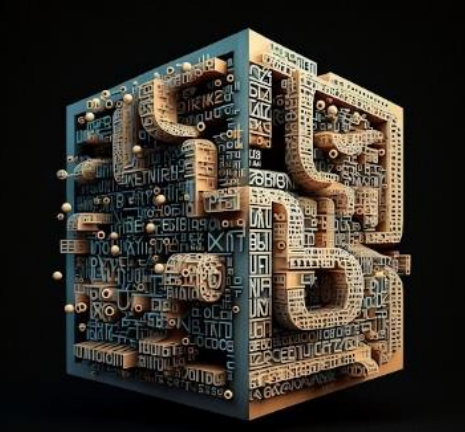


**Міністерство освіти та науки України
Харківський національний економічний
університет імені Семена Кузнеця**



**НАВЧАЛЬНО – НАУКОВИЙ ІНСТИТУТ
ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ**
кафедра: Мультимедійних систем і технологій

дисципліна
"ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ"
викладач: доцент, к.т.н.
Токарєв Володимир Володимирович



Тема. Теоретичні основи мультимедіа. Цифрове відео.

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ



Цифрове відео є одним із ключових напрямів розвитку сучасних інформаційних технологій, мультимедійних систем і цифрових комунікацій. Його використання охоплює:

- телебачення;
- освіту;
- кіновиробництво;
- комп'ютерні ігри;
- Відеоконференції;
- соціальні мережі;
- системи дистанційного навчання.



На відміну від аналогових технологій, цифрове відео базується на представленні зображення у вигляді дискретних числових даних, що забезпечує високу якість.

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

Можливість багаторазового копіювання без втрати якості та ефективну обробку за допомогою програмних алгоритмів. Основою цифрового відео є послідовність кадрів, кожен із яких являє собою цифрове зображення.



Важливими характеристиками цифрового відео є:

- роздільна здатність;
- частота кадрів;
- глибина кольору;
- бітрейт.

Розглянемо докладніше кожен з цих характеристик.

1.Роздільна здатність (Resolution) — це кількість пікселів (елементів зображення), що формують один відеокадр.

Записується як:

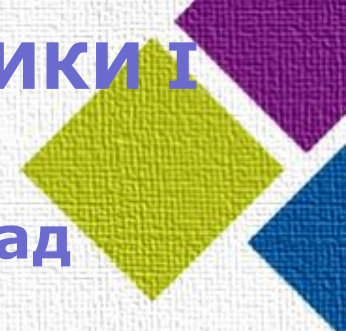
$$\textit{Resolution} = \textit{Width} \times \textit{Height}$$

де: **R** — роздільна здатність кадру;

W — ширина (**Width**);

H — висота (**Height**).

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ



Наприклад

1920×1080

W x **H**

означає

- **W** - **1920** пікселів у рядку;
- **H** - **1080** рядків.

Приклад_№1 - стандарт роздільної здатності відео - **Full HD**

$$1920 \times 1080 = 2073600 \approx 2.07 \text{MP}$$

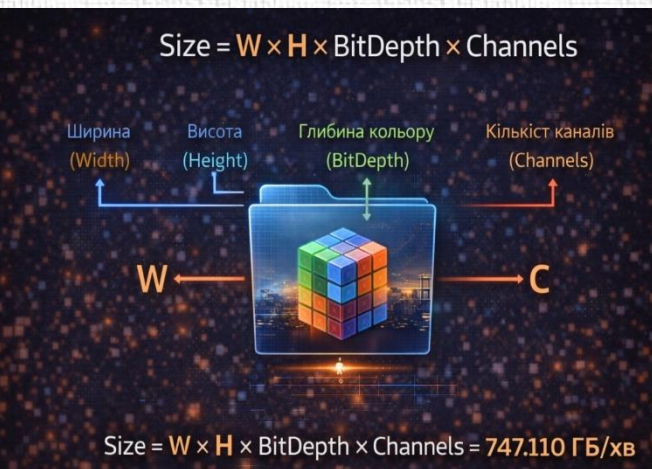
тобто: ≈ 2 мегапікселя.

Приклад_№2 - стандарт роздільної здатності відео - **4K UHD**

$$3840 \times 2160 = 8294400 \text{MP}$$

тобто: ≈ 8.3 - мегапікселя.

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ



Щільність інформації - ($\text{BitDepth} + \text{Pixels}$).
Обсяг одного **RAW** – кадру – (**Raw frame**) - це нестиснений або мінімально оброблений цифровий кадр, який містить первинні дані з сенсора камери або з буфера відеозахоплення, без стиснення кодеками типу **H.264/H.265** та без значної постобробки.

$$\text{Size} = W \times H \times \text{BitDepth} \times \text{Channels}$$

де:

Size - обсяг одного RAW-кадру;

W - ширина (Width);

H - висота (Height);

BitDepth - глибина кольору;

Channels - кількість каналів.

Розмір одного **RAW-кадру** або **цифрового зображення** без компресії, дорівнює ширині зображення, помноженій на висоту зображення, помноженій на глибину кольору та на кількість каналів.

Формула відповідає на питання, "Скільки інформації потрібно, щоб описати всі пікселі кадру?"

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

Чим RAW відрізняється від H.264

RAW	H.264
Без втрат	Сильна компресія
Максимальна деталізація	Втрата частини даних
Величезний розмір	Малий розмір
Гнучка кольорокорекція	Обмежена

Приклад_№1 - стандарт роздільної здатності відео - **Full HD**

W – 1920.

H – 1080.

RGB 8-bit.

1920 × 1080 × 8 × 3 = 49766400 bits : 8 bytes = 6.2MB один кадр

W H Bit C

W×H – це кількість пікселів у кадрі.

BitDepth - глибина кольору. Скільки біт інформації зберігається для одного каналу пікселя. **Наприклад:**

8 bit → 256 рівнів;

10 bit → 1024 рівні;

12 bit → 4096 рівнів.

Channels — канали кольору. Скільки компонент має піксель. **RGB** - **Red, Green, Blue.** Тобто – 3.

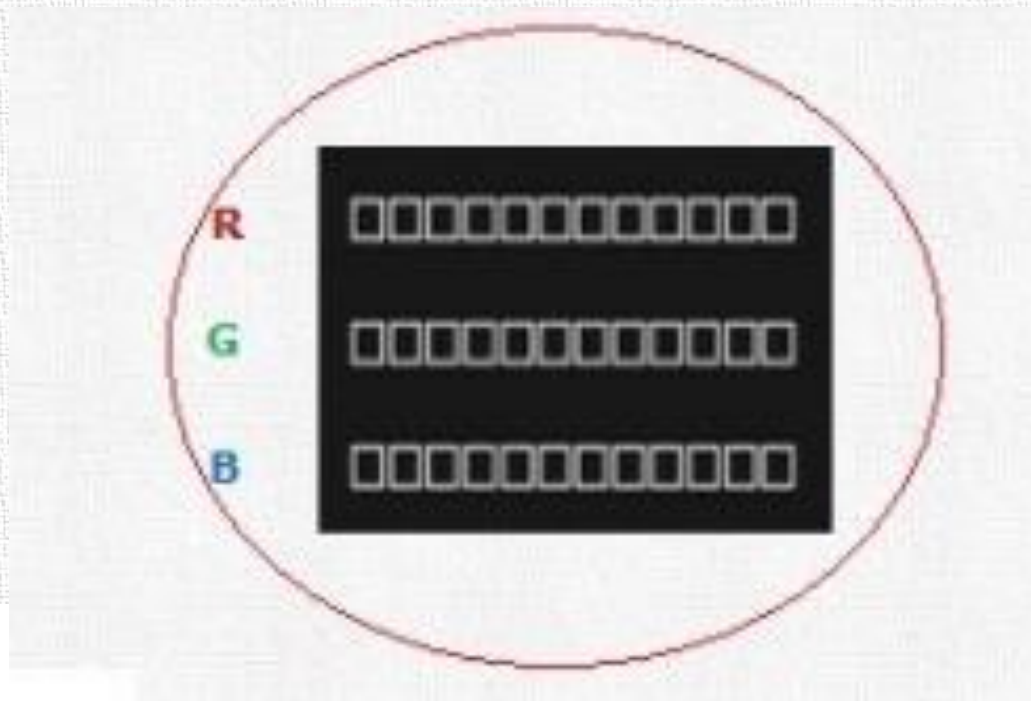
ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

Кожен піксель містить інформацію.
Наприклад:

RGB 8-bit: Red = 8 bit, Green = 8 bit, Blue = 8 bit.

Разом - 24bit = 3byte на один піксель.

один піксель



ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

Pixel Aspect Ratio (PAR) - співвідношення ширини пікселя до висоти пікселя:

$$DAR = \frac{W \times PAR}{H}$$

де: **DAR** - Display Aspect Ratio - співвідношення сторін зображення на екрані (що реально бачить користувач), наприклад **16:9** або **4:3**.

Не всі пікселі є квадратними, ця формула показує, як із цього впливає відображуване співвідношення сторін кадру.

1. Якщо піксель квадратний, то **PAR=1**.

2. Якщо піксель "витагнутий" по горизонталі або вертикалі, тоді **PAR ≠ 1**.

У деяких відеостандартах (особливо старих **SD**, **DVD**, телемовлення) пікселі могли бути **прямокутними**. Тому один і той же матричний розмір кадру (наприклад, **720x576**) може відображатися як **4:3** або **16:9** — залежно від **PAR**.

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ


Візьмемо PAL DVD: 720×576.



1. Якщо потрібно - **4:3**, то **PAR** приблизно - **1.067**.

2. Якщо потрібно - **16:9**, то **PAR** приблизно - **1.422**.

Один і той самий розмір кадру **720×576** може виглядати як **4:3** або **16:9** — через різний **PAR**.

 **PAR = 1**
(квадратний піксель)

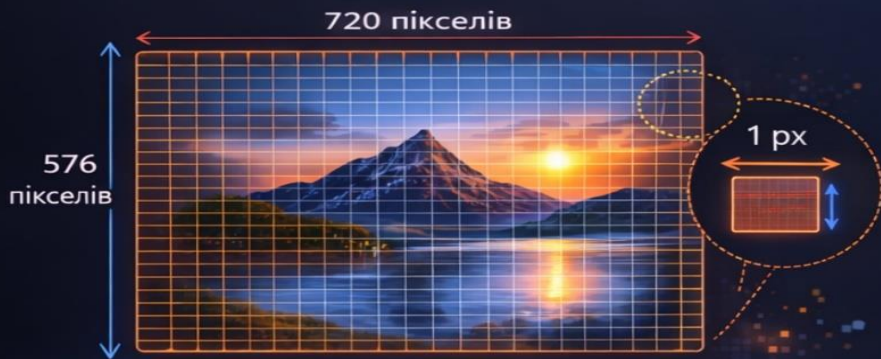
Один і той самий кадр: **720 × 576** пікселів
(W = 720, H = 576)

$$\text{PAR} = \frac{\text{Width (pixel)}}{\text{Height (pixel)}}$$

↓
Width

ВАРІАНТ 1 — **4:3** (стандарт DVD)

PAR = 1.067 (піксель ширший за висоту)



$$\text{DAR} = \frac{W \times \text{PAR}}{H}$$

$$= \frac{720 \times 1.067}{576} = 1.333 \approx 4:3$$

ВАРІАНТ 2 — **16:9** (широкоекранний)

PAR = 1.422 (піксель ще ширший)



$$\text{DAR} = \frac{720 \times 1.422}{576} = 1.777 \approx 16:9$$

W = 720
ширина (пікселів)



H = 576
висота (пікселів)



BitDepth + Channels
(не впливають на PAR!)

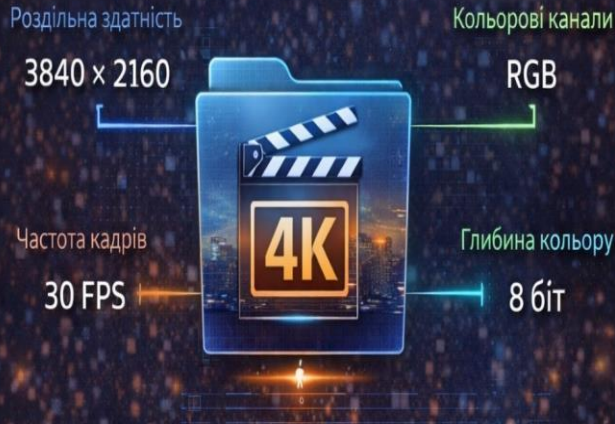


PAR ≠ 1 →
піксель прямокутний!



ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

Обсяг відео без компресії



Відео без компресії: $3840 \times 2160 \times 30 \times \text{RGB} \times 8 \text{ біт} = 747.110 \text{ ГБ/хв}$

$$\text{VideoSize} = \text{FrameSize} \times \text{FPS} \times \text{Time}$$

де: **VideoSize** - обсяг відео;
FrameSize - розмір кадру;
FPS - частота кадрів за секунду;
Time - тривалість.

- Приклад** - стандарт роздільної здатності відео - **Full HD**:
- FrameSize** - **6.2 MB**.
 - FPS** - **30 fps**.
 - Тривалість** - **1 хвилина**.

$$6.2 \times 30 \times 60 = 11160 \text{ MB} / 1024 \approx 10.89 \text{ GB}.$$

Саме тому потрібне стиснення.

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

На сьогоднішній день існує таке поняття, як **стандарт роздільної здатності**. **Стандарт роздільної здатності** — це узгоджений набір параметрів:

- кількість пікселів;
- співвідношення сторін;
- частота кадрів;
- колірний простір;
- розгортка.

Головна міжнародна організація

ITU — International Telecommunication Union - міжнародний союз електросвязи.

-**SD** - PAL:720×576. NTSC:720×480.

-**HD** - 1280×720 - став переходом від аналогового та **SD**-відео до **цифрового телебачення високої чіткості**. Його застосування було дуже широким і охопило практично всі сфери **мультимедійних технологій**.

-**Full HD** - 1920×1080 – YouTube, Онлайн-курси, Телебачення, Відеоконференції.

-**Ultra HD (4K)** - 3840×2160 - кіноіндустрія.

-**Cinema 4K** - 4096×2160 - кінотеатри.

-**8K** - 7680×4320 - великі екрани, VR.

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

ЯКІСТЬ ПЕРЕДАЧІ СВІТЛА ТА КОЛЬОРУ

-**SDR (Standard Dynamic Range)** — стандартний динамічний діапазон зображення. Це класичний формат відео і телебачення, який використовувався десятиліттями. Зображення має обмежений діапазон яскравості і кольорів. **8-bit** колір, вузький колірний простір. На практиці це означає світлі області - швидко "вигорають". Темні - втрачають деталі.

сонце → біла пляма, темна кімната → чорна маса.

-**HDR (High Dynamic Range)** — розширений динамічний діапазон. **HDR** дозволяє показувати дуже темні деталі, дуже яскраві деталі, одночасно. **10-bit** колір, широкий колірний простір. На практиці це означає видно фактуру хмар біля сонця, текстуру тіней, реалістичні відблиски.

Що означає - HDR 1080p

HDR - якість передачі світла та кольору;

1080p - кількість пікселів;

- прогресивна розгортка формує повний кадр одразу.

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

Progressive vs Interlaced - “Прогресивна розгортка vs Черезрядкова розгортка” - два способу формування кадру у відео.

Progressive: 1080p,
весь кадр.



Цілий кадр

Interlaced: 1080i,
поля.



t_0
 t_1 Поле 1 (непарні рядки)



t_1 Поле 2 (парні рядки)

1. **Progressive** - 1080**p**, весь кадр.
2. **Interlaced** - 1080**i**, поля.

1. Progressive (1080p) — “цілий кадр”. Прогресивна розгортка формує повний кадр одразу. Кадр має 1920×1080 пікселів. У момент часу t_0 знімається/відтворюється весь кадр.

У наступний момент t_1 — наступний повний кадр, і т. інш. Як це виглядає “в часі”:

-1080p30=30 повних кадрів/сек.

-1080p60=60 повних кадрів/сек.

Кожен кадр — “цілісний знімок” сцени в один момент.

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

Progressive vs Interlaced - “Прогресивна розгортка vs Черезрядкова розгортка” - два способу формування кадру у відео.

Progressive: 1080p,
весь кадр.



Цілий кадр

Interlaced: 1080i,
поля.



t_1 Поле 1 (непарні рядки)



t_1 Поле 2 (парні рядки)

1. **Progressive** - 1080**p**, весь кадр.
2. **Interlaced** - 1080**i**, поля.

2. Interlaced (1080i) — “поля”.
Черезрядкова розгортка ділить кадр на два поля (fields):

Field 1: непарні рядки (1,3,5,...).

Field 2: парні рядки (2,4,6,...).

1080i50 = 50 полів/сек. (25 “кадрів”, але з оновленням руху 50 раз/сек).

1080i60 = 60 полів/сек. (30 “кадрів”, але з оновленням руху 60 раз/сек.).

Один “повний кадр” складається з двох половин, знятих у різний час.



Association for
Computing Machinery

Advancing Computing as a Science & Profession

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

Progressive vs Interlaced - “Прогресивна розгортка vs Черезрядкова розгортка” - два способу формування кадру у відео.

Progressive: 1080p,
весь кадр.



Цілий кадр

Interlaced: 1080i,
поля.



t_1 Поле 1 (непарні рядки)



t_1 Поле 2 (парні рядки)

Combing

Чому на картинці показаний “combing” (гребінка).

Коли об’єкт швидко рухається (машина на рисунку), поля в **1080i** “бачать” його у двох різних положеннях. При спробі скласти їх в один кадр, контури починають “роздвоюватися” рядками → ефект гребінки. Це не дефект камери як такої — це наслідок **часового рознесення полів**.

Progressive vs Interlaced - “Прогресивна розгортка vs Черезрядкова розгортка”

Interlaced (i) — де було класично і де ще трапляється:

- телемовлення старих стандартів (особливо в країнах з PAL/NTSC спадщиною);
- архіви телевізійного контенту (новини, спорт, репортажі);
- частина broadcast-інфраструктури, яка історично будувалася під 1080i.

Причина. interlaced був “хитрим компромісом” для аналогового/раннього цифрового ТВ, зменшити смугу пропускання, але зберегти плавність руху.

Progressive (p) — це основний стандарт сьогодні:

- YouTube, стримінг, соцмережі;
- смартфони, ПК, монітори;
- сучасні камери, більшість “цифрових” пайплайнів;
- відеоігри та screen-recording;
- VR/AR контент (де затримка і чіткість критичні).

Причина. сучасні дисплеї (LCD/LED/OLED) — прогресивні, вони відображають повні кадри.

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

PPI - (**P**ixels **P**er **I**nch) — пікселів на дюйм, тобто **щільність пікселів** на екрані. Для дисплеїв:

$$PPI = \frac{\sqrt{W^2 + H^2}}{Diagonal}$$

PPI дорівнює кореню квадратному з суми квадратів ширини й висоти (в пікселях), поділеному на діагональ екрана (в дюймах). Отже, **PPI** = (скільки пікселів по діагоналі) / (скільки дюймів по діагоналі). Це і дає **пікселів на 1 дюйм**. Приклад розрахунку: монітор 24" Full HD (1920×1080). Пікселі по діагоналі:

1. Пікселі по діагоналі:

$$\sqrt{1920^2 + 1080^2} \approx 2202.9$$

2. PPI:

$$PPI = \frac{2202.9}{24} \approx 91.8$$

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

DPI - (Dots Per Inch) — точок на дюйм. Це щільність друкарських точок, які принтер (або поліграфічна система) здатен нанести на папір на відрізок 1 дюйм. **DPI** описує фізичну “дрібність” друку. 300 **DPI** означає - уздовж 1 дюйма принтер може розмістити до 300 “точок” (dot). **600 DPI** — ще дрібніше.

Сканери (часто вказують як “dpi сканування”). У сканерах **DPI** означає **дискретизацію оцифрування** - 300 DPI, 600 DPI, 1200 **DPI** — скільки точок сканер знімає з паперу на дюйм. Тут це ближче до **PPI**, але термін “dpi” історично закріпився. Якщо ви маєте зображення шириною **W_{px}** пікселів і хочете надрукувати його шириною **W_{in}** дюймів, тоді потрібна щільність:

$$DPI = \frac{W_{px}}{W_{in}}$$

Аналогічно по висоті:

$$DPI = \frac{H_{px}}{H_{in}}$$

DPI - дорівнює кількості пікселів по ширині (або висоті), поділеній на фізичну ширину (або висоту) в дюймах.



ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

2.Частота кадрів — **Frame Rate, FPS** — це кількість кадрів, які відеосистема знімає, кодує або відтворює за **1 секунду**. **1 FPS = 1Гц** у сенсі "подій за секунду". 1 FPS як частота:

$$FPS = \frac{N}{T}$$

де:

N — кількість кадрів;

T — час у секундах.

FPS - дорівнює кількості кадрів, поділений на час.

Період одного кадру (час між кадрами):

$$\Delta t = \frac{1}{FPS}$$

Тривалість одного кадру дорівнює одиниці, поділений на **FPS**.
Приклади.

30fps → $\Delta t = 1/30 \approx 0.0333$ c = **33.3** мс.

60fps → $\Delta t = 1/60 \approx 0.0167$ c = **16.7** мс.

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ



Кількість кадрів у відео заданої тривалості:

$$N = FPS \times T$$

де: T — час у секундах.

Bitrate — це швидкість передачі або запису відеоданих, тобто скільки біт інформації передається або зберігається за одну секунду відео. **Вплив на потік даних.** Для некодованого відео (або для приблизної оцінки):

$$Bitrate \propto (W \times H) \times FPS$$

Бітрейт прямо пропорційний добутку ширини кадру на висоту кадру та на частоту кадрів. Тобто при однаковій роздільній здатності збільшення **FPS** збільшує потребу в даних майже лінійно.

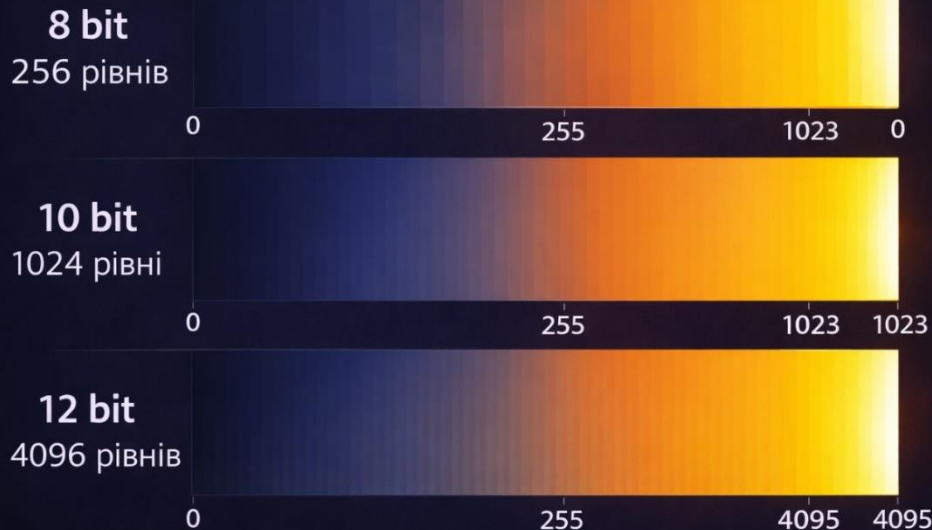
Короткі рекомендації:

- лекції, навчальні відео: **25/30 fps.**
- спорт, швидкі демонстрації: **50/60 fps.**
- кіно/художній стиль: **24 fps.**
- VR / ігри: **від 60 до 90 fps.**

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

3.Глибина кольору - **Bit Depth, Color Depth** - це ключова характеристика цифрового зображення/відео, яка визначає скільки дискретних рівнів може приймати яскравість або колірний канал. Від неї залежить плавність градієнтів, наявність banding, можливості **HDR** та якість кольорокорекції.

Глибина кольору (Bit Depth, Color Depth)



Глибина кольору показує, скільки біт виділено на кодування:

- яскравості;

- кольорових компонент наприклад R, G, B.

Чим більше бітів → тим більше можливих значень → тим точніше передаються відтінки.

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

Кількість рівнів для одного каналу. Якщо на канал виділено **b** біт, то кількість можливих рівнів:

$$L = 2^b$$

Приклади:

- b=8-біт, $L=2^8 = 256$ рівнів;
- b=10-біт, $L=2^{10} = 1024$ рівні;
- b=12-біт, $L=2^{12} = 4096$ рівнів.

Кількість можливих кольорів (для RGB).

Для **RGB**, якщо кожен з трьох каналів має **b** біт:

$$Colors = (2^b)^3 = 2^{3b}$$

Кількість кольорів дорівнює двійці в степені три помножити на b.

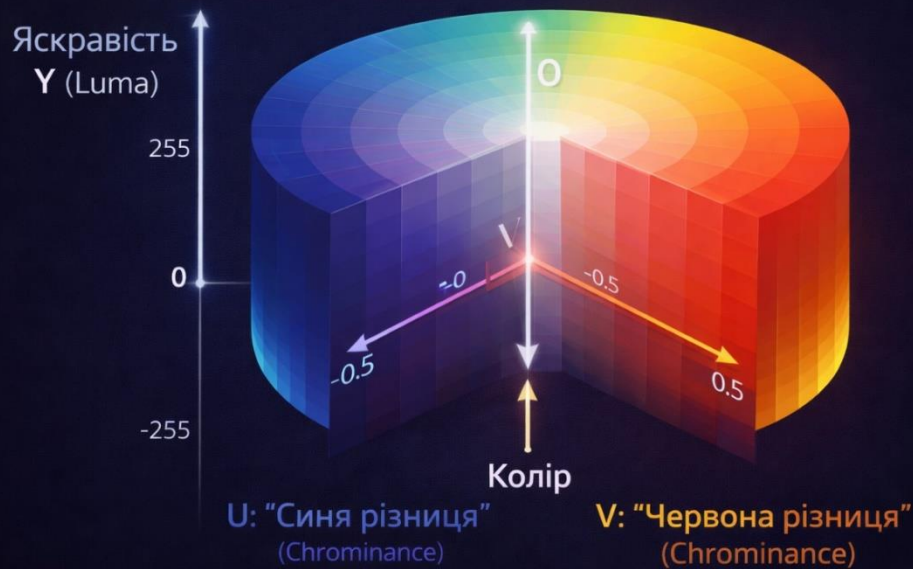
Приклади.

- 8-bit RGB**: $2^{24} \approx 16.7$ млн. кольорів;
- 10-bit RGB**: $2^{30} \approx 1.07$ млрд. кольорів;
- 12-bit RGB**: $2^{36} \approx 68.7$ млрд. кольорів.

Важливо - у відео часто використовується **YUV** - **вай-ю-ві** із компресією та квантизацією (**4:2:0**), там "кількість кольорів" оцінюється інакше, але правило 2^b для рівнів каналу зберігається.

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

Колірна модель YUV



YUV - вай-ю-ві — це сімейство колірних моделей (і практичних форматів представлення відео), де зображення розкладається на:

- Y** — яскравість;
- U** і **V** — кольорові різниці (**chrominance**, компоненти кольору). Людське око значно чутливіше до деталей яскравості, ніж до деталей кольору.

Тому відео кодують так, щоб:

- Y (яскравість) зберігати з високою роздільною здатністю, колір (U/V або Cb/Cr) можна зберігати грубіше (з меншим обсягом даних) — майже без помітної втрати якості. Це дає сильну економію бітрейту і розміру файлу.

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

Що означають компоненти YUV - вай-ю-ві

Y - приблизно відповідає "чорно-білому" зображенню - інтенсивність/яскравість.

U і **V** - це компоненти кольору, які кодують відхилення від яскравості, один канал описує "синю різницю", інший — "червону різницю".

У цифровому відео замість U/V зазвичай використовують Cb і Cr.

$$\mathbf{Cb} \approx (\mathbf{Blue} - \mathbf{Y}), \mathbf{Cr} \approx (\mathbf{Red} - \mathbf{Y}).$$

Як отримують **Y**, **Cb**, **Cr** з **RGB**. Точні коефіцієнти залежать від стандарту.

$$\mathbf{Яскравість} - \mathbf{Y} = \mathbf{0.2126R} + \mathbf{0.7152G} + \mathbf{0.0722B}$$

Y дорівнює сумі **R**, **G**, **B** із вагами; найбільша вага у зеленого, бо око найбільш чутливе до нього.

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

Позначення **4:a:b** — це стандартний спосіб описати хрома-субдискретизацію у форматах **YCbCr** (часто кажуть **YUV**). Воно показує, з якою частотою (щільністю) зберігаються кольорові компоненти **Cb/Cr** відносно яскравості **Y**.

Y - яскравість зберігають детально, а **Cb/Cr** - колір можна зберігати рідше, бо око менш чутливе до кольорових деталей.

Y — канал яскравості - містить основні деталі, контури, текстуру.

Cb і **Cr** — два кольорові канали ("синя різниця" і "червона різниця").

Як читати "4:a:b"

Формат **4:a:b** описує відносну кількість зразків на певній "опорній" ділянці зображення. У телевізійній традиції беруть рядок із 4 умовних "люткових" позицій (4 лума-зразки **Y** по горизонталі) і дивляться, скільки там:

-**Cb**-зразків;

-**Cr**-зразків.

Спрощено. **4** — це "база" для **Y** по горизонталі (**Y** завжди найдетальніший).

a — скільки зразків **Cb/Cr** беруть у першому рядку на ті ж 4 позиції.

b — скільки зразків **Cb/Cr** беруть у другому рядку (показує вертикальне зменшення).

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

У популярних форматах (**4:4:4**, **4:2:2**, **4:2:0**) числа однакові для **Cb** і **Cr**, тому пишуть одним "a" і "b", маючи на увазі обидва хрома-канали. Стиснення відео є необхідною умовою ефективного зберігання та передачі даних. Без компресії відеофайли займали б надзвичайно великий обсяг пам'яті. Для цієї операції застосовують кодеки. Кодеки використовують міжкадрове та внутрішньокадрове прогнозування, компенсацію руху та перетворення **DCT**.

DCT — Discrete Cosine Transform — дискретне косинусне перетворення. Це математичний метод перетворення сигналу (зображення, аудіо, відео) з просторової області у частотну область.

Простими словами - **DCT** дозволяє представити зображення або сигнал як суму косинусних хвиль різної частоти. У відео та зображеннях сусідні пікселі зазвичай схожі. **DCT** - перетворює блок пікселів у набір коефіцієнтів, концентрує більшість енергії в низьких частотах, дозволяє відкинути маловажливі високочастотні компоненти. Саме тому **DCT** лежить в основі: JPEG, MPEG, H.264/H.265, AV1 (там модифікації).

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

Бітрейт (Bitrate)



4. Бітрейт - це швидкість потоку даних, тобто скільки біт інформації передається або зчитується за 1 секунду під час відтворення/передачі/запису відео. Його можна уявити як "пропускну здатність" для відео: чим вона більша, тим більше інформації можна передати про зображення.

Високий бітрейт забезпечує кращу якість, але збільшує розмір файлу.

$$B = \frac{\text{bits}}{\text{second}}$$

Наприклад:

$$8 \text{ Mbps} = 8\,000\,000 \text{ біт/с.}$$

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

Якщо **бітрейт** приблизно сталий, то розмір файлу у **битах**:

$$Size_{bits} = B \times T$$

де:

B — бітрейт (bit/s);

T — тривалість (s).

Розмір файлу у байтах:

$$Size_{Bytes} = \frac{B \times T}{8}$$

Приклад:

Відео 10 хв. = 600 с, бітрейт 8 Mbps:

$$Size = \frac{8 \times 600}{8} = 600 \times 10^6 \approx 600MB$$

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

Відео зазвичай стискається кодеком (H.264/H.265/AV1) у послідовність кадрів/пакетів. Потік складається з:

- відеопотоку (video elementary stream);
- аудіопотоку (audio stream);
- інколи субтитрів/метаданих.

Загальний бітрейт:

$$B_{total} = B_{video} + B_{audio} + B_{overhead}$$

де **overhead** — службові дані контейнера/пакетизації.

Бітрейт відео прямо пропорційний добутку ширини кадру на висоту кадру, помноженому на частоту кадрів і на складність сцени. Тобто, чим більше пікселів у кадрі, більше кадрів за секунду і складніша сцена — тим більший потрібен **бітрейт**.

$$B_{video} \propto (W \times H) \times FPS \times Complexity$$

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

Де:

W×H — кількість пікселів у кадрі;

FPS — кадрів/сек;

Complexity — складність сцени (рух, шум, деталі, текстури).

Це пояснює, чому **4K** потребує приблизно в 4 рази більше даних ніж **1080p** (за інших рівних), **60fps** \approx у двічі важче ніж **30fps**.

Типи бітрейту

1.CBR (Constant Bitrate) — сталий бітрейт:

$$B(t) \approx \text{const}$$

Плюси:

-передбачуваний канал (зручно для **broadcast**, де смуга фіксована).

Мінуси:

-неефективно, прості сцени отримують "зайві біти", складні — можуть страждати.

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

2.VBR (Variable Bitrate) — змінний бітрейт **$B(t)$** змінюється залежно від сцени:

$B(t)$ змінюється залежно від сцени

Плюси:

-якість зазвичай вища при такому ж середньому розмірі.

Мінуси:

-бітрейт "стрибає", потрібен буфер.

У реальних кодерах часто використовують: Average bitrate (середній), Max bitrate (піковий).

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

Бітрейт і якість

Що реально змінюється. **Занадто малий бітрейт дає:**

- артефакти - "квадрати"/macroblocking;
- "мило" (втрата деталей);
- banding у градієнтах;
- "комариків" навколо контурів (ringing);
- змащення в русі.

Вищий бітрейт:

- більше деталей;
- менше артефактів;
- але після певного рівня зростання майже не помітне (diminishing returns).

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

Приклади як користуватися бітрейтом (практично)

Розрахунок потрібного бітрейту під обмежений розмір файлу:

Припустимо, треба 1 година відео+аудіо в 2 GB. 1 година = 3600с. Переведемо два гігабайти в мегабіти:

$$2 \text{ GB} \approx 2 \times 1024 \text{ MB} = 2048 \text{ MB} \approx 2048 \times 8 = \mathbf{16384 \text{ Mbit}}$$

$$B_{\text{video+audio}} = \frac{16384}{3600} \approx \mathbf{4.55 \text{ Mbps}}$$

Тобто середній бітрейт $\approx \mathbf{4.55 \text{ Mbps}}$ (включно з аудіо).

Якщо аудіо **AAC = 192 kbps**:

$$B_{\text{video}} \approx 4.55 - 0.192 = \mathbf{4.36 \text{ Mbps}}$$

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

Бітрейт для різних задач -орієнтири

Це практичні діапазони для **H.264/H.265**:

- ❖ **720p30**: ~2–5 Mbps
- ❖ **1080p30**: ~5–10 Mbps
- ❖ **1080p60**: ~8–15 Mbps
- ❖ **4K30**: ~15–35 Mbps
- ❖ **4K60**: ~25–60 Mbps

Бітрейт у стримінгу це адаптивний бітрейт. Відео готують у кількох профілях:

- 360p** - (низький);
- 720p**;
- 1080p**;
- 4K**.

Плеєр під час перегляду перемикає між ними залежно від швидкості інтернету, щоб уникати буферизації.

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

Як “правильно користуватися” бітрейтом при експорті

Коли задавати bitrate.

1. Якщо потрібен контроль розміру файлу або обмеження каналу.
2. Якщо робите стримінг або **broadcast**.

Коли краще не задавати bitrate напямую.

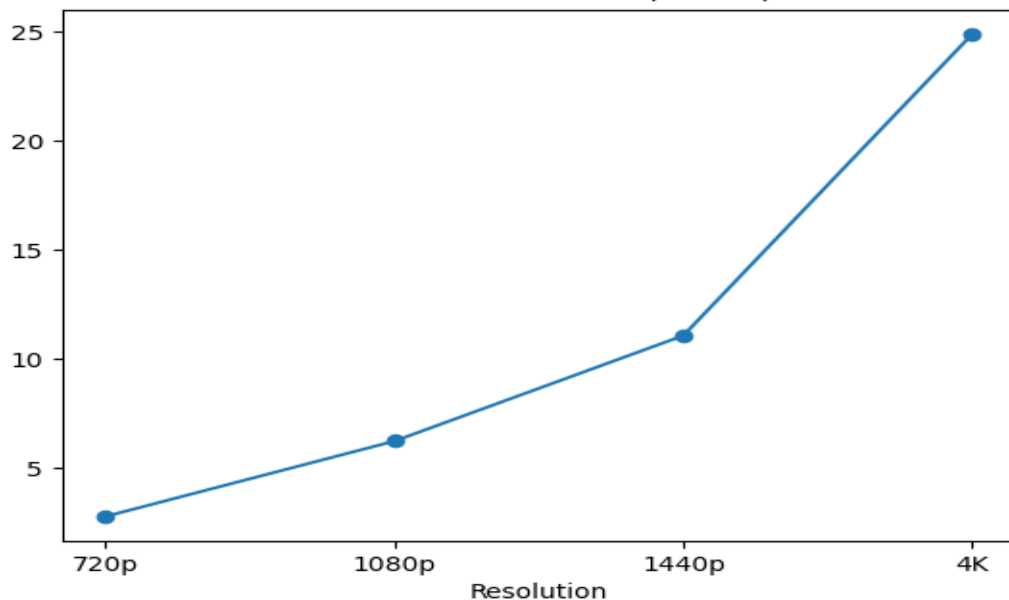
1. Коли важливіша візуальна якість, ніж точний розмір.

Висновок. Бітрейт визначає компроміс:

якість ↔ розмір файлу ↔ вимоги до мережі/диска/процесора

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

Bitrate vs Resolution (W × H)



1. Залежність бітрейту від роздільної здатності.

Графік Bitrate vs Resolution (W×H) показує по осі:

X — роздільна здатність (720p, 1080p, 1440p, 4K).

по осі:

Y — відносний бітрейт.

X — роздільна здатність - це стандартні роздільні здатності відео.

Формат	Роздільна здатність
720p	1280 × 720
1080p	1920 × 1080
1440p	2560 × 1440
4K	3840 × 2160

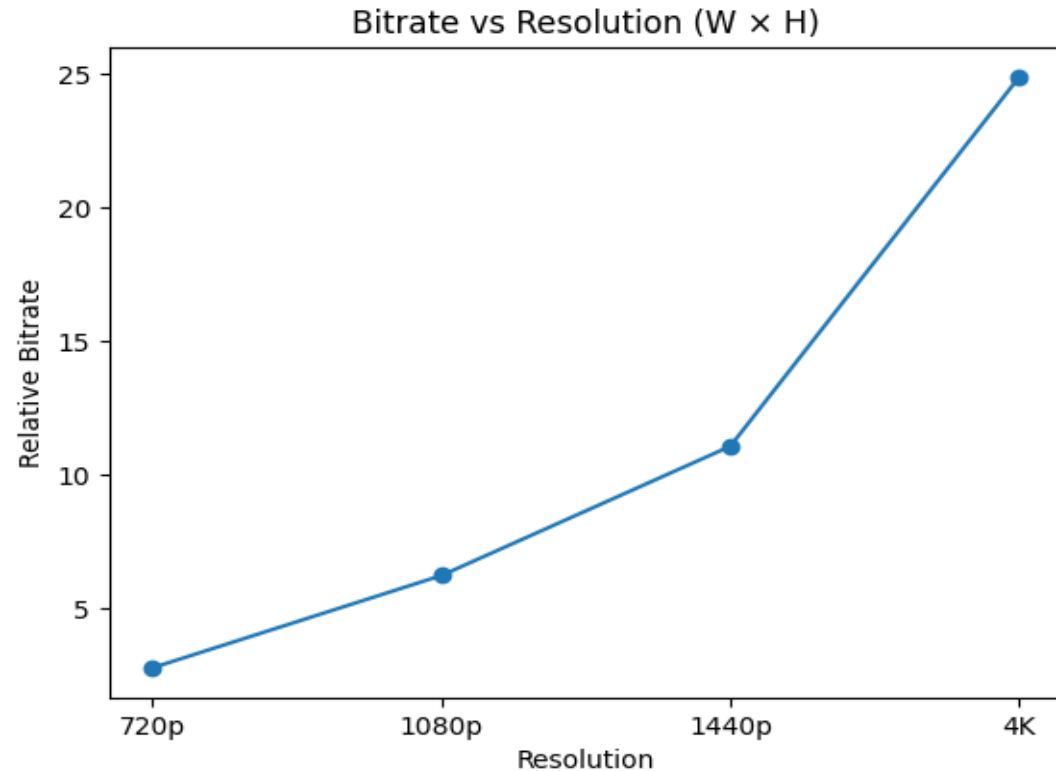


ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

Фізичний сенс:

$$R \propto W \times H$$

Кількість пікселів:



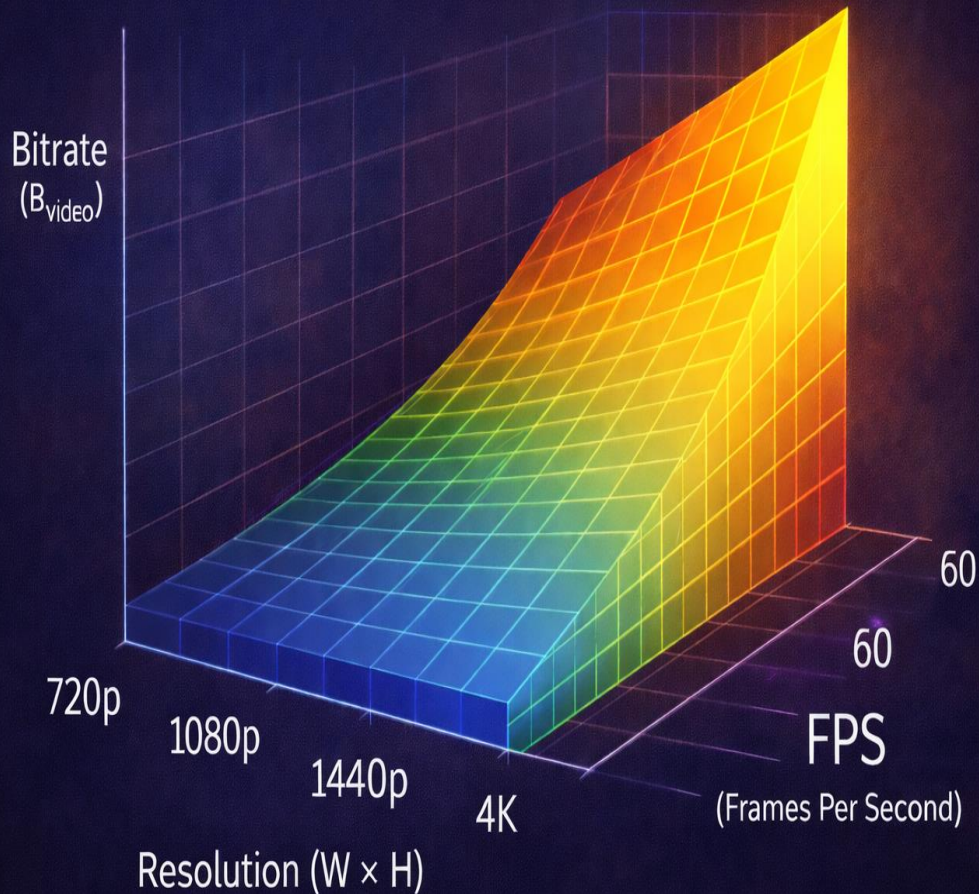
Формат	Пікселів
720p	0.92 MP
1080p	2.07 MP
1440p	3.69 MP
4K	8.29 MP

Графік показує - збільшення роздільної здатності відео призводить до різкого зростання необхідного бітрейту. Особливо це видно при переході до 4K. Отже 4K більше даних ніж 1080p тому потрібен більший бітрейт.

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

Графік демонструє основну залежність:

$$B_{video} \propto (W \times H) \times FPS$$



$$B_{video} \approx (W \times H) \times FPS$$

де:

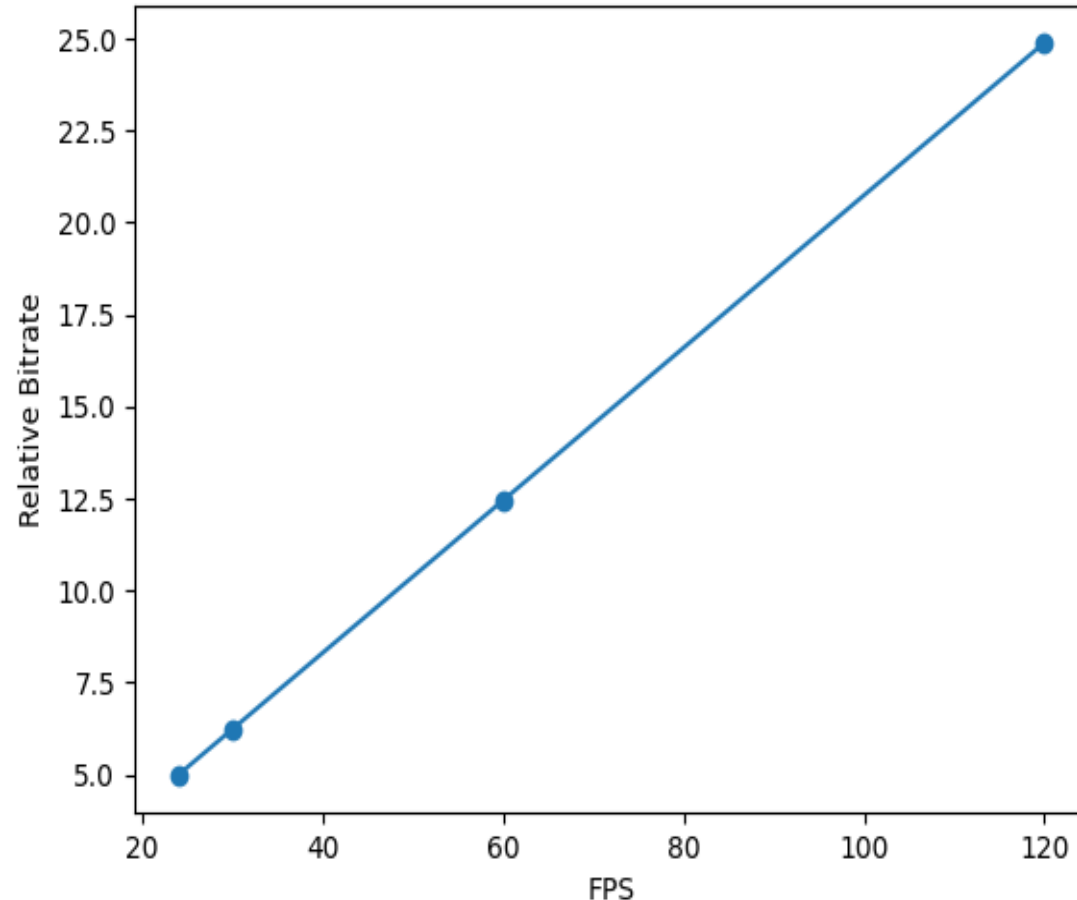
$W \times H$ — кількість пікселів;

FPS — кадрів за секунду.

На цьому графіку **FPS** вважається однаковим – **30fps**, тому показано лише вплив роздільної здатності.

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

Bitrate vs Frames Per Second (FPS)



2. Залежність бітрейту від частоти кадрів.

Графік показує:

- 24 fps
- 30 fps
- 60 fps
- 120 fps

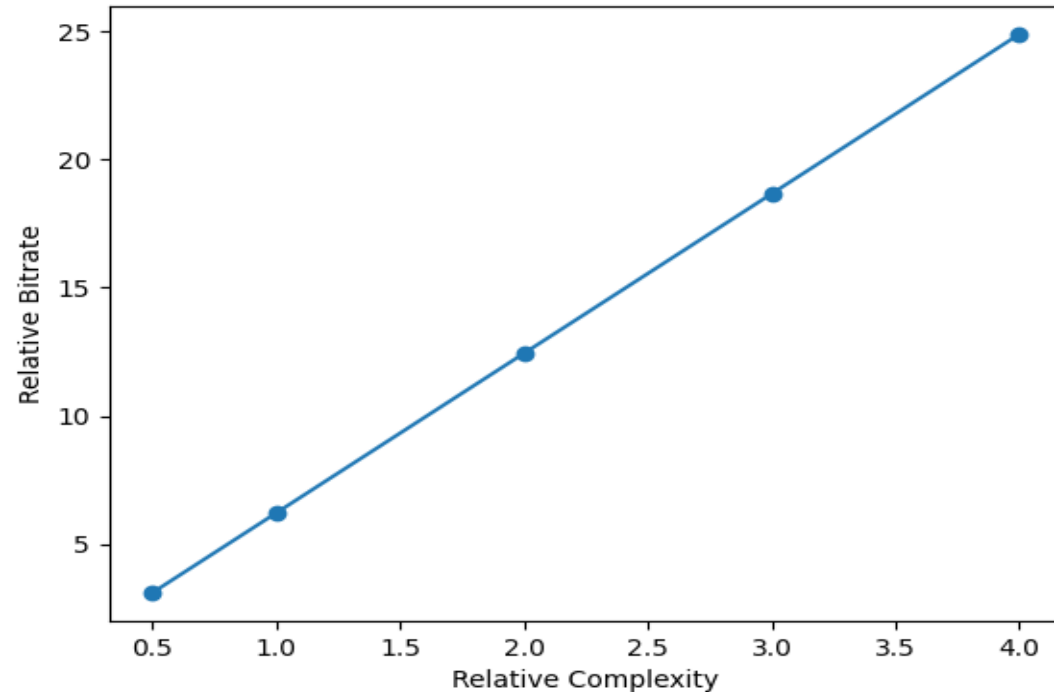
Залежність:

$$B \propto \text{FPS}$$

Якщо -30fps → 6 Mbps, то приблизно - 60fps → 12 Mbps. Тобто подвоєння FPS ≈ подвоєння бітрейту.

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

Bitrate vs Scene Complexity



3.Залежність бітрейту від складності сцени

На осі X - показані значення:

-0.5;

-1;

-2;

-3;

-4.

Це умовний індекс складності відео.

Він характеризує:

- кількість руху;
- кількість деталей;
- шум;
- текстури;
- зміни між кадрами.

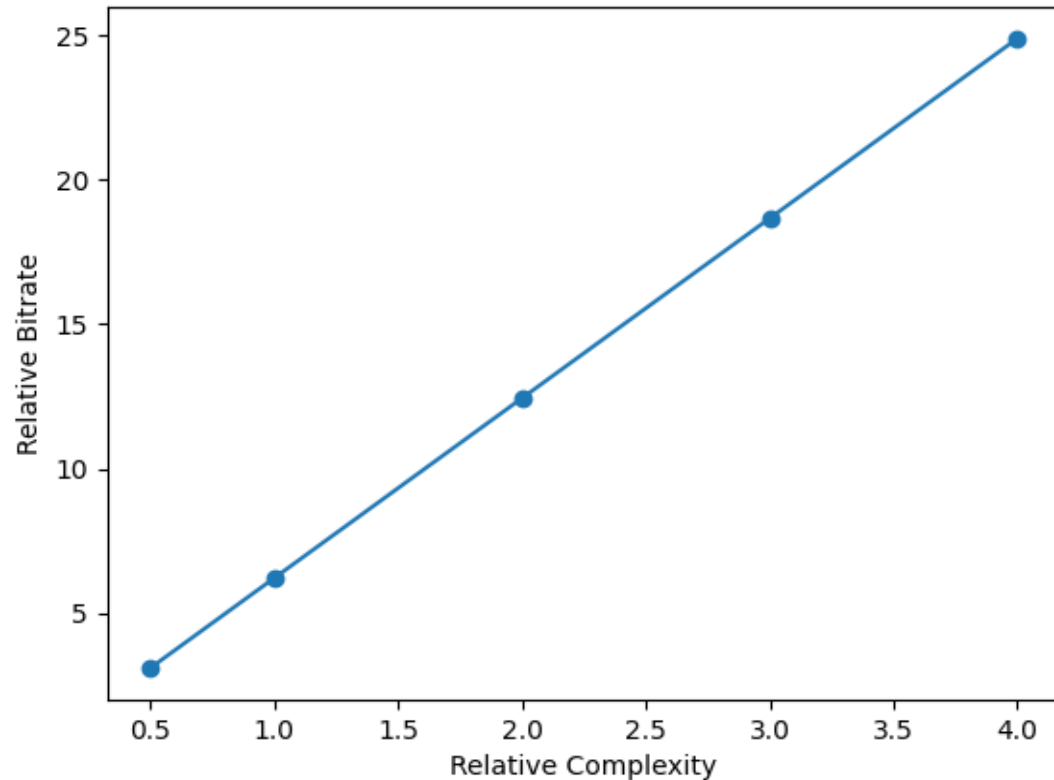


Association for
Computing Machinery

Advancing Computing as a Science & Profession

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

Bitrate vs Scene Complexity



Ось Y - показує наскільки збільшується обсяг даних, який потрібно передати або закодувати. Це відносний бітрейт відео. Це не абсолютні **Mbps**, а пропорційна шкала.

$$B_{video} = \propto Complexity$$

де:

B_{video} — бітрейт відео.

Чим складніша сцена, тим більше інформації потрібно передати.

Де використовується така модель. Подібні оцінки використовують у:

-YouTube;

-Netflix;

відеострімінгу;

відеоконференція і т.і.



Association for
Computing Machinery

Advancing Computing as a Science & Profession

ЦИФРОВЕ ВІДЕО, ЙОГО ХАРАКТЕРИСТИКИ І ФОРМАТИ ВІДЕОФАЙЛІВ

Приклади - залежність бітрейту від складності сцени

Тип сцени	Complexity
лекція	низька
новини	середня
футбол	висока
дощ/вода/листя	дуже висока

На бітрейт відео найбільше впливають:

- роздільна здатність;
- частота кадрів;
- складність сцени.