

ОПТИМІЗАЦІЯ КОНВЕЄРУ ОБРОБКИ ЗОБРАЖЕНЬ НА БАЗІ АРХІТЕКТУРИ ARM CORTEX-M55 З ВИКОРИСТАННЯМ ВЕКТОРНИХ РОЗШИРЕНЬ HELIUM

Стрімкий розвиток концепції периферійних обчислень (Edge Computing) зумовлює перенесення задач обробки даних безпосередньо на кінцеві пристрої, такі як мікроконтролери. Це вимагає від вбудованих систем високої продуктивності при суворих обмеженнях на енергоспоживання та об'єм пам'яті. Вирішенням цієї проблеми є використання сучасних архітектур, зокрема ARM Cortex-M55 із технологією MVE (M-Profile Vector Extension, або Helium), що дозволяє виконувати паралельні обчислення (SIMD) безпосередньо на мікроконтролері. Проте, розкриття потенціалу такого апаратного забезпечення вимагає глибокої адаптації та векторизації алгоритмів.

Метою роботи є розробка та оптимізація конвеєру (Pipeline) алгоритмів комп'ютерного зору (перетворення у відтінки сірого, розмиття Гауса, виділення контурів оператором Собеля) для виконання на цільовій платформі Infineon PSoC Edge E84.

Основна вимога до проєкту полягає у мінімізації кількості тактів процесора (cycles per pixel) та підвищенні показника IPC (Instructions Per Clock) при збереженні прийнятної візуальної якості вихідного зображення. Вихідна наївна реалізація, побудована на скалярних операціях з плаваючою комою, є вкрай неефективною для мікроконтролера та потребує рефакторингу.

На етапі технічної пропозиції сформовано наступне проєктне рішення, що базується на відмові від арифметики з плаваючою комою на користь цілочисельної арифметики та максимальному використанні інтрінсіків MVE (наприклад, `vldrbq_u16`, `vmلاق_n_u16`):

- перетворення у відтінки сірого (Monochrome): заміна дійсних коефіцієнтів (0.3, 0.59, 0.11) на цілочисельні множники з подальшим логічним зсувом вправо. Це дозволить обробляти по кілька пікселів одночасно за допомогою однієї векторної інструкції;

- розмиття Гауса (Gaussian Blur): використання попередньо розрахованого ядра з цілочисельними вагами (`uint8_t`). Для уникнення деградації продуктивності через умовні переходи (розгалуження) всередині циклу пропонується просторове розділення обробки: окремий алгоритм для

країв зображення та векторизований цикл без умовних переходів для центральної частини;

- виділення контурів (Sobel Edge Detection): векторизація двовимірної згортки ядрами Собеля (G_x , G_y). Окрім розділення обробки аналогічної до розмиття Гауса та переходу від дійсних до цілочисельних розрахунків, для усунення апаратного «вузького місця» у вигляді складних математичних обчислень, пропонується використовувати швидкі апроксимації, наприклад, замість квадратного кореня при знаходженні сили градієнта ($\sqrt{G_x^2 + G_y^2}$) можна використати суму модулів: $|G_x| + |G_y|$;

- оптимізація пам'яті: для зменшення навантаження на системну шину та ефективного використання швидкої пам'яті (TCM), пропонується дослідити можливість злиття циклів (Loop Fusion) або використання рядкових буферів (Line Buffers).

Особливої уваги в процесі оптимізації заслуговує проблема затримок при роботі з пам'яттю (так звана проблема «Memory Wall»), яка часто стає головним вузьким місцем у вбудованих системах комп'ютерного зору. Традиційне послідовне виконання алгоритмів, що вимагає повного збереження проміжних масивів пікселів після кожного етапу обробки, призводить до надмірного навантаження на системну шину. Використання методів злиття циклів спільно з рядковими буферами дозволяє локалізувати оброблюване вікно даних у надшвидкій тісно зв'язаній пам'яті процесора (TCM – Tightly Coupled Memory). Це зводить до мінімуму кількість звернень до основної оперативної пам'яті та гарантує безперервне постачання даних для конвеєра векторних обчислень.

Крім того, перехід до 8- та 16-бітної цілочисельної арифметики у поєднанні з архітектурою MVE (Helium) дозволяє виконувати одразу декілька операцій множення з накопиченням (MAC) за один такт процесора. Такий апаратно-програмний підхід не лише сприяє радикальному зменшенню метрики циклів на піксель (cycles/px), але й суттєво підвищує загальну енергоефективність системи. Зниження питомих енерговитрат на обробку одного кадру є критично важливим фактором для забезпечення тривалої автономної роботи периферійних IoT-пристроїв та робототехніки.

В результаті впровадження запропонованих підходів очікується багатократне прискорення часу виконання конвеєру у порівнянні з наївною скалярною реалізацією. Практична цінність роботи полягає у створенні готового, апаратно-оптимізованого програмно-апаратного модуля, який може бути інтегрований у реальні системи машинного зору на базі мікроконтролерів (наприклад, системи розумного міста, промислова автоматизація, дрони).

В подальших дослідженнях планується реалізувати запропоновану архітектуру в коді, провести профілювання на фізичному обладнанні за допомогою апаратних лічильників продуктивності (PMU).

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. ARM Cortex-M55 Processor. ARM Developer. URL: <https://developer.arm.com/Processors/Cortex-M55> (дата звернення: 11.04.2026).