

РОЗРОБКА ПРОГРАМНОЇ СИСТЕМИ АДАПТИВНОГО МУЛЬТИМЕДІЙНОГО СУПРОВОДУ ДЛЯ ЕЛЕКТРОННИХ ВИДАНЬ

Вінницький національний технічний університет

Анотація

У доповіді досліджується актуальна науково-практична проблема відсутності динамічної інтерактивності під час взаємодії користувачів із сучасними електронними виданнями. Запропоновано інноваційну архітектуру настільної програмної системи, яка здатна в режимі реального часу здійснювати семантичний аналіз текстового контенту та на його основі генерувати адаптивний фоновий аудіосупровід. Використання математичного апарату розрахунку частоти термінів дозволяє системі розпізнавати емоційні тригери, а алгоритми лінійного кросфейдингу забезпечують плавне перемикання мультимедійних сцен. Запропоноване рішення суттєво підвищує рівень занурення читача у літературний твір та розширює функціональні межі програмних засобів.

Ключові слова: електронні видання, мультимедійний супровід, частота терміна, кросфейдинг, обробка природної мови.

Abstract

The report investigates the urgent scientific and practical problem of the lack of dynamic interactivity during user interaction with modern electronic publications. An innovative architecture of a desktop software system is proposed, which is capable of performing semantic analysis of text content in real time and generating adaptive background audio accompaniment based on it. The use of the term frequency mathematical apparatus allows the system to recognize emotional triggers, and linear crossfading algorithms ensure smooth switching of multimedia scenes. The proposed solution significantly increases the reader's level of immersion in a literary work and expands the functional boundaries of software tools.

Keywords: electronic publications, multimedia accompaniment, term frequency, crossfading, natural language processing.

ВСТУП

Сучасна парадигма споживання цифрового контенту вимагає еволюції класичних форматів електронних видань. Стандартизовані цифрові контейнери забезпечують гнучкість візуального відображення тексту, проте концептуально залишаються статичними і позбавлені нативних механізмів для динамічної мультимодальної взаємодії з користувачем [1]. Відсутність інтерактивного звукового фону, який би синхронно корелював із сюжетною лінією твору, обмежує можливості когнітивного занурення читача. Актуальним науково-інженерним завданням є розробка програмного забезпечення, здатного здійснювати безперервний семантичний аналіз тексту на стороні клієнта та генерувати контекстний мультимедійний супровід без залучення сторонніх хмарних обчислень.

МАТЕМАТИЧНА МОДЕЛЬ АНАЛІЗУ ТА АУДІОМІКШУВАННЯ

Отриманий з презентаційного рівня неструктурований текст проходить етапи глибокої токенизації та жорсткої фільтрації шумової лексики. Ключовим етапом семантичного аналізу є визначення статистичної значущості кожного слова на основі інформаційно-пошукових метрик [2]. Для ідентифікації домінуючого емоційного маркера на сторінці застосовано алгоритм розрахунку локальної частоти терміна. Математична модель даного процесу описується формулою:

$$TF(t, d) = \frac{f_{t,d}}{(\sum_{t' \in d} f_{t',d})}$$

де, $TF(t, d)$ — частота терміна (показник того, наскільки часто конкретне слово зустрічається на сторінці); t — конкретне слово (токен), яке аналізується системою в даний момент (наприклад, "меч");

d (document) — поточний масив тексту (у нашому випадку — видима сторінка електронної книги); $f_{t,d}$ (frequency) — кількість фактичних входжень слова t на поточній сторінці d ; $(\sum_{t' \in d} f_{t',d})$ — загальна сумарна кількість усіх слів на поточній сторінці (знаменник дробу).

Отримані скалярні величини проходять етап компарації: обчислювальне ядро формує вектор валідних семантичних маркерів, відкидаючи статистично незначущі лексичні одиниці. Сформований масив тригерів ініціює роботу підсистеми генерації аудіосупроводу. Головною інженерною проблемою при динамічній зміні звукових доріжок є виникнення акустичних артефактів. Для забезпечення безшовного переходу розроблено алгоритм плавного перехресного загасання (crossfading) [3]. Даний метод керує пулом віртуальних аудіоканалів, амплітуда яких лінійно модулюється в часі. Розрахунок результуючої амплітуди звукового сигналу у динаміках користувача під час переходу між сценами описується так:

$$S_{out(t)} = (1 - \alpha(t)) \cdot S_{1(t)} + \alpha(t) \cdot S_{2(t)}$$

де, $S_{out(t)}$ — результуюча амплітуда звукового сигналу, який чує користувач у динаміках у момент часу t ; t — поточний момент часу в процесі переходу між двома аудіотреками; $S_{1(t)}$ — амплітуда (гучність) старого фонового треку, що завершується; $S_{2(t)}$ — амплітуда (гучність) нового фонового треку, що починає грати; $\alpha(t)$ — динамічний коефіцієнт мікшування, який лінійно зростає від 0 до 1 протягом усього часу переходу; $1 - \alpha(t)$ — коефіцієнт згасання, який автоматично зменшує гучність старого треку від максимуму до нуля.

ВИСНОВКИ

Розроблена гібридна програмна система успішно вирішує науково-прикладну проблему статичності сучасних електронних видань. Імплементація наведених математичних моделей дозволила створити високоефективну систему локального аналізу контенту. Проведене тестування підтвердило здатність системи безпомилково ідентифікувати ключові емоційні тригери та здійснювати безшовне мікшування аудіопотоків. Запропоновані технічні рішення відкривають нові перспективи для створення інтерактивних читацьких середовищ.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Salmela M., Koskimaa R. *Digital Reading Environments and Interactive Narratives*. London : Palgrave Macmillan, 2023. 245 p.
2. Jurafsky D., Martin J. H. *Speech and Language Processing*. 3rd ed. Stanford : Stanford University Press, 2024. 638 p.
3. Lerch A. *An Introduction to Audio Content Analysis*. 2nd ed. Hoboken : Wiley-IEEE Press, 2022. 368 p.

Мазур Максим Сергійович, студент групи 4ПІ-22Б, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький національний технічний університет, м. Вінниця, email: flash23100@gmail.com

Дудник Олександр Олександрович, науковий керівник, кандидат технічних наук, доцент кафедри програмного забезпечення, Вінницький національний технічний університет, м. Вінниця.

Mazur Maksym Serhiyovych, student of group 4PI-22B, Faculty of Information Technologies and Computer Engineering, Vinnytsia National Technical University, Vinnytsia, email: flash23100@gmail.com

Dudnyk Oleksandr Oleksandrovich, scientific supervisor, PhD, Associate Professor of the Software Department, Vinnytsia National Technical University, Vinnytsia.