

СЕКЦІЯ 1 АКУСТИЧНІ ТА МУЛЬТИМЕДІЙНІ СИСТЕМИ

ОБҐРУНТУВАННЯ ВИКОРИСТАННЯ SPARK AR ДЛЯ СТВОРЕННЯ ІНТЕРАКТИВНОГО КОНТЕНТУ

д.т.н., проф. С.М. Порошин, І.О. Салатун, НТУ «ХПІ», м. Харків

У сучасних умовах розвитку соціальних мереж створення інтерактивного контенту, зокрема AR-масок, стає важливим елементом залучення користувачів та підвищення їхньої взаємодії з брендами. AR-маски дозволяють користувачам взаємодіяти з віртуальними об'єктами в реальному часі, змінювати свій вигляд та створювати унікальні візуальні ефекти, що робить їх популярними в таких соціальних мережах, як Instagram, Facebook та Snapchat. Задача створення таких масок передбачає використання програмного забезпечення, яке дозволяє розробникам інтегрувати елементи доповненої реальності безпосередньо в соціальні платформи.

Для реалізації цього завдання можна використовувати декілька інструментів, таких як Spark AR, Lens Studio та Adobe Aero. Проте найкращим рішенням для створення AR-масок для соціальних мереж є Spark AR.

Spark AR є однією з найпопулярніших платформ для створення контенту з використанням технологій доповненої реальності (AR), розробленою компанією Meta (раніше Facebook). Платформа була створена для того, щоб полегшити інтеграцію AR-контенту в соціальні мережі, зокрема Facebook та Instagram. Ця інтеграція дозволяє користувачам активно взаємодіяти з AR-масками й ефектами, що значно підвищує залученість аудиторії. Завдяки легкості використання та гнучкому функціоналу, Spark AR став важливим інструментом як для індивідуальних користувачів, так і для компаній, маркетологів, брендів і контент-креаторів.

Однією з основних переваг Spark AR є функція трекінгу обличчя, яка забезпечує точне відстеження мимічних рухів та виразів обличчя в реальному часі. Це дозволяє створювати реалістичні та динамічні маски, які адаптуються до змін виразу обличчя користувача. Відстеження обличчя є основним елементом багатьох AR-ефектів, що використовуються для косметичних фільтрів, зміни зовнішності або для створення інтерактивних ігор. Завдяки цій технології користувачі можуть отримувати персоналізований AR-досвід, а креатори — створювати контент, що активно реагує на зміни в поведінці користувача.

Іншою важливою особливістю платформи є інтеграція 3D-елементів, яка дозволяє розробникам додавати до AR-ефектів тривимірні об'єкти, що робить контент більш реалістичним та інтерактивним. Ця можливість дозволяє створювати віртуальні аксесуари, предмети або навіть персонажів, що взаємодіють з користувачем у реальному часі. Використання 3D-технологій значно розширює потенціал творчого самовираження, дозволяючи брендам і

креаторам впроваджувати інноваційні рішення для своїх проєктів.

Також важливою складовою Spark AR є підтримка звуку. Платформа дозволяє розробникам додавати звукові ефекти або реакції на звуки в реальному світі, що підвищує рівень взаємодії з користувачами. Наприклад, маски можуть реагувати на голосові команди або відтворювати звукові ефекти у відповідь на певні дії. Це відкриває нові можливості для створення інтерактивних AR-ефектів у музичних та маркетингових кампаніях, дозволяючи контенту інтегруватися з аудіовізуальним середовищем користувача.

Анімації у Spark AR є ще однією потужною функцією, яка дозволяє створювати динамічні візуальні ефекти в реальному часі. Інструменти для анімації дозволяють задавати складні моделі поведінки для 3D-об'єктів, створюючи ефекти з рухомими елементами, що реагують на дії користувача або зовнішні події. Це дає можливість інтегрувати віртуальні об'єкти в реальний світ з високим рівнем реалістичності, що значно розширює спектр можливостей для розробки контенту.

Окремо варто згадати, що Spark AR виділяється своєю доступністю для широкого загалу. Завдяки візуальному інтерфейсу та великій кількості готових шаблонів, навіть користувачі без програмістських навичок можуть створювати високоякісний AR-контент. Це суттєво знижує поріг входу в сферу доповненої реальності, залучаючи до неї не тільки професіоналів, але й аматорів, які прагнуть експериментувати з новими технологіями.

Однією з ключових сфер застосування Spark AR є маркетинг і реклама. Доповнена реальність стала важливою складовою багатьох рекламних кампаній, оскільки вона дозволяє створювати інтерактивний контент, який привертає увагу користувачів та забезпечує високий рівень залучення. Бренди можуть використовувати AR для створення віртуальних продуктів, фільтрів або конкурсів, які заохочують користувачів взаємодіяти з контентом та ділитися ним у соціальних мережах. Це забезпечує вірусний ефект і підвищує видимість бренду в цифровому просторі.

Крім того, Spark AR надає інструменти для аналізу ефективності AR-контенту, що є надзвичайно важливим для маркетологів і бізнесу. Ці інструменти дозволяють відстежувати, як користувачі взаємодіють з ефектами, скільки разів вони були використані, і якою була їхня реакція. Така аналітика допомагає оптимізувати стратегії взаємодії з аудиторією та підвищувати ефективність рекламних кампаній.

Таким чином, Spark AR представляє собою багатофункціональну та доступну платформу для створення AR-контенту, яка має великий вплив на сучасний цифровий світ. Її ключові функції, такі як трекінг обличчя, інтеграція 3D-елементів, підтримка звуку та анімації, дозволяють створювати високоякісний та інтерактивний контент, доступний для широкої аудиторії. Завдяки низькому порогу входу та потужному функціоналу, Spark AR стала

інструментом не тільки для професійних розробників, але й для бізнесу, маркетологів і креативних користувачів по всьому світу.

ВИКОРИСТАННЯ НЕЙРОННИХ МЕРЕЖ ДЛЯ ВИДІЛЕННЯ ВОКАЛЬНОЇ ПАРТІЇ З МУЗИЧНИХ КОМПОЗИЦІЙ

К.В. Бабак., к.т.н., доц. С.О. Шейко, ХНУРЕ, м. Харків

Розділення звуку на окремі джерела, як-то вокал і інструментальна частина, є складним завданням, оскільки сигнали часто змішуються на рівні гармоній, частот і ритму. Формально ця проблема відома як розділення джерел звуку або розділення сигналу (audio source separation). Вона полягає у відновленні або реконструкції одного або декількох вихідних сигналів, які в результаті лінійного або згорткового процесу змішані з іншими сигналами. У цій галузі досліджень багато практичних застосувань, у тому числі покращення якості звуку (мовлення) та усунення шуму, музичні ремікси, просторовий розподіл звуку, ремастеринг і т.ін. Звукоінженери іноді називають цю техніку розшаруванням (demixing).

Раніше для цього використовували класичні методи обробки сигналів, такі як спектральний аналіз і частотна фільтрація. Але ці методи були обмеженими і не завжди забезпечували необхідну якість, особливо коли вокал та інструменти перетинаються у частотному спектрі. Сьогодні для даної задачі починають використовуватися вступають нейронні мережі. З цієї теми є велика кількість ресурсів, від сліпого поділу сигналів з аналізом незалежних компонентів (ICA) до напівкерованої факторизації невід'ємних матриць і закінчуючи пізнішими підходами на основі нейронних мереж.

Для розділення вокалу і «мінуса» використовують кілька архітектур нейронних мереж, серед яких найпопулярнішими є рекурентні нейронні мережі (RNN), згорткові нейронні мережі (CNN) та їх комбінації. Також популярними є архітектури на основі U-Net та Wave-U-Net. Ці моделі працюють на основі спектрограм – візуалізації звуку в частотно-часовому вимірі (рис. 1), що дозволяє мережі виділяти необхідні патерни вокалу та інструментів.

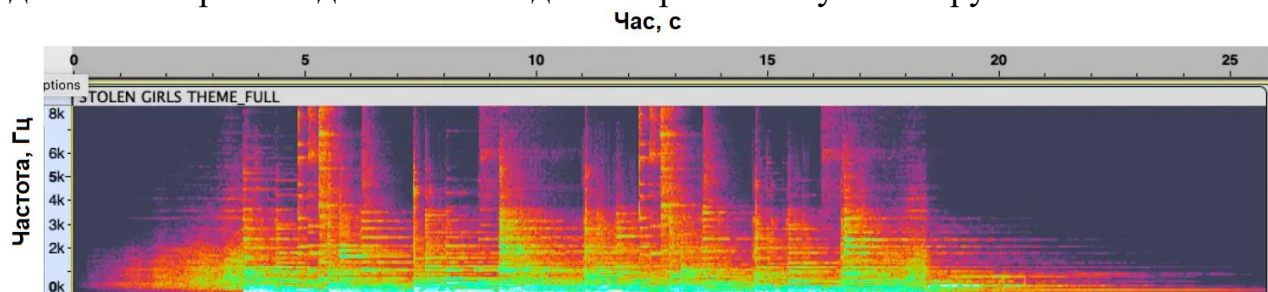


Рис. 1. Приклад спектрограми

Зі спектрограми ми можемо визначити три основні елементи:

– фундаментальна частота (f_0), яка визначається частотою вібрації голосових зв'язок;