

СТВОРЕННЯ КОРПУСУ МУЗИЧНИХ ТВОРІВ НА ОСНОВІ МЕТОДІВ КЛАСТЕРИЗАЦІЇ

Сторожук С.О., Далекій Ю.Л.

Науковий керівник – к.т.н., доц. Ткаченко О.М.

В сучасних комп'ютерних мережах основний об'єм трафіку припадає на мультимедійну, зокрема, аудіо інформацію. Зростання обсягу мультимедійної інформації, що передається і обробляється в комп'ютерних системах, зумовила необхідність автоматизації процесів аналізу і пошуку даних. В сучасних системах обробки аудіо інформації виникає необхідність швидкого пошуку музичних творів на основі аудіо контенту у віддалених базах даних (БД) великого розміру. Суть пошуку музичного твору на основі аудіо контенту полягає в тому, щоб автоматично отримувати файли аудіо записів музичних творів, подібних до заданого аудіо запису під час запиту.

Враховуючи великі обсяги аудіо інформації в БД, велике значення має швидкість пошуку. Тому зазвичай відліки аудіо сигналу ділять на невеликі частини (фрейми). Для кожного фрейму виконується спектральний аналіз, на основі якого обчислюється значення вектора параметрів. Таким чином, основою аналізу аудіо є обчислення векторів параметрів на фреймах. Вирішено було здійснювати параметризацію аудіо за допомогою мел-частотних кепстральних коефіцієнтів (MFCC).

В більшості випадків музичний твір характеризується певною періодичністю, що полягає в наявності ідентичних або дуже схожих за текстом та характером мелодії фрагментів. Відповідно можна говорити про надлишковість даних, якими описується музичний твір, і можливість скоротити кількість параметрів для його опису. З огляду на це доцільним є застосування методів кластерного аналізу. Це дозволить значно зменшити обсяги пам'яті, необхідні для їх зберігання. В даній роботі під час формування шаблонів музичних творів використовувався вдосконалений метод кластеризації k-середніх, який, на відміну від класичного, дозволяє отримати розв'язок, наближений до глобального мінімуму спотворення шляхом послідовного запуску k-середніх для центроїдів 1,2, ..., k.

Отже, в процесі формування шаблонів корпусу музичних творів необхідно було:

1. видалити з аудіо файлів тишу з початку та кінця записів;
2. поділити на фрейми по 20мс з перекриттям 10мс;
3. для кожного фрейму розрахувати вектор параметрів MFCC розмірності 13;
4. виконати кластеризацію векторів параметрів MFCC.

В результаті кожен шаблон було представлено 1000 кластерів MFCC.