

В.М. ТЕРЕЩЕНКО

Київський національний університет імені Тараса Шевченка, Київ, Україна,
e-mail: *v.ter@knu.ua*.

П.В. ТИТАРЧУК

Київський національний університет імені Тараса Шевченка, Київ, Україна,
e-mail: *pavlo.tytarchuk@knu.ua*.

В.В. ГОРДІЙЧУК

Київський національний університет імені Тараса Шевченка, Київ, Україна,
e-mail: *vitaliyk-17@knu.ua*.

Я.В. ТЕРЕЩЕНКО

Київський національний університет імені Тараса Шевченка, Київ, Україна,
e-mail: *y.tereshchenko@knu.ua*.

СКЕТЧ-АНІМАЦІЯ З ТЕКСТУ: МЕТОДИ, КОНТРОЛЬ ТА МОБІЛЬНЕ РОЗГОРТАННЯ

Анотація. Розглянуто наявні підходи та методи генерації анімованих скетч-відео на основі текстового опису з урахуванням можливості запуску на мобільних пристроях та механізмів керованої генерації. Метою дослідження є огляд та аналіз сучасних методів, визначення критичних викликів та окреслення шляхів створення цілком автономних систем скетч-анімації на периферійних пристроях. Науковий внесок охоплює: структуровану таксономію методів за типом вхідних даних, механізмом керування та архітектурою; критичний аналіз придатності алгоритмів для роботи на мобільних платформах; визначення ключових технічних бар'єрів та напрямів подальших досліджень. Проведено аналіз публікацій за період 2023–2025 рр., що охоплює підходи на основі дифузійних моделей, фреймворки керованої генерації відео та методи оптимізації для мобільних систем, зокрема квантування, проріджування (pruning) та дистиляцію знань. Результати дослідження свідчать про те, що хоча сучасні методи забезпечують високу якість генерації, вони потребують значних обчислювальних ресурсів, які перевищують можливості мобільних пристроїв у 10–100 разів. Сучасні нейронні процесори мобільних пристроїв у поєднанні з методами компресії моделей відкривають перспективні шляхи прискорення обчислень, проте потрібні значні архітектурні інновації. Серед ключових прогалин у дослідженнях виокремлено часову узгодженість у легковагових моделях, компроміси між детальним контролем та ефективністю, а також обмеження щодо анімації множини об'єктів.

Ключові слова: скетч-анімація, генерація відео з тексту, дифузійні моделі відео, мобільне розгортання, керована генерація, виконання моделей на пристрої, компресія моделей, NPU.

DOI 10.34229/KCA2522-9664.26.5.3

ВСТУП

Індустрія анімації зазнає трансформаційних змін із появою технологій генеративного штучного інтелекту. Традиційне виробництво двовимірної (2D) анімації є трудомістким процесом, що охоплює розроблення персонажів, анімацію, створення проміжних кадрів та колоризацію [1].

Внаслідок конвергенції технологічних тенденцій цей огляд є особливо своєчасним. По-перше, дифузійні моделі досягли безпрецедентної якості у 2024–2025 рр. із системами Sora, Kling та Veo, що демонструють фотореалістичну генерацію відео зі складним розумінням руху [2–4]. По-друге, нейронні процесори мобільних пристроїв (NPU) досягли продуктивності 40–50 трлн операцій на секунду (TOPS), уможливаючи виконання обчислень (inference) дедалі складніших моделей безпосередньо на пристрої [5]. По-третє, підвищення доступності інструментів створення контенту сформувало попит на системи анімації, що не потребують професійних навичок чи високовартісного обладнання [6].

Мета дослідження полягає в аналізі сучасного стану генерації анімованих скетч-відео з текстового опису з фокусом на механізми керованої генерації, що