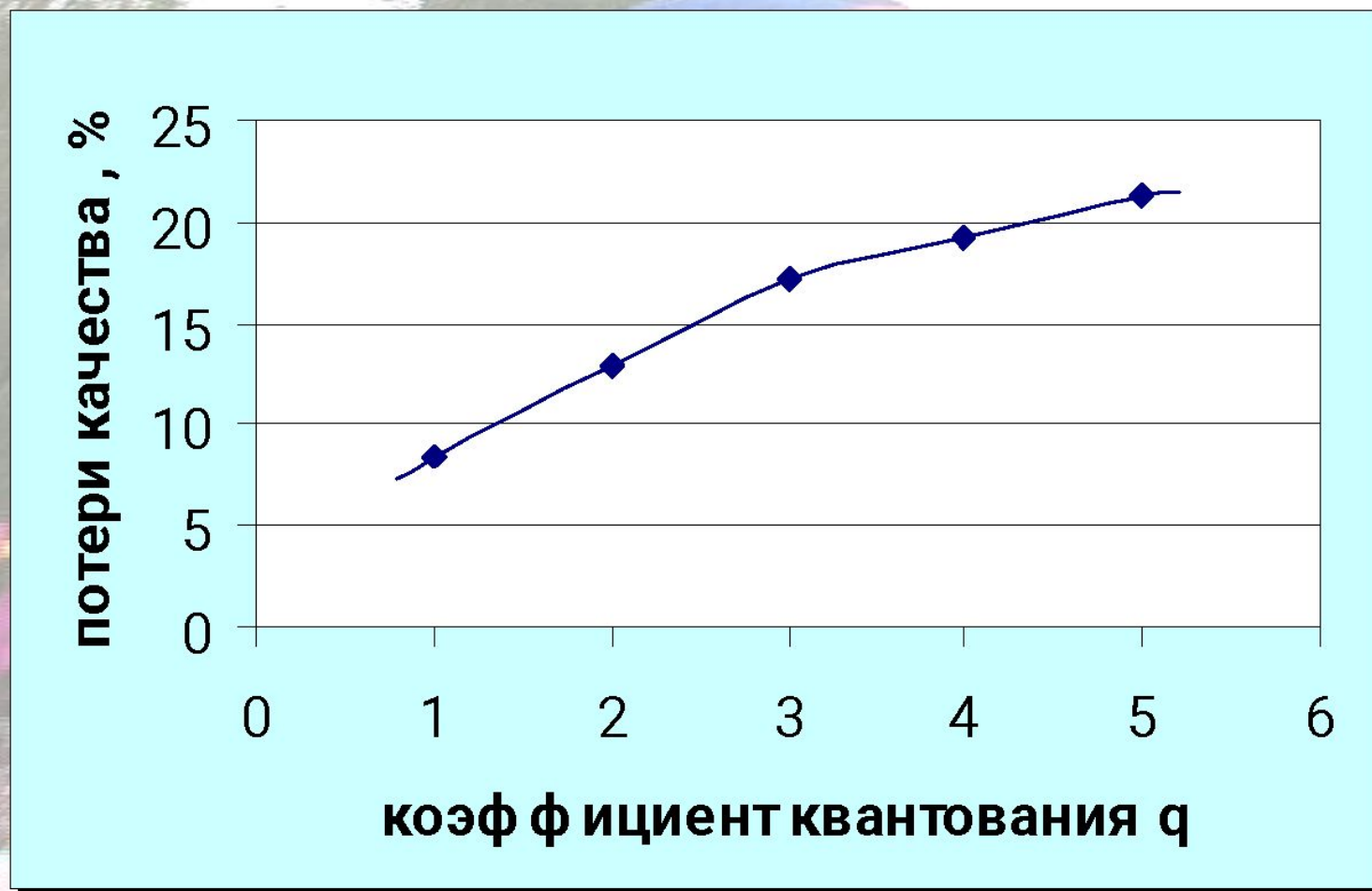
23	36	36	38	47	44	45	63
36	45	47	51	54	39	30	41
31	27	23	32	43	40	44	62
52	35	30	44	51	38	33	43
150	98	58	50	41	25	28	44
122	64	31	39	38	25	28	41
180	100	50	47	36	17	21	34
160	83	42	53	49	31	33	44

Исходная матрица
изображения

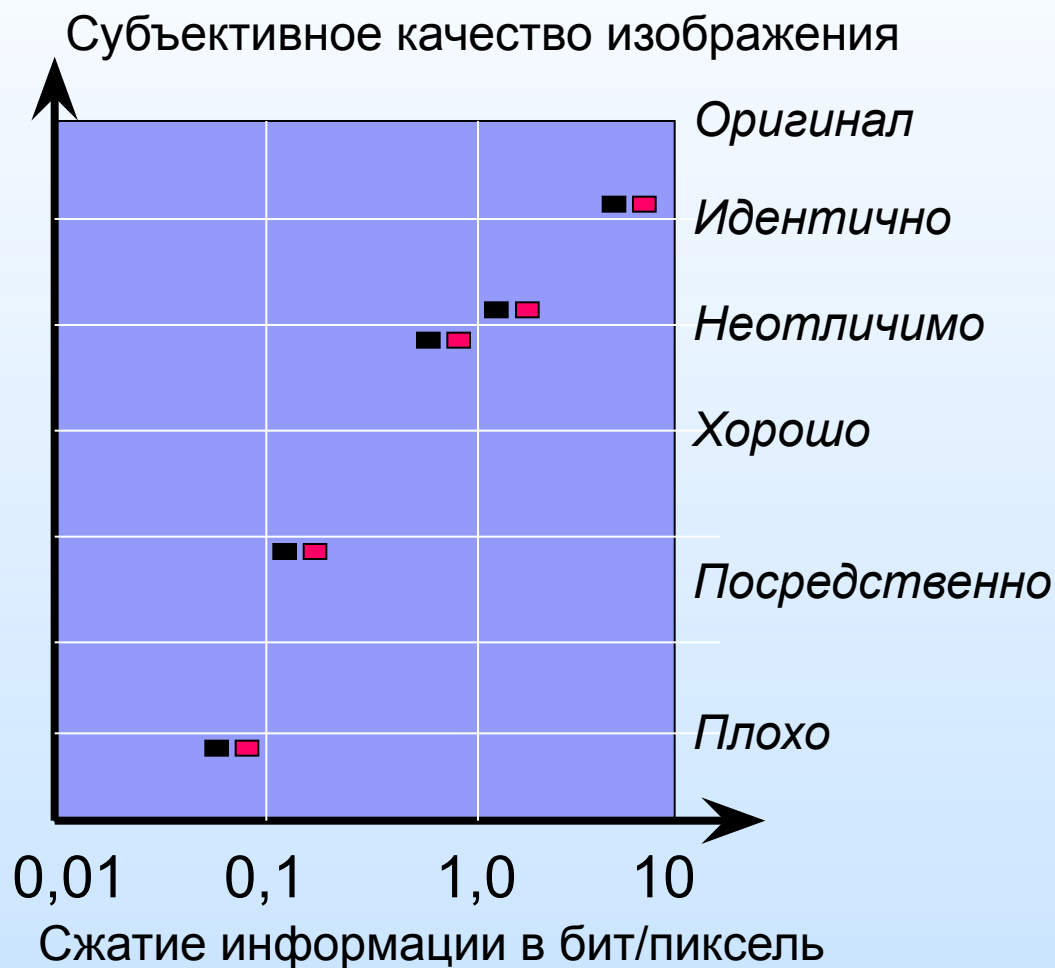
24	41	43	45	44	42	48	62
51	35	38	49	49	38	39	51
24	12	18	40	50	39	38	44
19	41	33	42	45	32	38	51
192	91	63	53	43	22	27	40
114	60	36	37	41	28	33	37
213	84	46	40	35	21	24	35
168	72	51	52	52	43	31	37

Потери ~ 17 %

Зависимость потерь компрессированного изображения от коэффициента квантования



Качество компрессированного изображения для различных значений сжатия информации

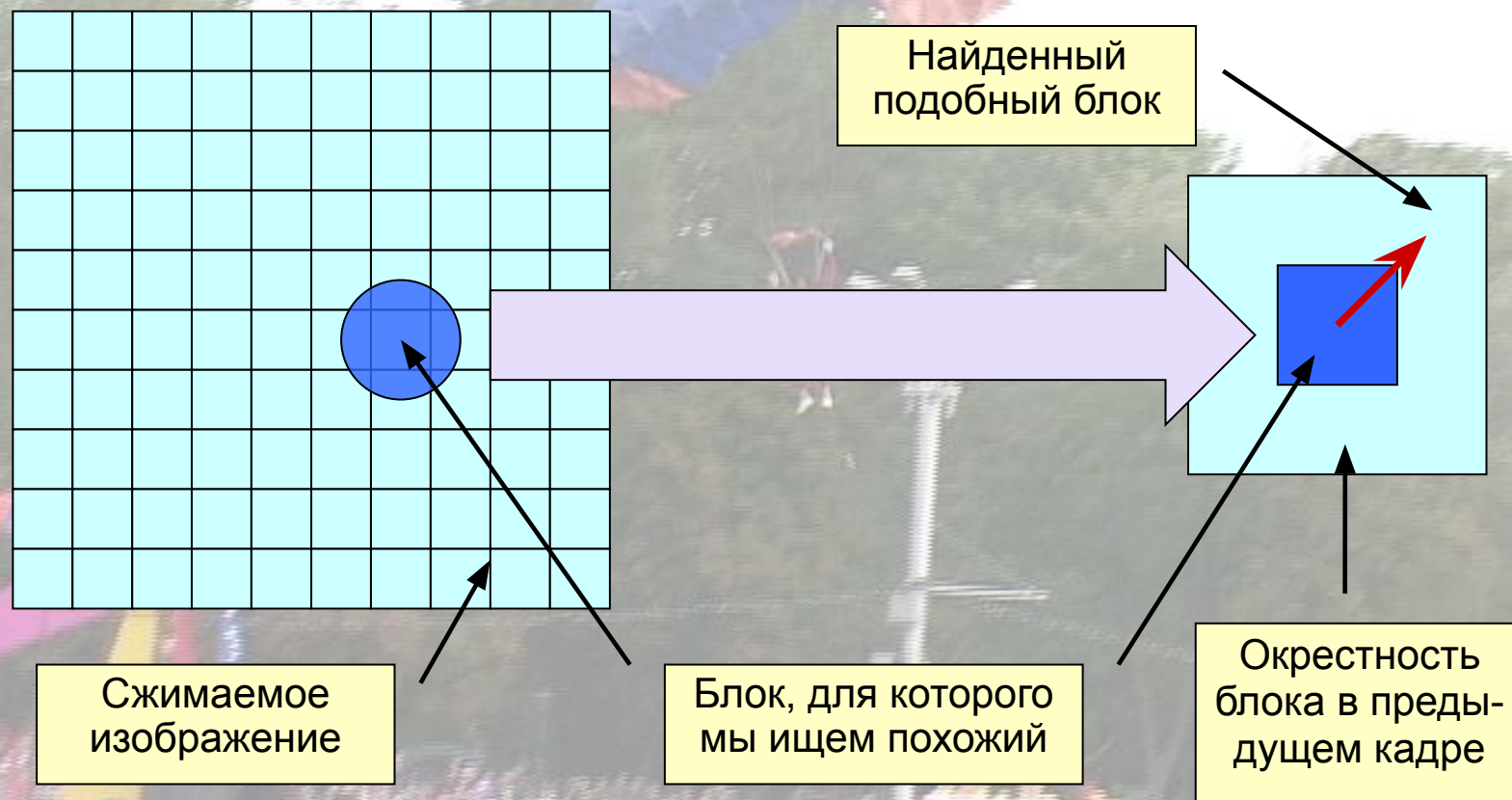


■ - черно-белое изображение

■ - цветное изображение

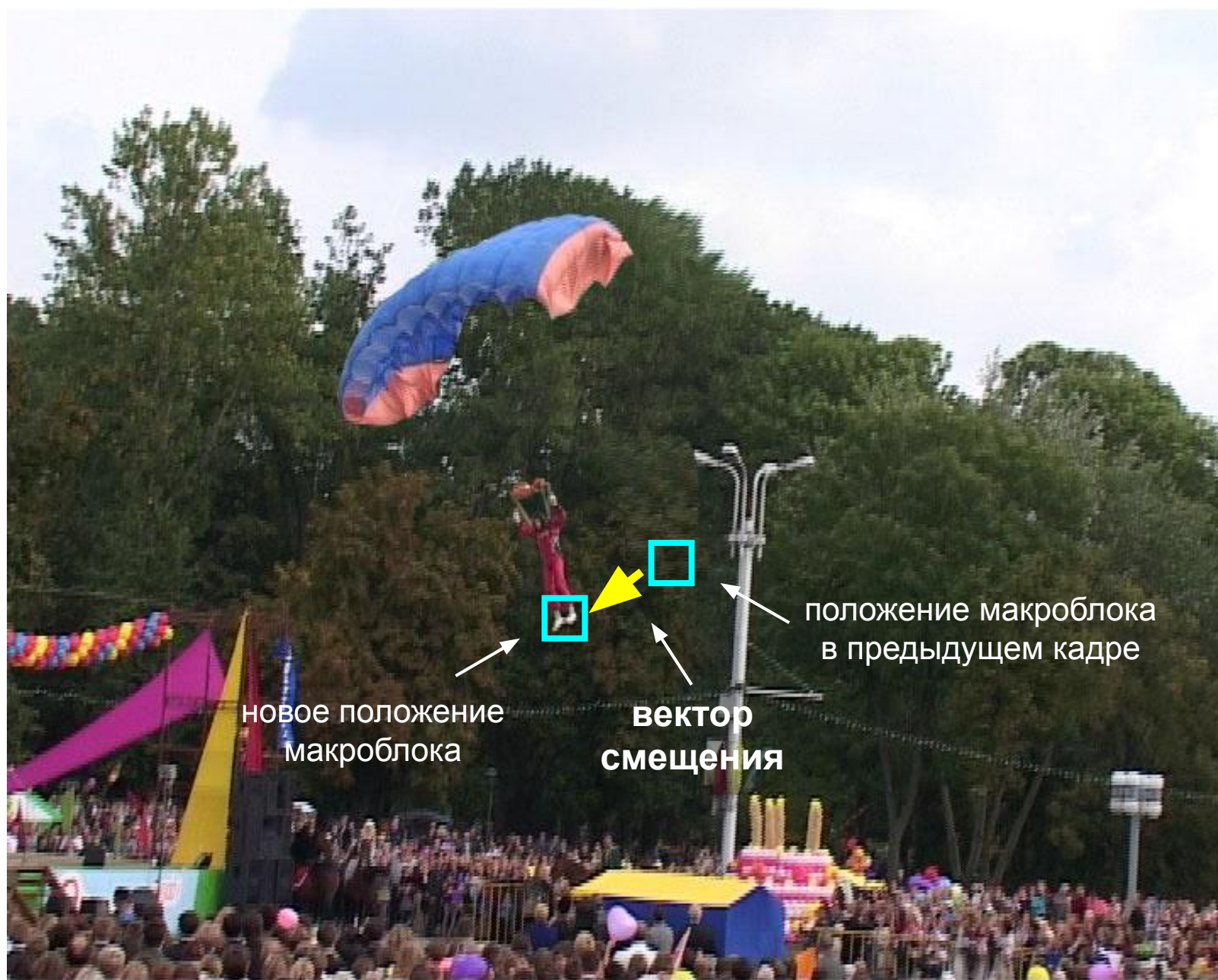


Использование векторов смещений блоков





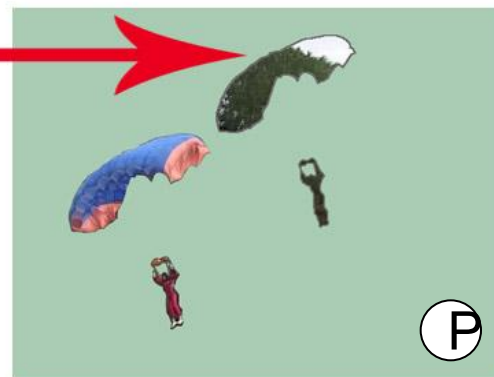
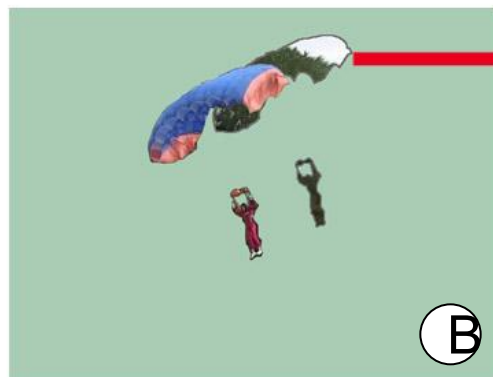
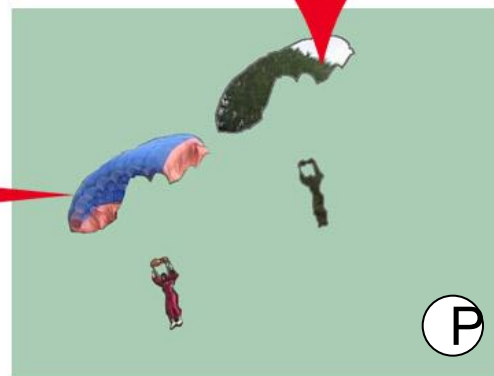




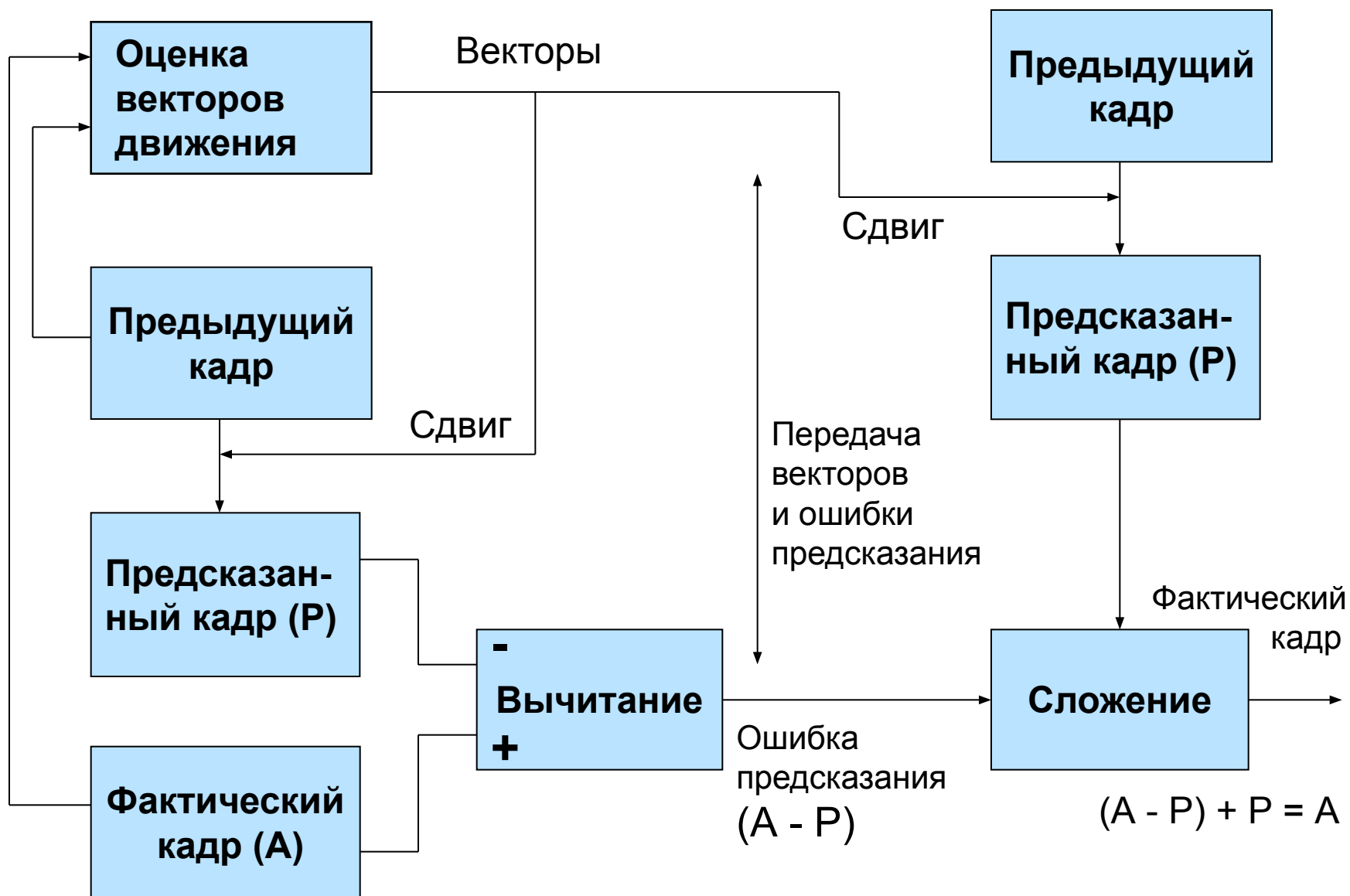
новое положение
макроблока

вектор
смещения

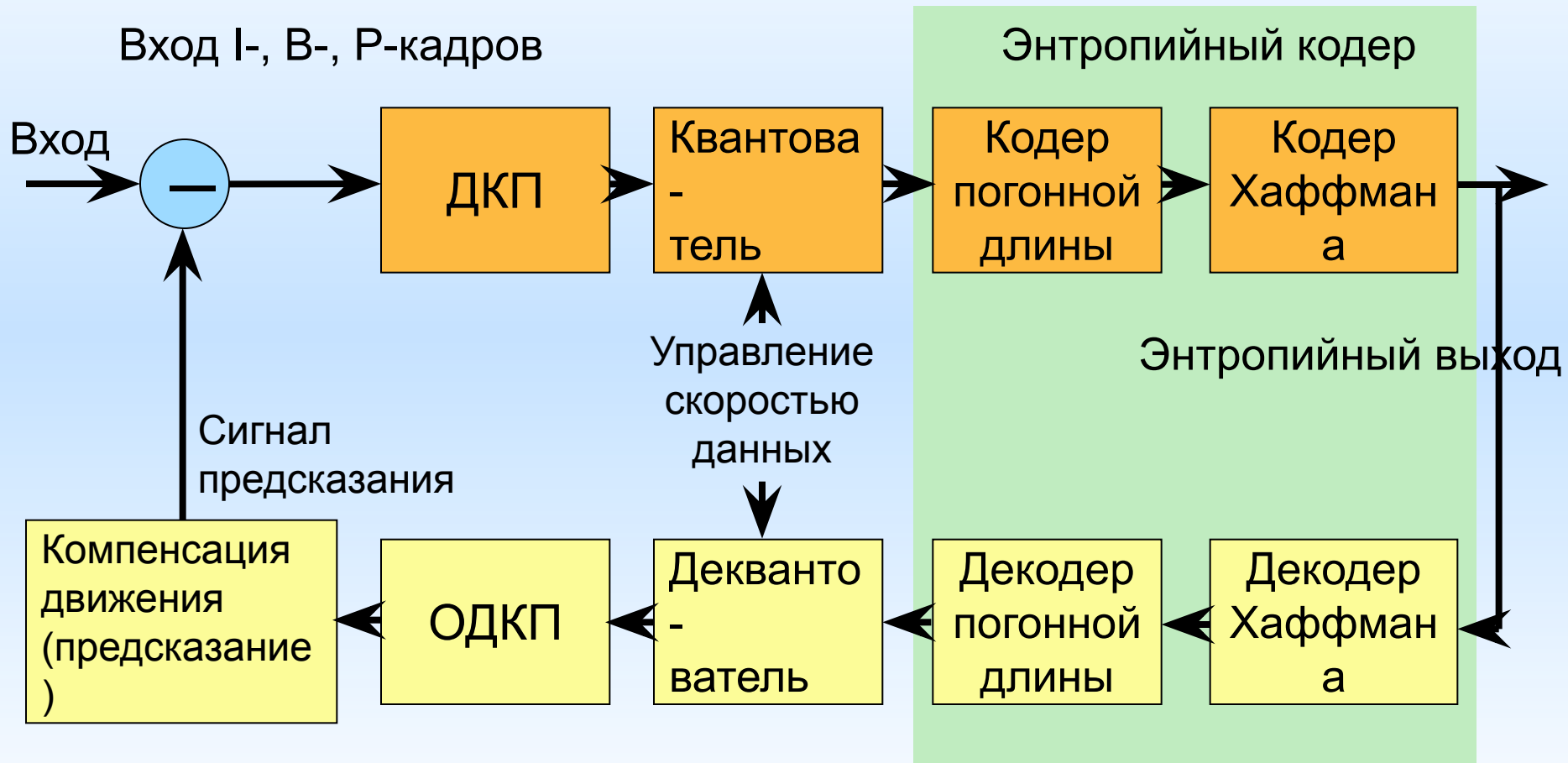
положение макроблока
в предыдущем кадре

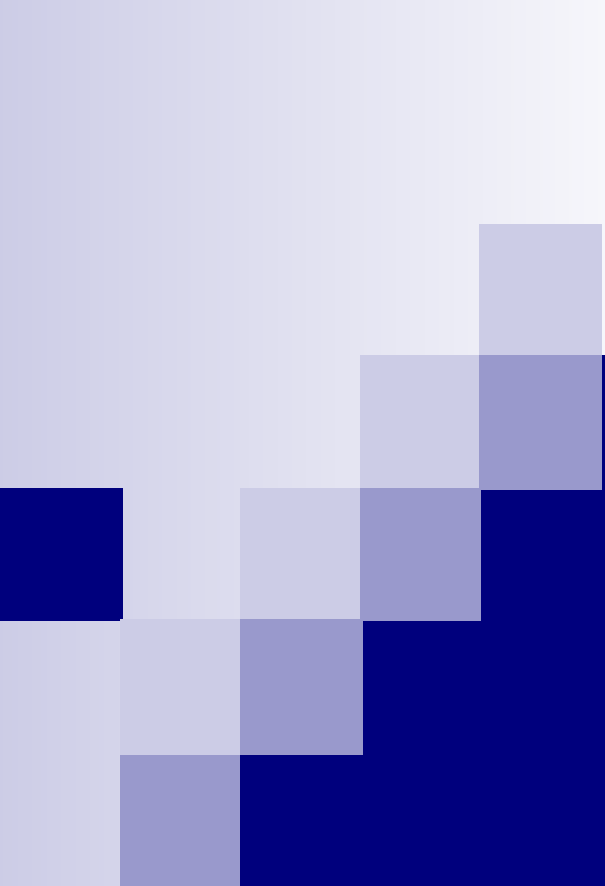


Формирование нового положения макроблока с помощью векторов движения

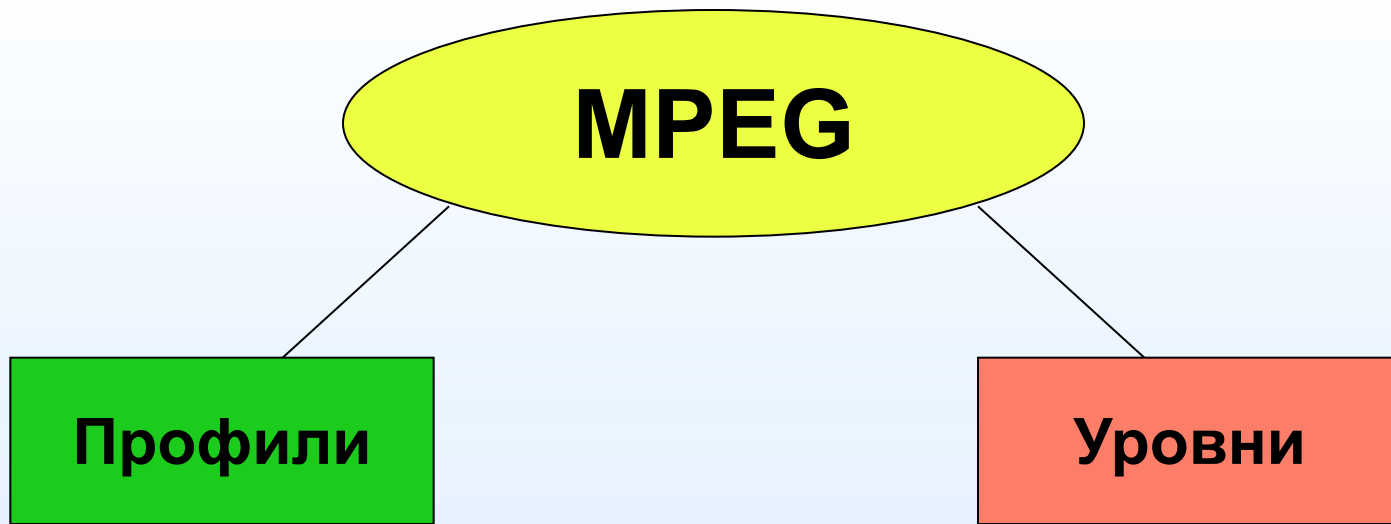


Элементы кодера MPEG





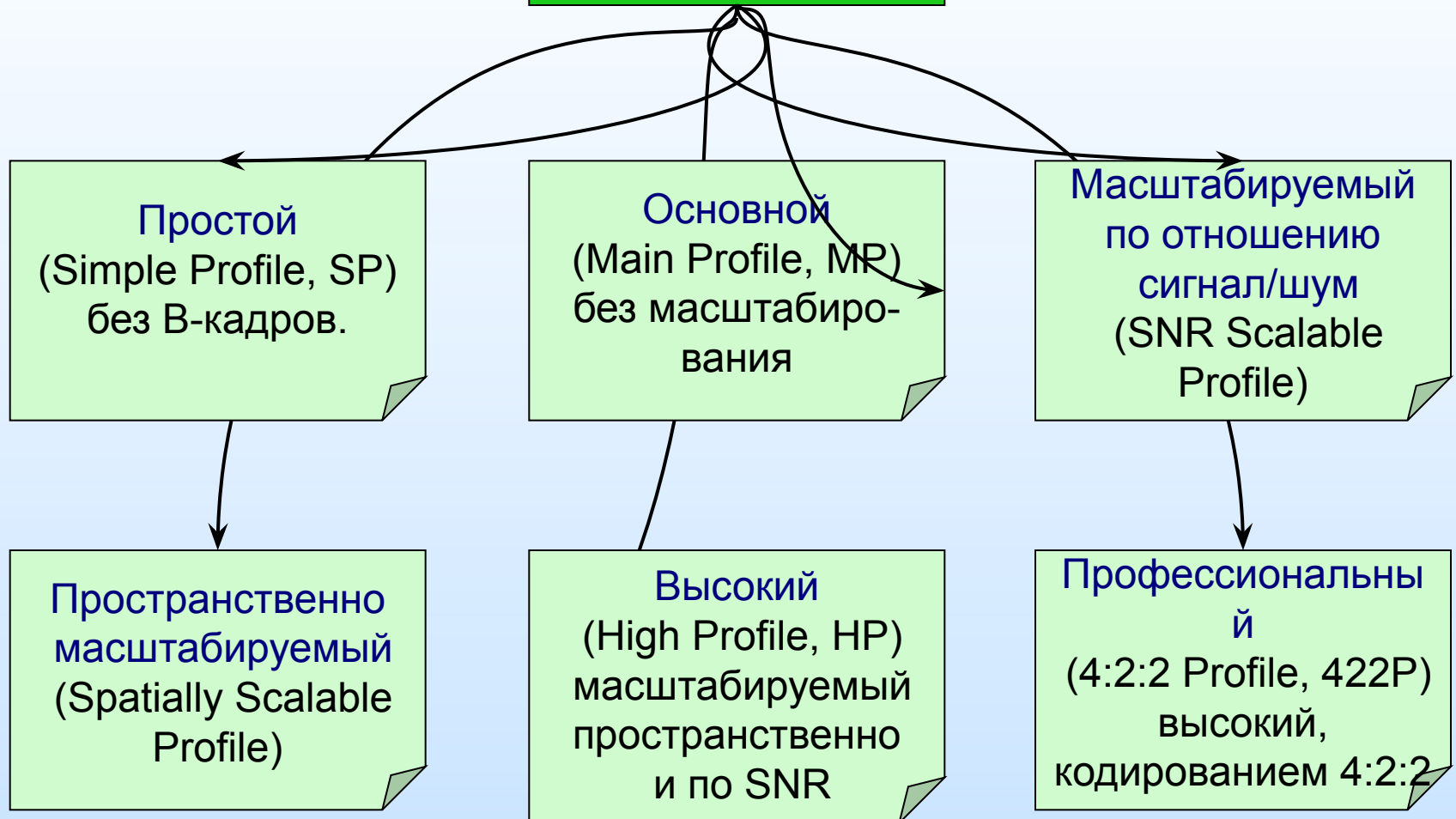
Профили и уровни



Профиль – подмножество структуры битового потока сжатого видео-изображения. Профиль определяет набор методов.

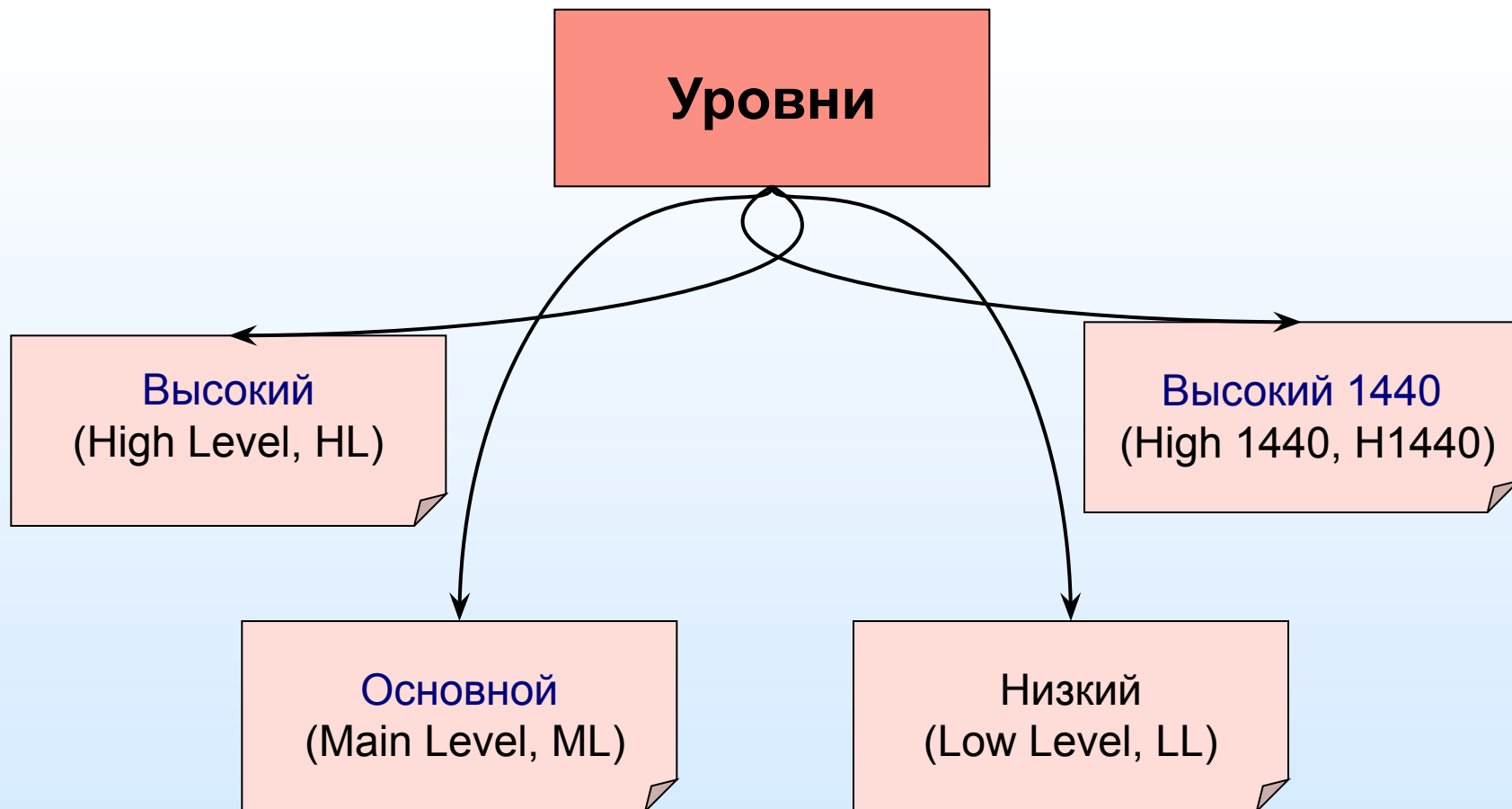
Уровень – ряд ограничений, применяемых к параметрам MPEG-потока в выбранном методе, например, разрешение выходного изображения, частота кадров и т.п.

Профили



Функции профилей MPEG

Профиль Функции	Simple	Main	4:2:2 Profile	SNR	Spatial	High
I-кадры						
P-кадры						
B-кадры						
Разделение по SNR						
Разделение по разрешению						
представлени	4:2:0	4:2:0	4:2:2	4:2:0	4:2:2	4:2:0

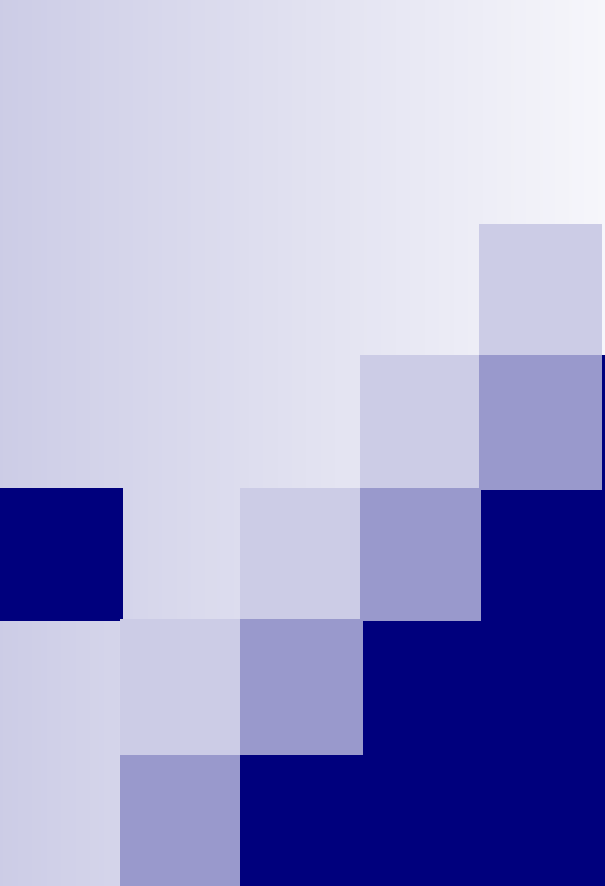


Каждый уровень соответствует тому или иному классу телевизионных систем: уровни **HL** и **H1440** предусмотрены для **HDTV** (high-definition television – телевидения высокой чёткости). Уровень **ML** соответствует **обычному телевидению**, а **LL** так называемому **телевидению ограниченной четкости**.

Максимальные значения ограничений, накладываемых на уровни и профили MPEG.

Возможные комбинации профилей и уровней

Профиль Уровень	Simple	Main	4:2:2 Profile	SNR	Spatial	High
High		1920x1152 80 Мбит/с			1920x1152 100 Мбит/с	
High 1440		1440x1152 60 Мбит/с		1440x1152 60 Мбит/с	1440x1152 80 Мбит/с	
Main	720x576 15 Мбит/с	720x576 15 Мбит/с	720x576 15 Мбит/с		720x608 20 Мбит/с	720x608 50 Мбит/с
Low		352x288 4 Мбит/с	352x288 4 Мбит/с			



MPEG-2

Стандарт MPEG-2 состоит из 9 частей. Первые три части достигли статуса международного стандарта:

- ISO/IEC DIS 13818—1 Информационная технология – Обобщенное кодирование динамических изображений и звукового сопровождения –

Часть 1: Системная часть.

- ISO/IEC DIS 13818—2 Информационная технология - Обобщенное кодирование динамических изображений и звукового сопровождения –

Часть 2: Видеосигнал.

- ISO/IEC 13818—3:1995 Информационная технология - Обобщенное кодирование динамических изображений и звукового сопровождения –

Часть 3: Звуковой сигнал.



Система MPEG-2 выполняет следующие функции:

1. Мультиплексирование отдельных потоков видео и аудио данных в единый поток данных.
2. Обеспечение способов синхронизации компонентов потоков, которые формируют аудио- и/или видеосервисы.
3. Преобразование непрерывного потока бит в пакеты так, чтобы ошибки в этих битах не могли распространяться за границы единичного пакета.
4. Обеспечение наличия программно-зависимой информации (Program Specific Information, PSI) для декодирования входящего потока.

Режимы кодера MPEG-2

```
graph TD; A[Режимы кодера MPEG-2] --> B[с постоянной скоростью потока данных]; A --> C[с постоянным уровнем качества декодированного изображения];
```

с постоянной скоростью
потока данных

Непрерывное изменение коэффициентов матрицы квантования.

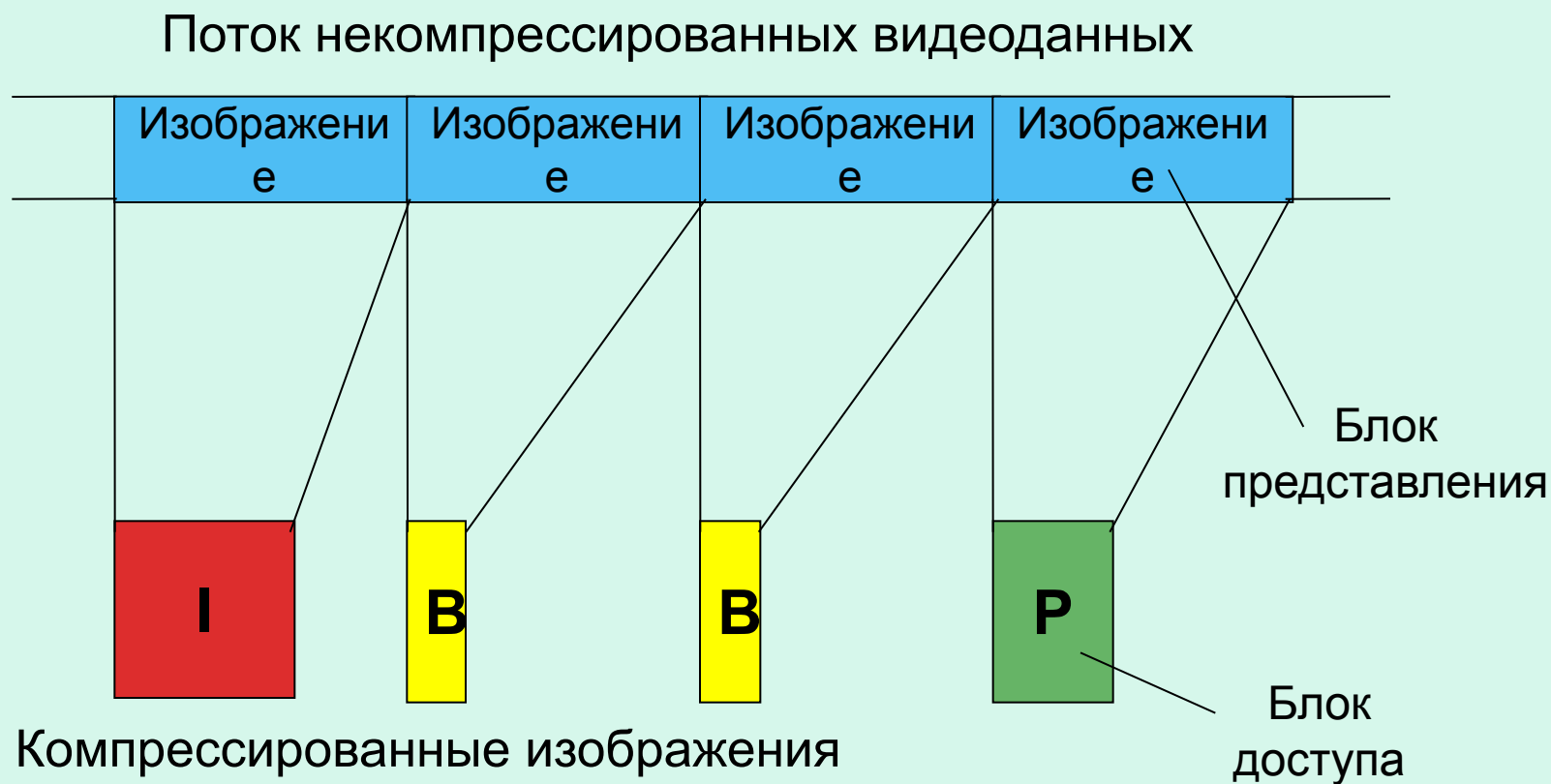
Используется в системах передачи сжатых изображений по каналам связи с фиксированной пропускной способностью, в системах цифрового спутникового, кабельного и наземного телевизионного вещания.

с постоянным уровнем качества
декодированного изображения

Фиксированная матрица квантования.

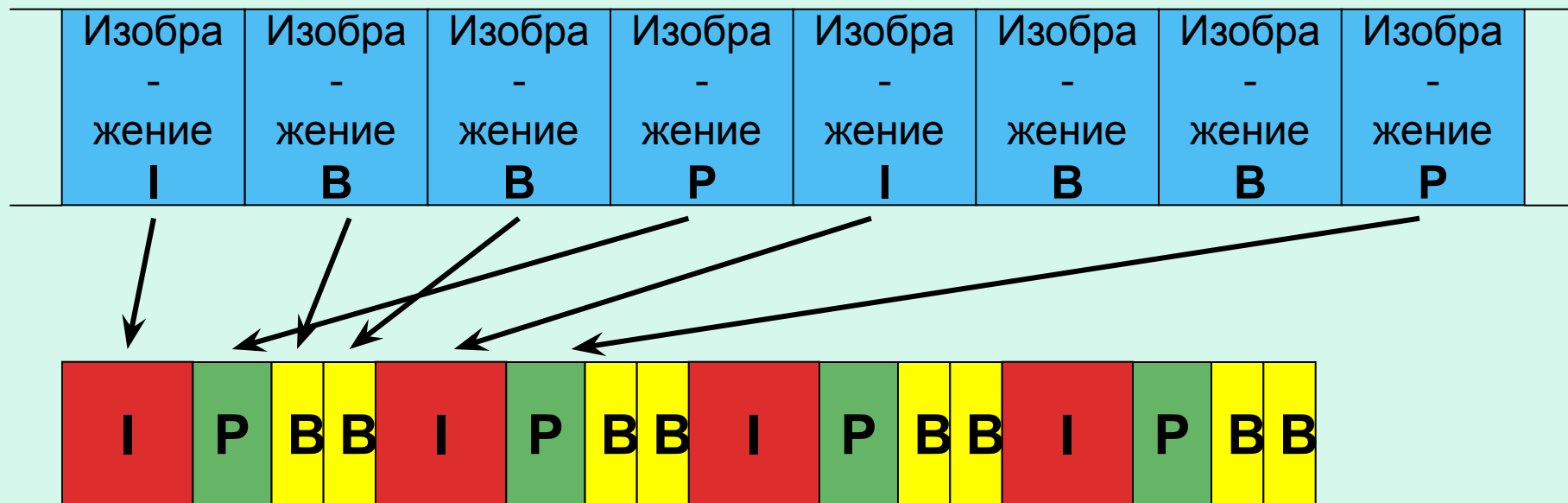
Используется при записи компрессированных потоков видеоданных на дисковые накопители в условиях отсутствия ограничений на объем записанных данных.

В стандарте MPEG-2 изображения (**блоки представления**) рассматриваются как результат декодирования компрессированных изображений - **блоков доступа**.



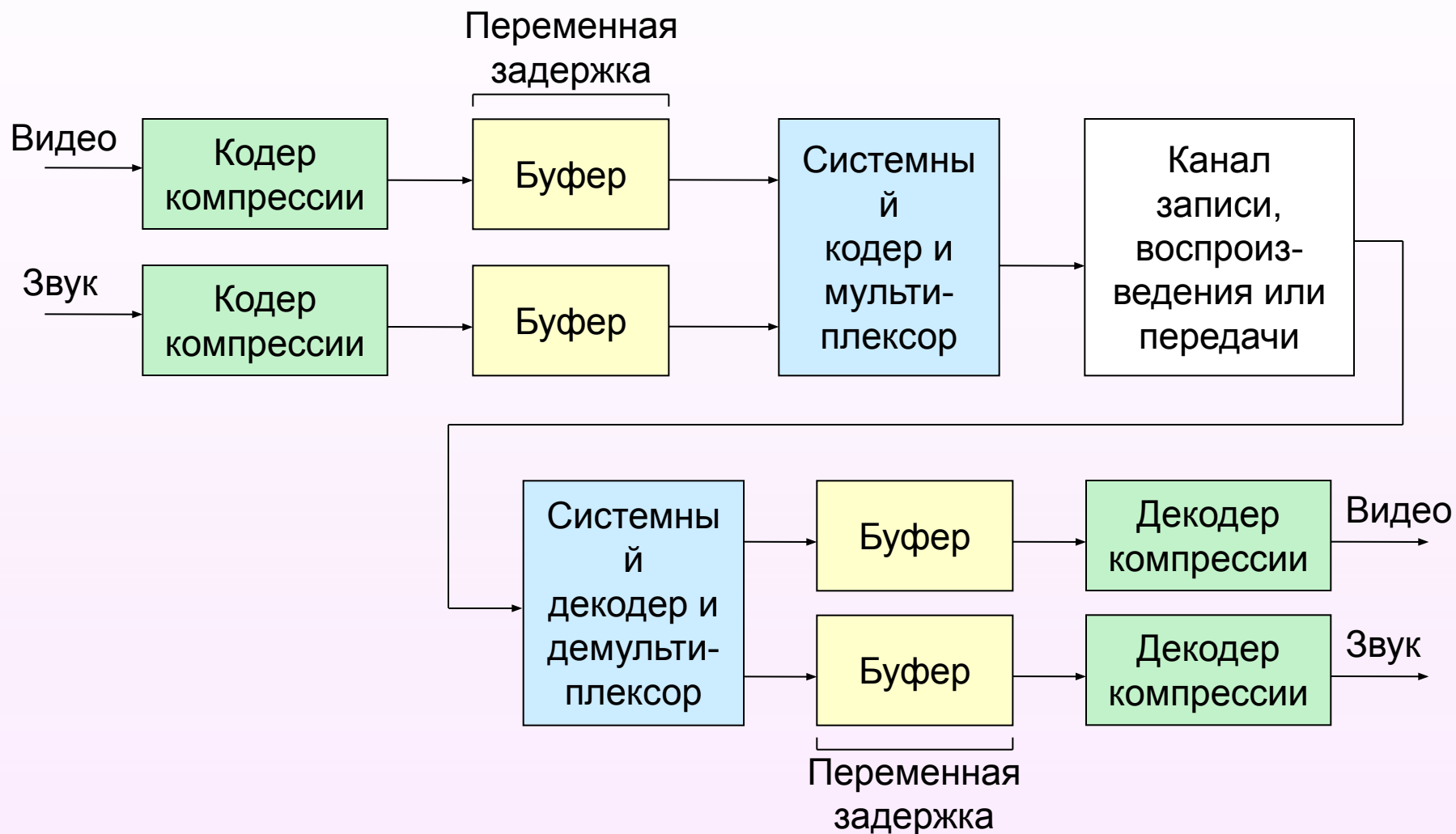
Преобразование блоков представления в блоки доступа в процессе компрессии

Поток видеоданных

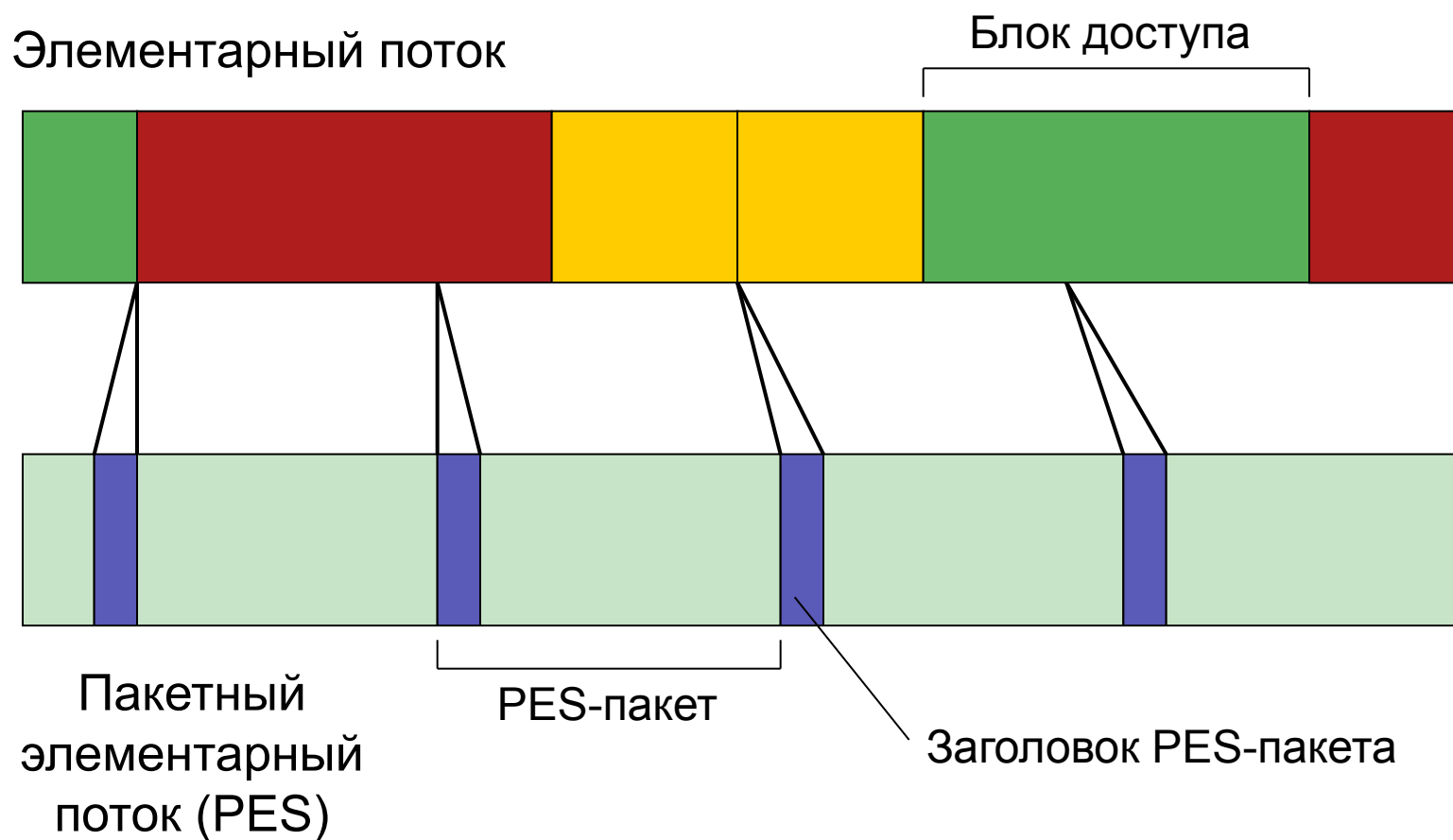


Элементарный поток видеоданных

Синхронизация. Принцип компрессии с постоянной задержкой



Первый шаг на пути получения единого потока – формирование **пакетного элементарного потока PES (Packetised Elementary Stream)**, представляющего собой последовательность PES-пакетов:



РЭС-пакет



Возможные формы единого потока данных MPEG-2

```
graph TD; A([Возможные формы единого потока данных MPEG-2]) --> B[программный поток]; A --> C[транспортный поток];
```

программный поток

транспортный поток

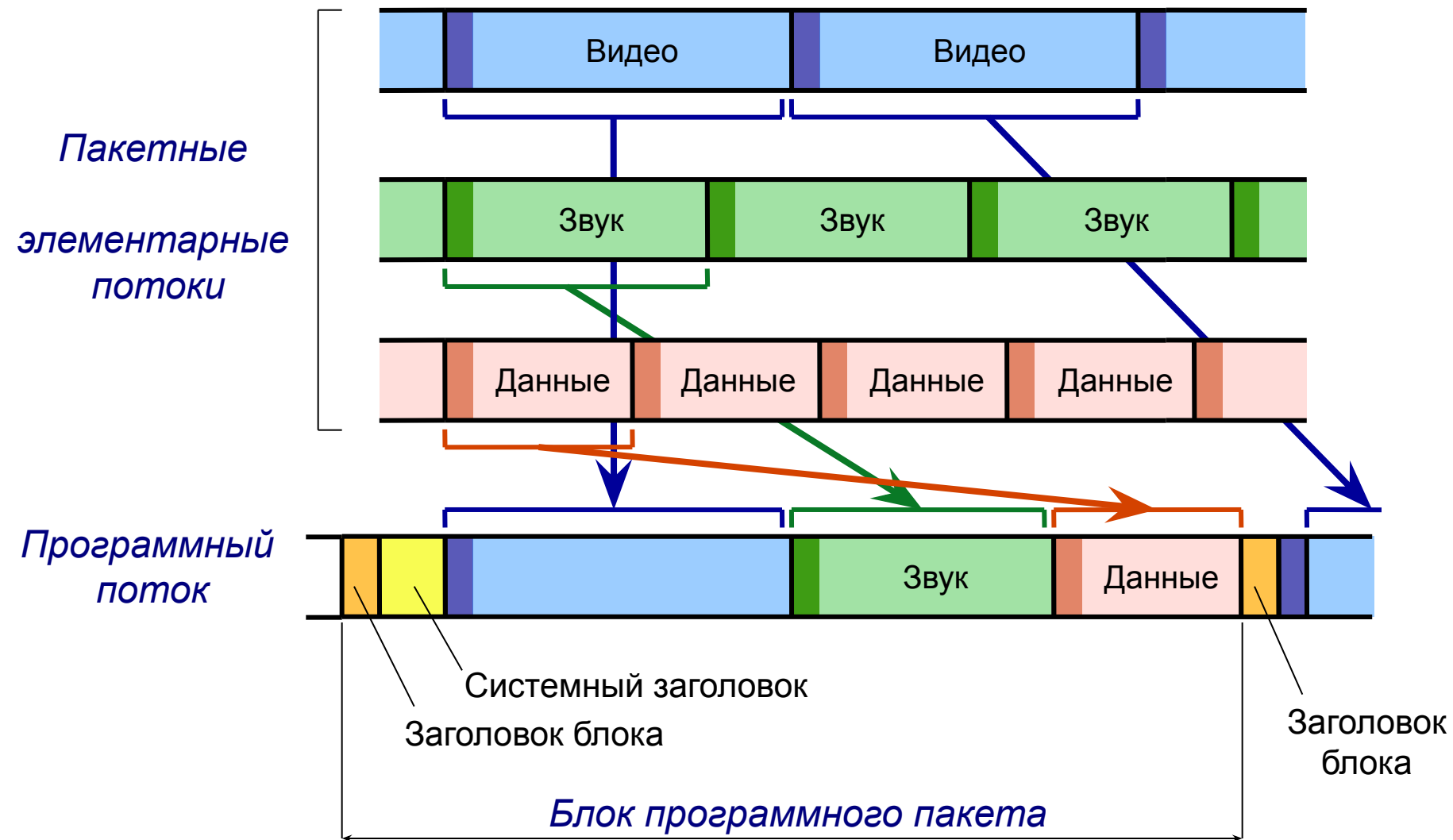
Программный поток

Объединяет элементарные потоки, образующие телевизионную программу. При формировании программного потока образуются блоки из PES-пакетов.

Блок содержит:

- *заголовок блока* – содержит **опорное системное время**, поэтому должен появляться не реже, чем через 0,7 с.
- *системный заголовок* – содержит информацию о характеристиках программного потока (скорость передачи данных, число видео и звуковых элементарных потоков)
- определенное количество *PES-пакетов*

Формирование программного потока



Транспортный поток

Объединяет пакетные элементарные потоки, переносящие данные нескольких программ с независимыми временными базами.

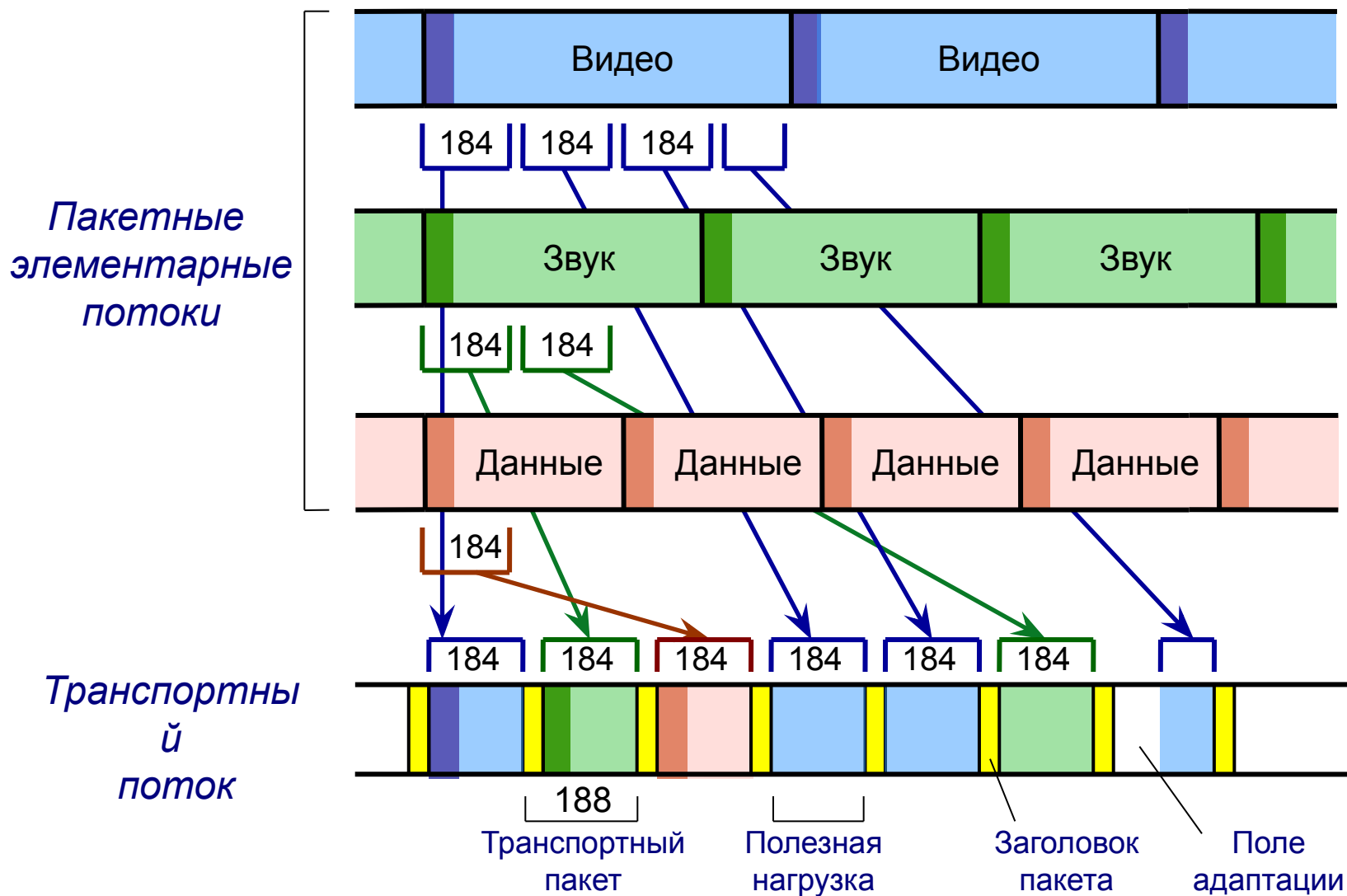
Состоит из коротких пакетов фиксированной длины (188 байтов).

Элементарные потоки видео, звука и дополнительный данных (например, телетекст) разбиваются на фрагменты, равные по длине полезной нагрузке транспортного пакета (184 байта) и мультиплексируются в единый поток.

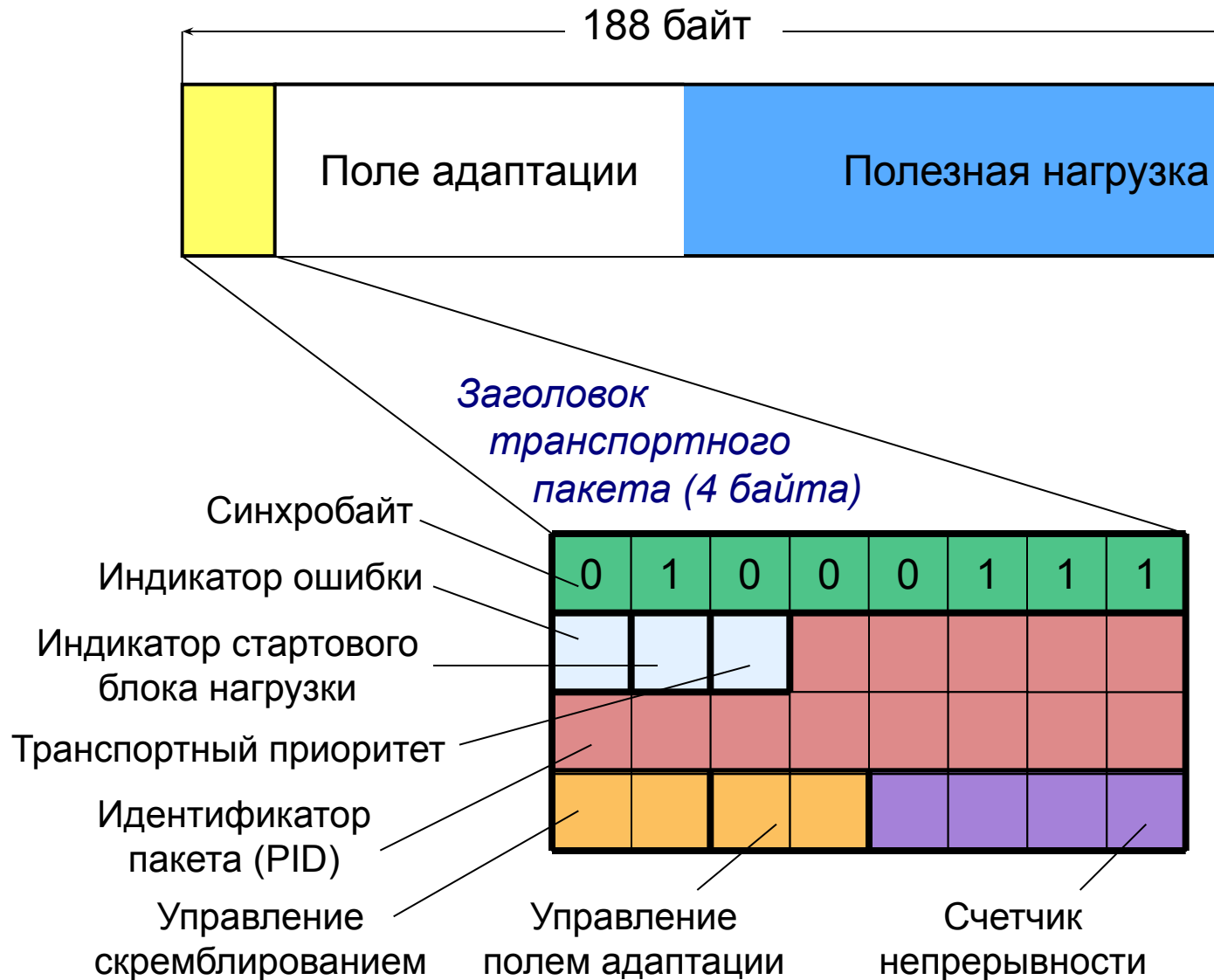
Этот процесс подчиняется ряду ограничений:

- первый байт каждого PES-пакета элементарного потока должен быть первым байтом полезной нагрузки транспортного пакета;
- каждый транспортный пакет может содержать данные лишь одного PES-пакета;
- если PES-пакет не имеет длину, кратную 184 байтам, то один из транспортных пакетов не заполняется данными PES-пакета полностью. В этом случае избыточное пространство заполняется полем адаптации.

Структура транспортного потока



Структура транспортного пакета



Заголовок транспортного потока

Пакетизированные
элементарные
потoki (PES)

Персональные
данные

Программно-зависимая
информация (PSI)

Таблица соединения программ (PAT)
Таблица состава программы (PMT)
Таблица условного доступа (CAT)
Таблица сетевой информации (NIT)

Персональные
данные

Поле адаптации
(Adaptation Field)

Персональные
данные

Иерархическая идентификация программ

Идентификатором принадлежности транспортного пакета к определенному элементарному потоку является значение **PID**.

Для распознавания элементарных потоков и объединения их в телевизионные программы служит **программно-зависимая информация PSI**, которая должна обязательно передаваться в транспортном потоке.

В системной спецификации MPEG-2 определено **4 типа таблиц с программно-зависимой информацией**:

- таблица соединения программ **PAT** (Program Association Table)
- таблица состава программы **PMT** (Program Map Table)
- таблица сетевой информации **NIT** (Network Information table)
- таблица условного доступа **CAT** (Conditional Access Table)

Представление таблицы соединения программ

PID=0

Транспортный пакет с таблицей соединения программ (PAT)

Таблица сообщает список номеров всех программ, которые содержатся в транспортном потоке, и указывает **идентификаторы пакетов**, в которых находятся **PMT-таблицы** с информацией о программах и элементарных потоках, из которых они складываются.

Номер программы **0** зарезервирован, он используется для указания на **PID пакета с сетевой информацией NIT** о сетях передачи транспортного потока, частотах каналов, характеристиках модуляции и т.п.

В данном примере **PID пакета с NIT** равен **16**, а **PID пакета с информацией PMT** о программе **1** равен **21**.

Таблица соединения программ (PAT)

Номер программы	Значение PID для таблицы плана программы (PMT)
0	16
1	21
2	22
•	•
•	•
•	•

Представление таблицы состава программы

PID=21

Транспортный пакет с таблицей
состава программы (PMT)

В данном примере показана PMT
для программы 1 с PID, равным 21.

Таблица PMT указывает сведения о
программе и тех элементарных потоках,
из которых она складывается.

В этом примере элементарный поток
видеоданных программы 1 переносится
пакетами с PID=50, поток звука –
пакетами с PID=51, дополнительных
данных – пакетами с PID=52.

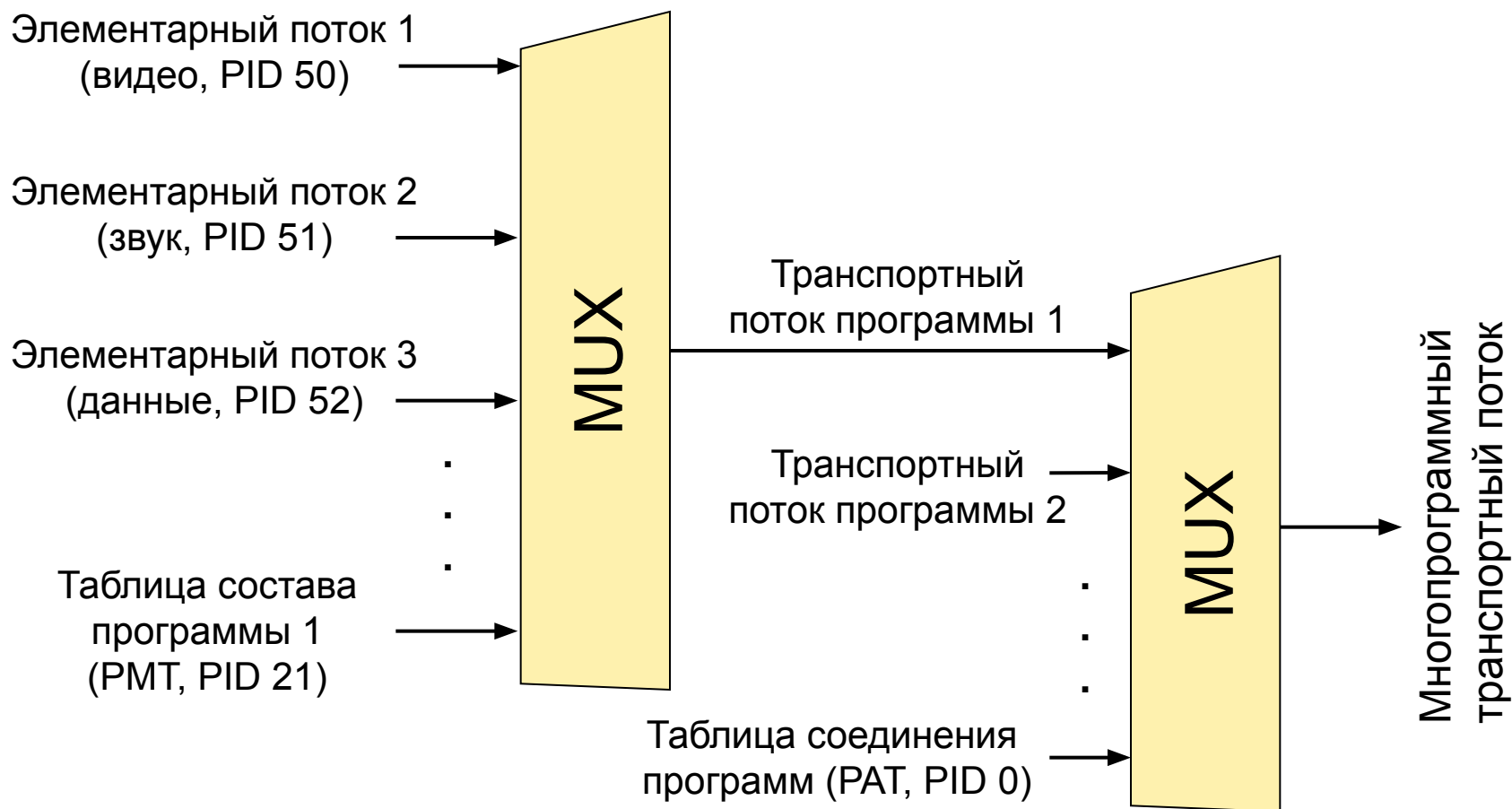
В таблице также указывается PID
транспортных пакетов, переносящих
метки опорного времени данной
программы (обычно эти пакеты имеют
тот же PID, что и поток видео).

Таблица состава
программы (PMT)

Элементарный поток	Значение PID
Видео	50
Звук	51
Данные	52
.	.
.	.
Опорное время	50

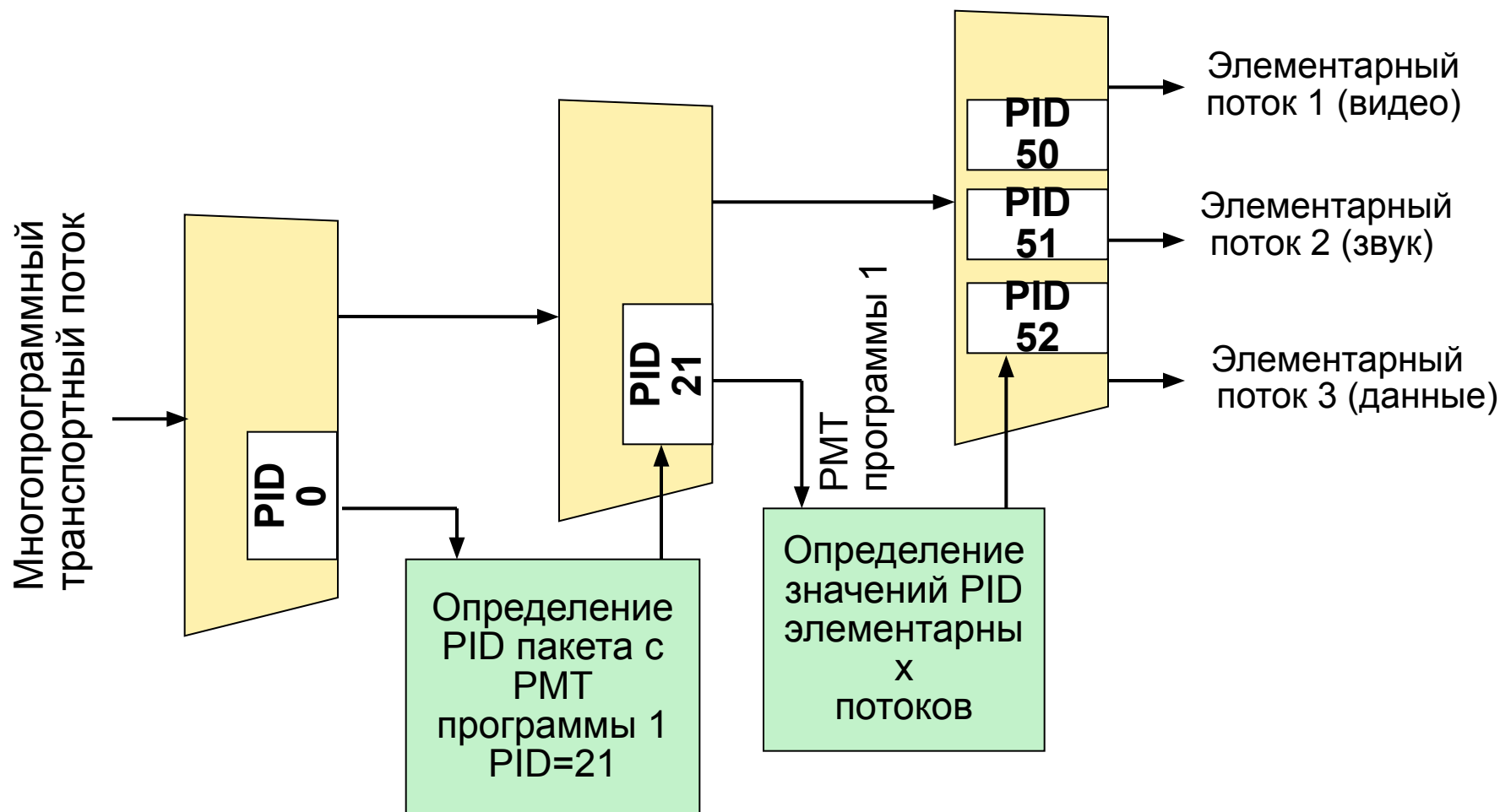
Модель мультиплексирования элементарных и транспортных потоков

(Значения PID соответствуют примерам таблиц, указанным выше)

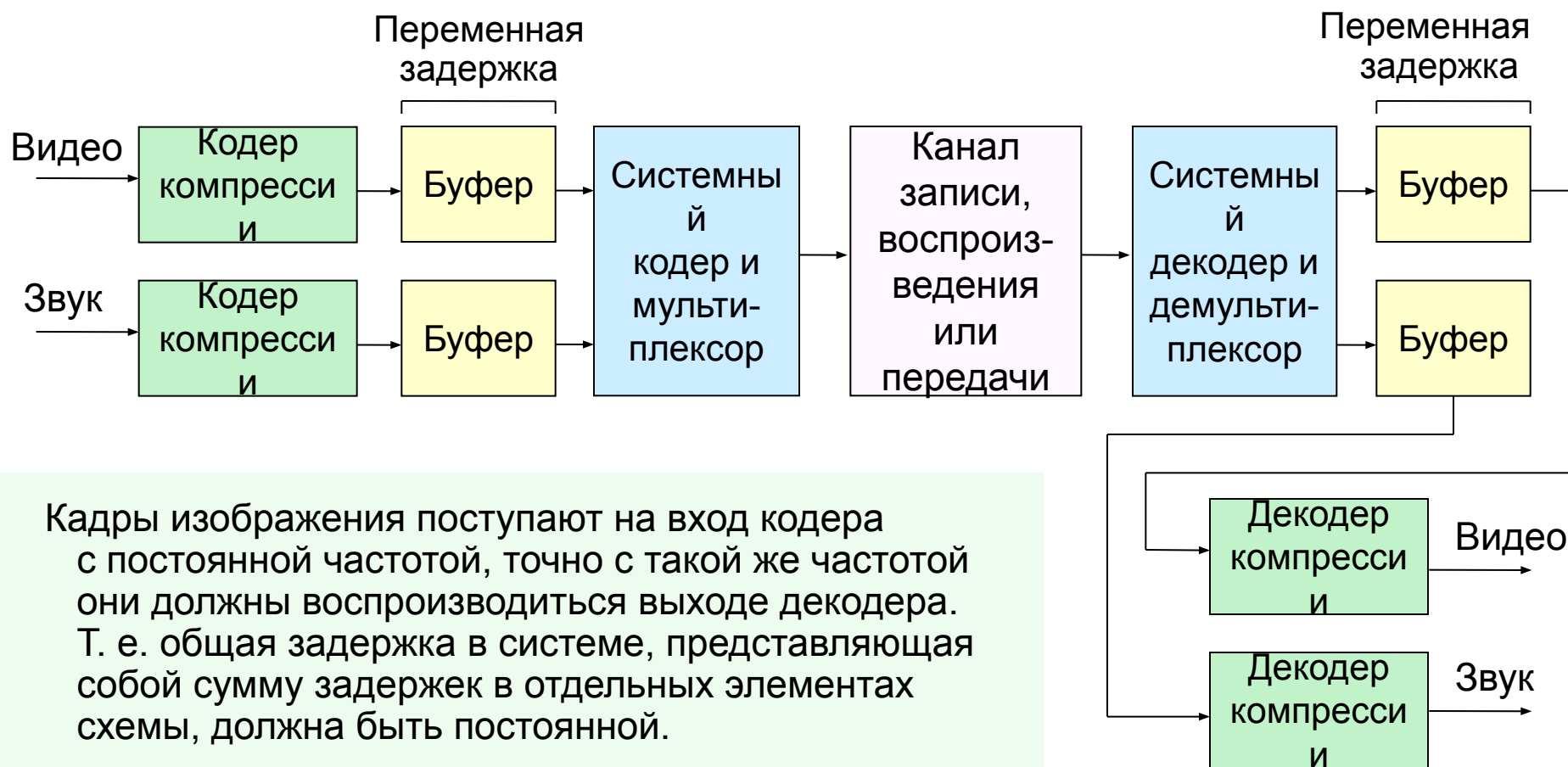


Модель демультимплексирования транспортного потока и получения элементарных потоков программы 1

(Значения PID соответствуют примерам таблиц, указанным выше)



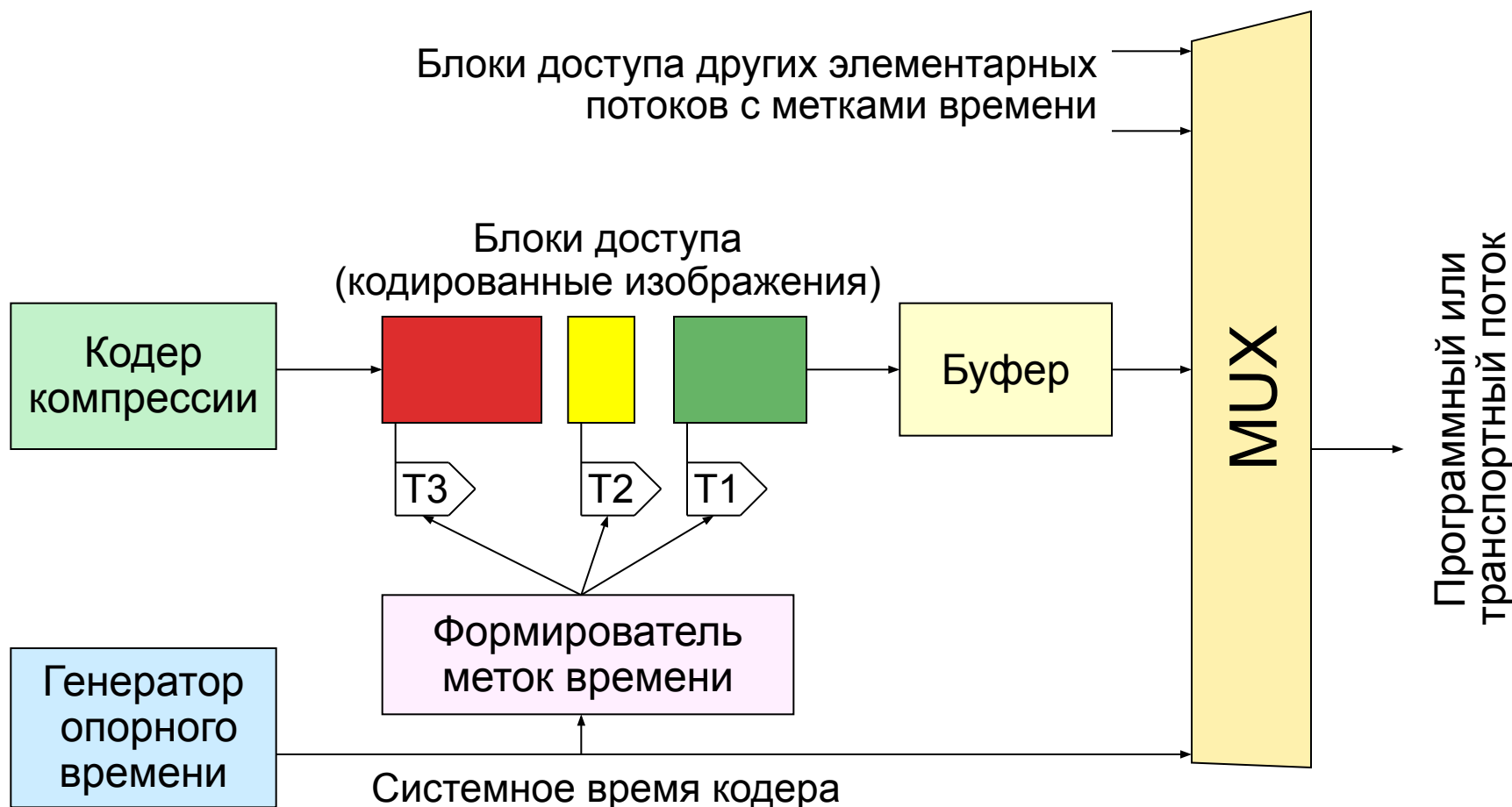
Принцип компрессии с постоянной задержкой



Кадры изображения поступают на вход кодера с постоянной частотой, точно с такой же частотой они должны воспроизводиться выходе декодера. Т. е. общая задержка в системе, представляющая собой сумму задержек в отдельных элементах схемы, должна быть постоянной.

Это достигается за счет использования буфера в кодере и в декодере. Данные поступают в буфер с переменной скоростью, а выходят - с постоянной.

Компенсацию задержек и синхронизацию обеспечивают **метки времени**, которые ставятся в соответствие каждому блоку доступа и сообщают декодеру точное время, когда блок доступа должен быть извлечен из буфера декодера и декодирован.



Метки времени

```
graph TD; A[Метки времени] --> B[Метки времени представления PST]; A --> C[Метки времени декодирования DST];
```

Метки времени
представления PST

Метки времени
декодирования DST

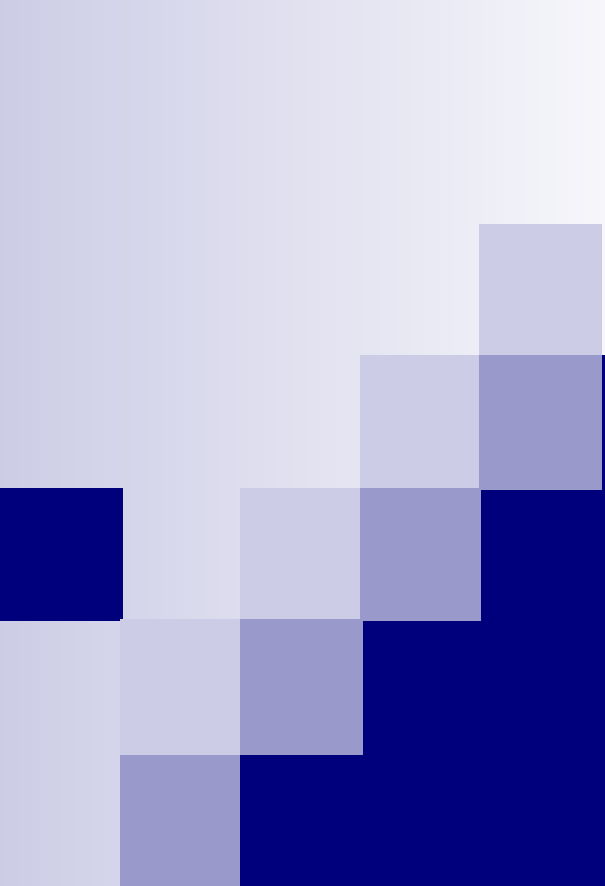
Определяют момент времени, в который декодированный блок доступа (кодированное изображение или звук) должен быть предъявлен зрителю.

Для всех элементарных потоков, кроме видео, PST – это единственные необходимые метки.

Необходимы для потока видеоданных.

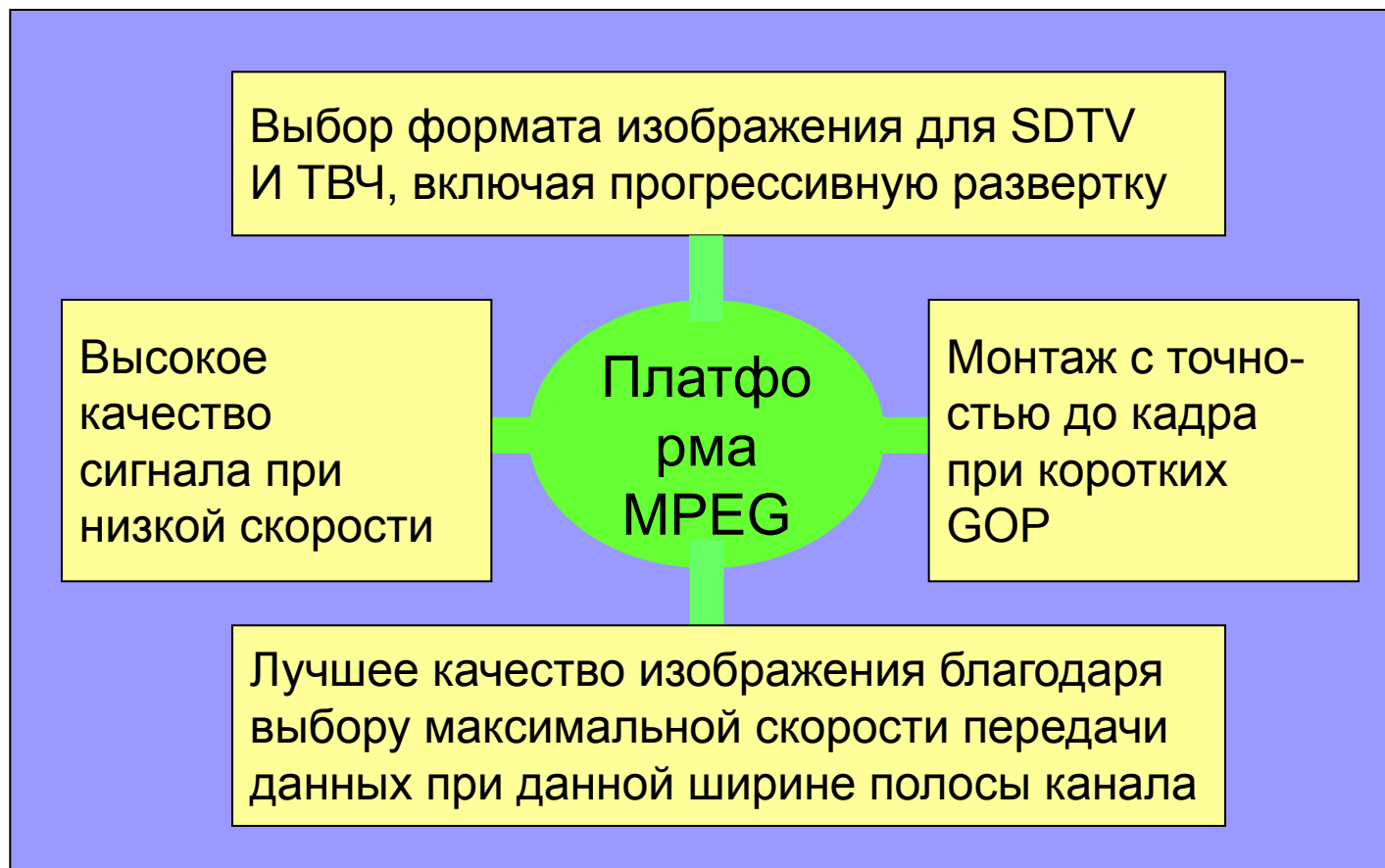
Определяют моменты времени, в которые блоки доступа извлекаются из буфера и декодируются, но не предъявляются зрителю. (Время предъявления назначают метки PST.)

Необходимы изображениям типа I и P, которые должны декодироваться раньше, чем B-изображения.



Применение MPEG-2

Использование MPEG-2 в инфраструктуре телевидения



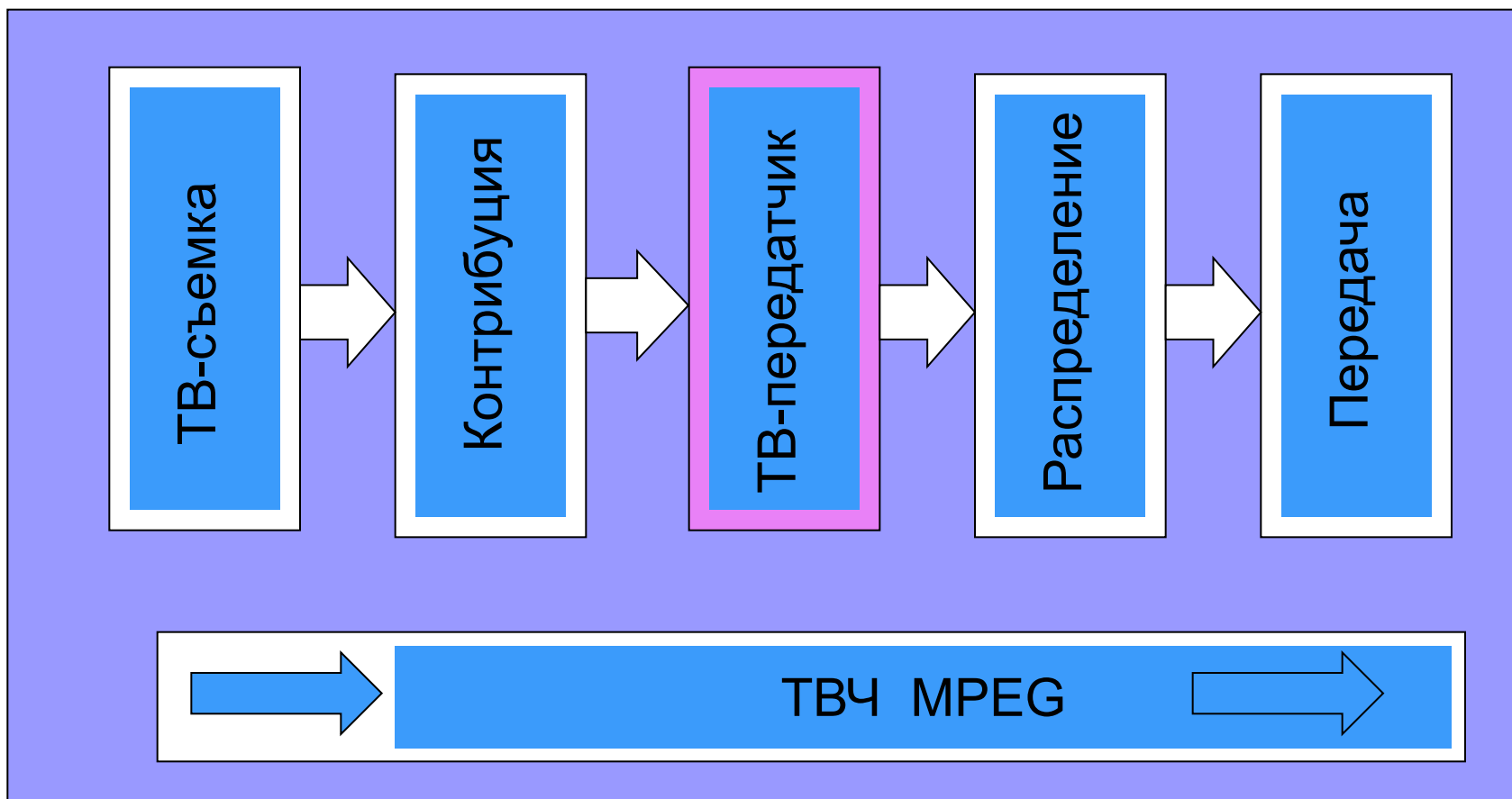
Основные аспекты кодирования MPEG-2 для обработки сигналов



Основные особенности цифровых ТВ-систем

- Эффективность затрат
- Рентабельность
- Качество сигнала
- Возможности применения
- Возможность масштабирования
- Совместимость

Ступени прохождения сигналов в телевидении



Применение MPEG-2

Спутниковое телевидение



Архивирование



Кабельное телевидение



Нелинейный монтаж

