

УДК 621.397:004.032.6

**SPAİK: НЕЙРОМЕРЕЖЕВА ОПТИМІЗАЦІЯ БІТРЕЙТУ
ВІДЕО ДЛЯ МОБІЛЬНИХ ПРИСТРОЇВ ЧЕРЕЗ
SUPER-RESOLUTION ТА ПЕРЦЕПТУАЛЬНІ ВТРАТИ**

Гашимов І.А.

e-mail: ilkin.hashymov@nure.ua

Науковий керівник – к.т.н, проф. каф. МІРЕС Зубков О.В.

Харківський національний університет радіоелектроніки, каф. МІРЕС
м. Харків, Україна

This thesis focuses on neural-assisted bitrate reduction for mobile viewers in SPAİK (Streaming Platform for Adaptive Interactive Communications). The main problem is to decrease video bitrate without perceptible quality degradation on small screens, where classical encoder optimizations often trade sharpness for compression. We propose a practical pipeline that combines modern codec efficiency (AV1) with learning-based super-resolution (SR) on the client side. The approach transmits a lower-resolution stream and reconstructs perceived detail using SR models trained with perceptual losses, which are known to correlate better with human perception than pixel-wise losses. The work outlines microservice integration for encoding and optional neural enhancement, targeted at a TypeScript/Golang backend and multi-platform clients. Evaluation is planned using a perceptual video quality metric (VMAF) to quantify the bitrate–quality trade-off and to guide tuning decisions.

Для стримінгових платформ критичною метою є не лише «доставити відео», а й зробити це економно щодо стосовно мережі користувача, особливо на мобільних пристроях. У дипломній еволюції SPAİK («Спайк») планується розширення клієнтів (веб, мобільний, настільний) і перехід до мікросервісів у бекенді. Вихідною точкою виступає курсова реалізація платформи стримінгу як вебзастосунку з можливістю подальшої декомпозиції.

Проблема формулюється так: зменшити середній бітрейт відео без помітної втрати якості на малих екранах. Практично це означає: при однакових або близьких умовах перегляду на смартфоні користувач не повинен помічати «замилення», блокінг чи втрату деталей, але мережеве споживання має зменшитися.

Ключова ідея – використати те, що для мобільного перегляду часто прийнятні нижчі просторові роздільності, а втрати деталей можливо частково компенсувати learning-based super-resolution на клієнті.

В дослідженні була виконана розрахункова оцінка потенціалу економії бітрейту. Наведемо оцінку вимог до бітрейту через кількість пікселів (це не гарантія бітрейту, а опорна інженерна інтуїція для планування експериментів).

Крок 1. Кількість пікселів для 1080p: $1920 \times 1080 = 2\,073\,600$ пікселів.

Крок 2. Кількість пікселів для 540p (половина по ширині й висоті):
 $960 \times 540 = 518\,400$ пікселів.

Крок 3. Відношення: $2\,073\,600 / 518\,400 = 4,0$.

Отже, при зменшенні лінійної роздільності вдвічі кількість пікселів зменшується у 4 рази. Це створює простір для значного зниження бітрейту (ступінь залежить від кодека, контенту та цільової якості), після чого SR-модель може відновлювати сприйману різкість.

У якості «компенсатора» низької роздільності використовувався нейромережовий компонент perceptual loss і SR. Класичне навчання SR за MSE (піксельна помилка) часто приводить до «пластикових» або розмитих деталей. У роботі [2] запропоновано перцептуальні втрати (perceptual losses), які вимірюють схожість у просторі ознак попередньо натренованої мережі та дають візуально більш узгоджені результати для задач перетворення зображень, зокрема super-resolution.

Подальший розвиток SR у напрямі «реалістичних текстур» демонструє ESRGAN, де введено архітектурні й функціональні модифікації (зокрема зміни у perceptual loss), що спрямовані на відновлення природніших текстур у порівнянні з попередніми GAN-підходами.

Для SPAIK це означає практичний шлях: передавати на мобільний клієнт нижчу роздільність, зменшуючи мережеве навантаження, і відновлювати сприймані деталі SR-моделлю на пристрої (або опційно на edge-сервісі).

Для практичної реалізації кодування і вимірювання якості використовувався кодек AV1 і VMAF як інженерний критерій.

Щоб підсилити ефект від зниження роздільності, доцільно спиратися на сучасні кодеки. AV1 позиціонується як відкритий відеокодек, розроблений для високоефективного стиснення. Офіційна специфікація визначає формат бітстриму та процес декодування, що важливо для відтворюваності експериментів і сумісності клієнтів у різних реалізаціях.

Для оцінювання «помітності» втрат потрібен перцептуально орієнтований показник. VMAF (Video Multi-Method Assessment Fusion) описаний як практичний метрик, що корелює із суб'єктивними оцінками й застосовується для вимірювання якості та оптимізації кодування. Це дає формалізовану основу для вибору компромісу між бітрейтом і якістю у SPAIK.

У термінах експериментальної методики жослідження пропонується дві контрольні гілки:

- **Baseline:** AV1 (або інший обраний кодек) у нативній для мобільного роздільності/бітрейті.

- **Neural-assisted:** нижча роздільність + SR (perceptual loss) → порівняння VMAF та бітрейту.

Далі була запропонована інтеграція в мікросервісну архітектуру SPAIK. З огляду на план відмови від сторонніх медіасервісів і перехід до

мікросервісів, нейрокомпонент доцільно винести в ізольовану підсистему, щоб не ускладнювати контрольну площину платформи. У рамках стеку JavaScript/TypeScript/Golang практичним виглядає такий поділ (це саме проектна пропозиція): сервіс кодування/пакування (AV1), сервіс профілів якості (лідери/лестниці бітрейтів), сервіс (або SDK на клієнті) нейропокращення для мобільного, а також сервіс метрик (облік VMAF/часу декодування/енергоспоживання). Специфікації AV1 та метрика VMAF задають перевірювану основу для відтворюваних експериментів.

У тезі запропоновано реалістичний для дипломного впровадження шлях зменшення бітрейту на мобільних клієнтах SPAIK: поєднання ефективного кодека (AV1) із client-side super-resolution, натренованим на perceptual loss (Johnson et al.) та/або ESRGAN-підходах, із валідацією якості через VMAF. Така постановка дає вимірюваний критерій «економія бітрейту $\leftrightarrow$ перцептуальна якість» і добре узгоджується з мікросервісною еволюцією платформи.

Список використаних джерел:

1. Alliance for Open Media. AV1 Bitstream & Decoding Process Specification : вебресурс. URL: <https://aomediacodec.github.io/av1-spec/> (дата звернення: 12.03.2026).
2. Johnson J., Alahi A., Fei-Fei L. Perceptual Losses for Real-Time Style Transfer and Super-Resolution. ECCV 2016. URL: <https://svl.stanford.edu/assets/publications/pdfs/JohnsonECCV16.pdf> (дата звернення: 12.03.2026).
3. Wang X., Yu K., Wu S. та ін. ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks. ECCVW 2018. URL: https://openaccess.thecvf.com/content_ECCVW_2018/papers/11133/Wang_ESRGAN_Enhanced_Super-Resolution_Generative_Adversarial_Networks_ECCVW_2018_paper.pdf (дата звернення: 12.03.2026).
4. Rassool R. VMAF Reproducibility: Validating a Perceptual Practical Video Quality Metric. IEEE BMSB 2017. URL: https://realnetworks.com/sites/default/files/vmaf_reproducibility_ieee.pdf (дата звернення: 12.03.2026).
5. Alliance for Open Media. AV1 Video Codec : вебсторінка. URL: <https://aomedia.org/specifications/av1/> (дата звернення: 12.03.2026).