



Лекция 6. Запись и воспроизведение видео информации

1. Понятие видео и характеристики видеосигнала
2. Стандарты видео
3. Представление видеоинформации в компьютере
4. Сжатие и распаковка видеоданных.
5. Этапы создания видеофильмов



1. Понятие видео и характеристики видеосигнала

В́идео (от **лат.** *video* — смотрю, вижу) — электронная технология формирования, записи, обработки, передачи, хранения и воспроизведения **сигналов изображения**, основанная на принципах **телевидения**, а также **аудиовизуальное произведение**, записанное на физическом носителе.



1. Понятие видео и характеристики видеосигнала

Видеозапись — электронная технология записи визуальной информации, представленной в форме видеосигнала или **цифрового потока** видеоданных, на физический **носитель** с целью сохранения этой информации и возможности последующего её воспроизведения и отображения на **устройстве вывода** (**монитора**, экрана или **дисплея**).

Результатом видеозаписи является видеограмма или видеофонограмма.



Характеристики видеосигнала

- **Частота кадров**
- **Стандарт разложения**
- **Разрешающая способность**
- **Соотношение сторон экрана**



Частота кадров

Количество (частота) кадров в секунду — это число неподвижных изображений, сменяющих друг друга при показе 1 секунды видеозаписи и создающих эффект движения объектов на экране. Чем больше частота кадров, тем более плавным и естественным будет казаться движение. Минимальный показатель, при котором движение будет восприниматься однородным — примерно 16 кадров в секунду (это значение индивидуально для каждого человека).



В звуковом **кинематографе** частота съёмки и проекции стандартизирована с 1932 года и составляет 24 кадра в секунду.

Системы телевидения **PAL** и **SÉCAM** используют 25 кадров в секунду (25 fps или 25 **Герц**), а система **NTSC** использует 30 кадров в секунду (точнее 29,97 fps). Компьютерное видео хорошего качества, как правило, использует частоту 30 кадров в секунду. Верхняя пороговая частота мерцания, воспринимаемая **человеческим мозгом**, в среднем составляет 39-42 Герц и индивидуальна для каждого человека, а также зависит от условий наблюдения.



Некоторые современные профессиональные видеокамеры могут снимать с частотой до 120 кадров в секунду. Специальные камеры снимают с частотой до 1000 кадров в секунду.

Сверхскоростные киносъёмочные аппараты могут снимать несколько миллионов кадров в секунду. В них киноплёнка неподвижна и расположена на внутренней поверхности специального барабана, а изображение развёртывается вращающейся призмой.



Существует и бескадровое видео. Принцип работы заключается в следующем: светочувствительные сенсоры с большой частотой передают данные о своём состоянии, которые параллельно записываются на носитель. Отдельных кадров при этом нет — только массивы информации с каждого из датчиков (**пикселей**) об их изменении во времени. При воспроизведении также нет кадров — на экране пиксели меняют свой цвет в соответствии с записанными массивами. Если пиксель не менял цвет, то он и не обновляется. Для наилучшего просмотра такого видео требуется специальный монитор.



Стандарт разложения

Стандарт разложения определяет параметры **телевизионной развёртки**, применяемой для преобразования двумерного изображения в одномерный **видеосигнал** или поток данных. В конечном счёте, от стандарта разложения зависит количество элементов изображения и кадровая частота.

Развёртка может быть **прогрессивной** (построчной) или **чересстрочной**.



При прогрессивной развёртке все горизонтальные линии (строки) изображения отображаются поочередно одна за другой.

При чересстрочной развёртке каждый кадр разбивается на два **поля** (полукадра), каждое из которых содержит чётные или нечётные строки. За время одного кадра передаются два поля, при этом увеличивается частота мерцания **кинескопа** выше физиологического порога заметности.

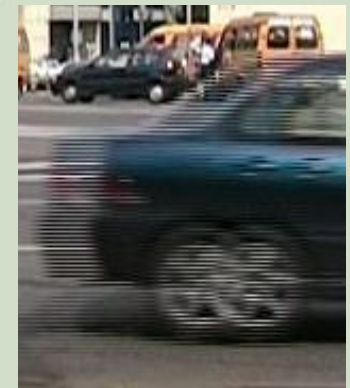


Чересстрочная развёртка была компромиссом, чтобы иметь возможность передачи по каналу с ограниченной **полосой пропускания** изображения с достаточно большой разрешающей способностью.

Несмотря на недостатки, чересстрочная развёртка используется до сегодняшнего дня в **телевидении стандартной четкости** вследствие повсеместного распространения телевизоров, поддерживающих только такие стандарты.



Недостатками чересстрочной развертки являются расщепление вертикальных границ горизонтально движущихся объектов (эффект «гребёнки» или «расчёски») и заметность мерцания на тонких фактурах.



Чересстрочную развёртку часто называют на английский манер *интерлэйс* (**англ. *interlace***) или ***интерлэйсинг***.



Устройства воспроизведения видеоизображения, оснащенные разверткой 100 Гц мерцают с частотой, которая не воспринимается глазом. В таких приемниках изображение с чересстрочной разверткой отображается с удвоением кадров.

ЖК- и **LED-**мониторы (**телевизоры**) вообще избавлены от мерцания. В таких приборах можно говорить только о скорости обновления изображения, поэтому чересстрочная развертка в них является лишь мерой условности, не влияющей на отображение.

Для подавления эффектов, возникающих при просмотре чересстрочного видео на экране с прогрессивной разверткой, применяются специальные алгоритмы **деинтерлейсинга**.



Новые цифровые стандарты телевидения, например, **HDTV** предусматривают прогрессивную развёртку. Новейшие технологии, позволяют имитировать прогрессивную развёртку при показе видео с чересстрочной. Последнюю обычно обозначают символом «i» после указания вертикального разрешения, например 720×576i×50. Прогрессивную развёртку обозначают символом «p», например 720p (означает видео с разрешением 1280×720 с прогрессивной разверткой).

Также для отличия частоты кадров или полей может обозначаться такими же символами кадровая частота, например **24p**, 50i, 50p.



Разрешающая способность

До наступления цифровой эры видео, горизонтальная **разрешающая способность** аналоговой системы видеозаписи измерялась в вертикальных телевизионных линиях (**ТВЛ**) при помощи специальных **телевизионных испытательных таблиц** и обозначала количество элементов в строке видео изображения, зависящее от частотных характеристик устройства записи.

Вертикальная разрешающая способность в изображении заложена в стандарте разложения и определяется количеством строк.



Соотношение сторон экрана

Соотношение ширины и высоты кадра (англ. *aspect ratio*) — важнейший параметр любой видеозаписи. С конца XIX века немые кинофильмы и, в последующем, фильмы «классического» формата, имели соотношение сторон экрана 4:3 (4 единицы в ширину к 3 единицам в высоту; в кинематографе записывается как 1,33:1).



Соотношение сторон экрана

Считалось что, экран с таким соотношением сторон близок к полю зрения человеческого глаза. Появившееся вскоре телевидение переняло это соотношение и почти все аналоговые телесистемы (и, следовательно, **телевизоры**) имели соотношение сторон экрана 4:3. Первые **компьютерные мониторы** также унаследовали телевизионный стандарт соотношения сторон.



Соотношение сторон экрана

В кинематографе уже в начале **1950-х годов** появляется **панорамное, широкоэкранное и широкоформатное кино.**

Широкоэкранные кинематографические системы обладали соотношениями сторон до **2,75:1**, стремясь к максимальному «эффекту присутствия», чтобы сделать границы кадра малозаметными. Главная причина заключается в том, что поле бинокулярного зрения человека приближается к соотношению **2:1**.



Соотношение сторон экрана

Смена соотношений сторон экрана требовала или обрезки изображения при помощи **пансканирования**, или добавления пустых полей сверху и снизу, чтобы вписать фильм в экран **леттербоксингом**.

Оба способа приводили к потерям частей изображения или его качества. На сегодняшний день классический формат 1,37:1 вообще не используется в кинематографе, полностью уступив своё место **кашетируемому** кадру 1,85:1.



Соотношение сторон экрана

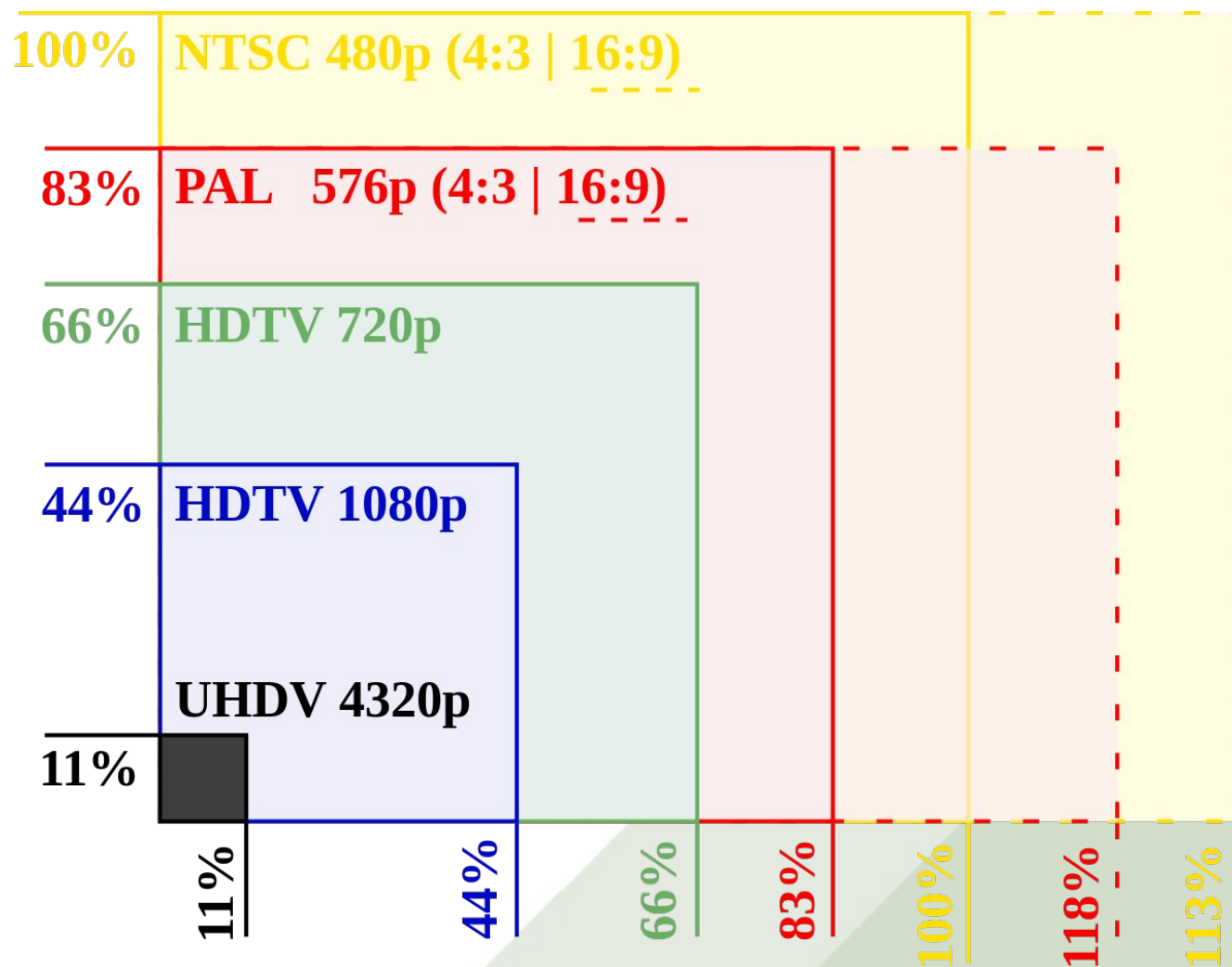
При выборе соотношения сторон экрана **телевидения высокой чёткости** был одобрен стандарт 16:9 (1,78:1), более близкий распространённым форматам кино. **Цифровое телевидение** стандартной чёткости в основном так же ориентируется на соотношение 16:9, применяя **цифровое анаморфирование**. Всё это, по замыслу создателей, призвано глубже погрузить зрителя в атмосферу просматриваемого видеофильма.

Есть и альтернативные объяснения переходу на широкий формат: стремление к ухудшению качества пиратских видеокопий и телевизионных копий.



Цифровое анаморфирование — технология передачи и записи широкоэкранного изображения цифрового телевидения при помощи стандартов разложения, изначально рассчитанных на классическое соотношение сторон экрана 4:3.

При этом информационная ёмкость такого кадра используется наиболее эффективно за счёт трансформации соотношения сторон пикселя. Кроме цифрового телевидения стандартной чёткости технология применяется при мастеринге DVD дисков и является цифровым аналогом оптического анаморфирования.



Соотношение сторон пикселя разных стандартов с анаморфированием и без.
Пунктиром обозначен размер пикселя **576i** и **480i** анаморфированного видео



2. Стандарты видео

PAL, NTSC, SECAM - это видеостандарты цветного телевидения, принятые в различных странах.

В 1953 году в США появился первый из стандартов - система **NTSC**. Разрешение экрана для этого стандарта составляет 720x480 пикселей, частота 30 кадров в секунду (kps). Такая частота обусловлена тем, что промышленный ток в США имеет частоту 60 Гц.

Система NTSC распространена в США, Канаде, Японии, Мексике и некоторых других странах.

В 1956 году, Франция разработала стандарт **SECAM**. Поскольку частота промышленного тока в Европе составляет 50 Гц, частота кадров для SECAM - 25 kps.

В 1967 году началось вещание цветного телевидения данного стандарта во Франции и Советском Союзе.



Стандарт **PAL** разработан в Германии в 1967 году и используется во всех странах Европы кроме вышеназванных, а также в Азии, Австралии и некоторых странах Африки и Южной Америки.

Стандарты PAL и SECAM имеют одинаковую частоту кадров - 25 kps и одинаковое разрешение экрана - 720x576



Стандарты различают по числу строк: PAL и SECAM - 625 строк, NTSC - 525 строк. Не все строки активные.

Для PAL из 625 строк активными являются 576, а оставшиеся - вспомогательные, нужны для установки луча на начало и подготовки к приему новой последовательности строк. Отображение кадра на экран происходит в два прохода, а, значит, кадр состоит из двух **полукадров (полей)**.

Каждый полукадр длится $1/50$ секунды, а, соответственно, кадр - $1/25$ (для стандарта PAL).

Значит, частота кадров составляет 25 kps.

Для NTSC длительность полукадра $1/30$ секунды, кадра - $1/30$ сек.



DVB-T (Digital Video Broadcasting — Terrestrial) - стандарт для передачи **цифрового** телевидения и радио, используемый в Европе, Азии, Африке, Австралии.

В Северной Америке аналогичный стандарт называется ATSC, в Японии - ISDB, в Китае — DMB-T/H.



Основные преимущества цифрового телевидения:

более высокое качество, устойчивость сигнала к помехам, уменьшение в несколько раз мощности потребления электроэнергии цифровым передатчиком, возможность трансляции Full-HD, применения стереозвука, увеличение количества программ на одном частотном канале.



3. Представление видеоинформации в компьютере

Когда говорят о видеозаписи, прежде всего имеют в виду движущееся изображение на экране телевизора или монитора.

Преобразование оптического изображения в последовательность электрических сигналов осуществляется видеокамерой. Эти сигналы несут информацию о яркости и цвете отдельных участков изображения. Они сохраняются на носителе в виде изменения намагниченности видеоленты (аналоговая форма) или в виде последовательности кодовых комбинаций электрических импульсов (цифровая форма).



Процесс превращения непрерывного сигнала в набор кодовых слов называется аналого-цифровым преобразованием:

- **дискретизации**, когда непрерывный сигнал заменяется последовательностью мгновенных значений через равные промежутки времени;
- **квантования**, когда величина каждого отсчёта заменяется округлённым значением ближайшего уровня;
- **кодирования**, когда каждому значению уровней квантования, полученных на предыдущем этапе, сопоставляются их порядковые номера в двоичном виде.



По своей сути **видеофайл — это набор статичных изображений**, меняющих друг друга с определенной частотой. Каждое статичное изображение является отдельным кадром видео. Это действительно так, если мы говорим о несжатом видео. Однако в таком формате никто не хранит фильмы.

Кадр видео формата **PAL** состоит из **720** точек по горизонтали и **576** по вертикали. То есть один кадр состоит из **414720** точек.

Для хранения цвета каждой точки в памяти отводится **24** бита (по **8** бит для каждой из составляющих RGB).

Следовательно, для хранения одного кадра понадобится **9 953 280** бит (или примерно **1,2** Мбайт).

То есть секунда несжатого видео в формате **PAL** будет занимать почти **30** Мбайт. А один час такого видео — более **100** Гбайт.



В основном, видео хранят в видеофайлах, в которых применены различные алгоритмы сжатия информации. Благодаря этим технологиям видеофайл можно сжимать в десятки и сотни раз практически без потери качества картинки и звука.



AVI (Audio Video Interleave) — это контейнерный формат, поэтому в нём могут содержаться аудио/видео, сжатые различными комбинациями кодирования.

Контейнер — это файл с каким либо расширением, служащий для хранения в цифровом виде преобразованной аналоговой информации. Контейнер является файлом некоего стандарта, в котором одновременно может содержаться несколько различных типов информации.



AVI файл может содержать различные виды компрессированных данных (например, **DivX** для видеоинформации и **MP3** для аудио), в зависимости от того, какой кодек используется для кодирования/декодирования. В файле с расширением **AVI** может храниться несжатое видео, видео в форматах **DV**, **MPEG-4**, **DivX**, **Xvid** и даже **MPEG-1** и **MPEG-2**. Кроме того, файл формата **AVI** может, например, содержать в себе только звук. То есть файлы формата **AVI** являются контейнером для хранения данных различного типа.



DivX — технология видеозаписи, позволяющая создавать и просматривать медиа файлы с высокой степенью сжатия. Это технология сжатия, которая делает фильм в **8-12** раз меньшим с небольшой потерей качества.

DivX широко используется для сжатия компьютерных видеофайлов и файлов **DVD**, чтобы они помещались на стандартный CD.



MPEG (Moving Picture Expert Group) — формат, предназначенный для сжатия звуковых и видеофайлов для загрузки или пересылки, например, через Интернет.

Разработан *Экспертной группой кинематографии*, которая занимается разработкой стандартов кодирования и сжатия видео- и аудиоданных.

Существуют разные

стандарты **MPEG: MPEG-1, MPEG-2, MPEG-3, MPEG-4, MPEG-7.**

Например, **MPEG-4** — стандарт, благодаря которому фильмы можно кодировать в хорошем качестве при низкой скорости передачи данных. Стандарт **MPEG-4** в основном используется для вещания (потокковое видео), записи фильмов на компакт-диски, видеотелефонии (видеотелефон) и широковещания, в которых активно используется сжатие цифровых видео и звука.



Для записи видеоинформации на магнитную ленту используются форматы:

- **VHS** (Video Home System) не более 240 строк при аналоговом способе записи;
- **S-VHS** (Super VHS) усовершенствованная VHS, разрешение до 400 строк;
- **Video8**, цифровая запись методом импульсно-кодовой модуляции, 260 строк;
- **Hi-8** , усовершенствование Video8, до 400 строк.



Для записи аудио-видео CD используются форматы:

- **VCD** (Video Compact Disc);
- **SVCD** (Super VCD);
- **DVD** (Digital Compact Disc);
- **BD** (Blue-Ray Disc).



4. Сжатие и распаковка видеоданных

Концепция одновременной записи видео и звука в формате AVI (Audio Video Interleave) – заключается, запись аудио- и видеосигналов в одном файле отдельно для каждого кадра – сначала звук первого видеокадра, затем изображение, затем звук следующего кадра и т.д. При этом информация внутри кадров или последовательности кадров может быть не сжата или скомпрессована.



**Для изображений достаточно передавать
только разностную информацию:**

- о положении объекта, направлении и величине смещения;
- о новых элементах фона (открывающихся за объектом по мере его движения).



В MPEG-кодировке принципиально формируются три типа кадров:

- **I (Intra)**, выполняющие роль опорных и сохраняющие полный объем информации о структуре изображения;
- **P (Predictive)**, несущие информацию об изменениях в структуре изображения по сравнению с предыдущим кадром (типов I или P);
- **B (Bi-directional)**, сохраняющие только самую существенную часть информации об отличиях от предыдущего и последующего изображений (только I или P).



Последовательности I-, P- и B-кадров объединяются в фиксированные по длине и структуре группы кадров - **GOP** (*Group of Pictures*).

GOP - начинается с I-кадра и с определенной периодичностью содержит P-кадры. Ее структуру описывают как M/N , где M – общее число кадров в группе, а N – интервал между P-кадрами.



Фазы общего стандарта

- **MPEG-1** предназначен для записи синхронизированных видео- и звукового сопровождения на CDROM с учетом скорости считывания 1,5-3,5Мбит/сек;
- **MPEG-2** предназначен для обработки видеоизображения, соизмеримого по качеству с телевизионным при пропускной способности системы передачи данных в пределах 3-15 Мбит/сек;



Фазы общего стандарта

- **MPEG-3** предназначался для использования в системах телевидения высокой четкости HDTV при разрешении 1920x1080 при 30 кадр/сек со скоростью потока данных 20-40 Мбит/сек;
- **MPEG-4** является стандартом для низкоскоростной передачи и задает принципы работы с цифровым представлением медиаданных для областей;
- **MPEG-7** обеспечивает стандарт для описания различных типов мультимедийной информации, чтобы обеспечивать эффективный и быстрый ее поиск. MPEG-7 официально называют **Multimedia Content Description Interface**

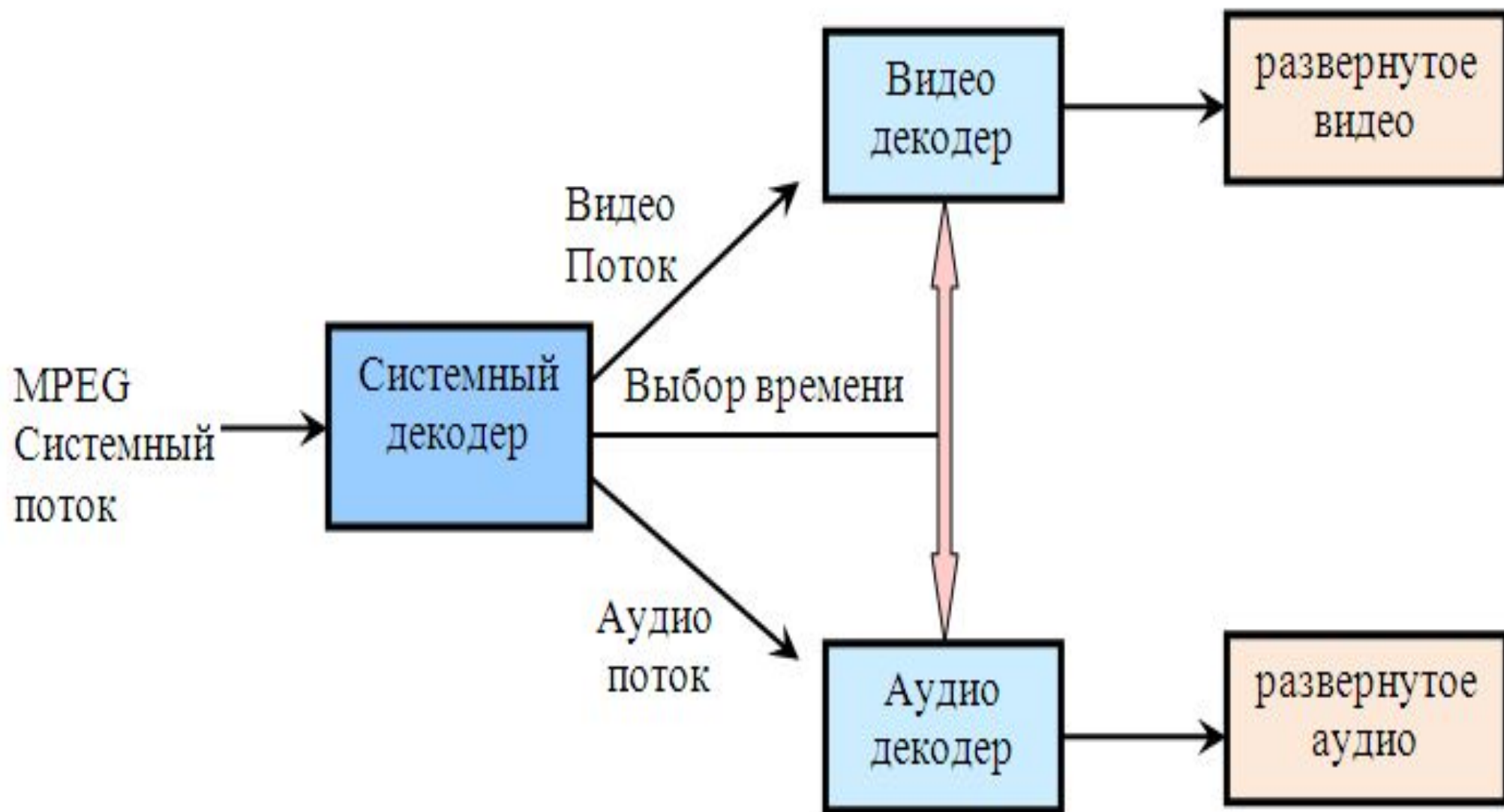


Синхронизация и объединение звука и видео, осуществляется с помощью системных потоков, который включает в себя:

- **Системный слой**, содержащий временную и иную информацию с целью разделения и синхронизации видео и аудио;
- **Компрессионный слой**, содержащий видео и аудио потоки.



Принцип синхронизации и объединения звука и видео





Для синхронизации аудио и видео потоков в системный поток встраивается таймер, работающий с частотой 90кГц SCR (*System Clock Reference*) :

- метка, по которой происходит увеличения временного счетчика в декодере;
- метка начала воспроизведения (PDS, *Presentation Data Stamp*) вставляются в изображения или в звуковой кадр, чтобы «объяснить» декодеру, когда их воспроизводить.



MPEG-2

В MPEG-2 внесены изменения в аудиоканал:

- Реализованы новые виды частот 16 кГц, 22,05 кГц, 24 кГц;
- Добавлена поддержка многоканальности;
- Появился стандарт **AAC** (*Advanced Audio Coding, прогрессивное кодирование звука*), обеспечивающий очень высокое качество звука со скоростью 64 Кбит/сек на канал.

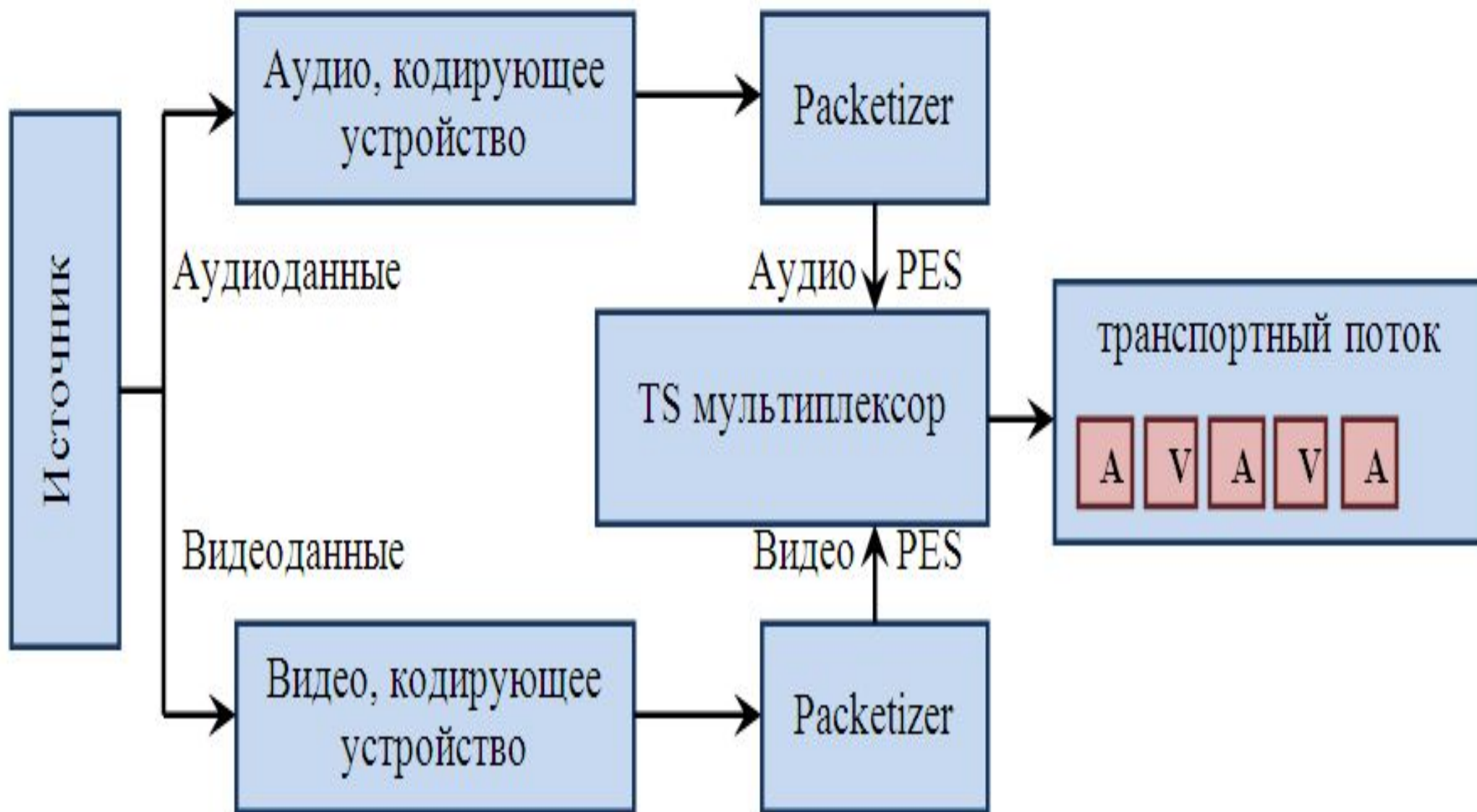


Системный уровень MPEG-2 обеспечивает два уровня объединения данных:

1. Первый уровень - **PES (Packetized Elementary Stream)** - разбивает звук и видео на пакеты.
2. Второй уровень делится на:
 - **MPEG-2 Program Stream** (совместим с MPEG-1 System) - для локальной передачи в среде с низким уровнем ошибок;
 - **MPEG-2 Transport Stream** - внешнее вещание в среде с высоким уровнем ошибок, передает транспортные пакеты двух типов (сжатые данные - PES - и сигнальную таблицу PSI).



Передача транспортных пакетов по стандарту MPEG-2





5. Этапы создания видеофильмов

- ☐ реальная съемка (или компьютерная генерация) первичных кадров;
- ☐ монтаж кадров (обрезка и вставка кадров в нужное место);
- ☐ генерация титров и спецэффектов;
- ☐ редактирование звука;
- ☐ сборка готового фильма и конвертирование его в нужный цифровой формат;
- ☐ запись в аналоговом или цифровом виде на внешние носители (магнитную ленту, CD или DVD).