

**Магістрант Конкевич А.О., к.т.н., доцент Волкогон В.П.**

**Національний технічний університету України  
«Київський політехнічний інститут»**

## **РОЗПІЗНАВАННЯ РУХОМИХ ОБРАЗІВ ДЛЯ КЕРУВАННЯ ВИКОНАННЯМ МУЗИЧНИХ ТВОРІВ**

### **Вступ**

Інтелектуальні системи в музиці знайшли широке застосування: це системи написання/обробки музичних творів, системи розпізнавання виконання для Midi-датчиків цифрової обробки сигналів (DSP), системи розпізнавання повного аранжування (багатьох інструментів) і багато інших.

Також при відтворенні фонограм застосовується багато неінтелектуальних інформаційних систем і обладнання— секвенсори MIDI, синтезатори, семплери, драм- машини та ін. Керування такими системами також примітивне — зазвичай на рівні "Пуск/Стоп".

У музиці застосовується два основні підходи до керування грою — сильно структуроване керування (виконання — натискання клавіш, притискання струн та ін.), звичайно застосовується до інструментів; слабо структуроване керування (жестове) — для спілкування музикантів між собою.

Розпізнавання жестикуляції має свою специфіку у порівнянні зі звичайним розпізнаванням зображень. Найперше, незначну роль грає ідентифікація предметів, що знаходяться в кадрі, значно важливіше їх переміщення. Рух предметів, окрім “аналогового” змісту, тобто впливу на якийсь параметр пропорційно величині руху, може мати також “символьний” зміст, тобто коли деякий закінчений рух має сенс окремої команди. Отже, потрібна математична модель, яка водночас може як ідентифікувати закінчені команди, так і визначати аналогові параметри поточного руху. Окрім того, розпізнавання має виконуватися швидко, атримка виходу повинна не перевищувати 20-40 мс, тому алгоритм має бути обчислювально оптимізований для можливості виконання у реальному часі. Конкретні рухи, які використовуються для подання музичних команд, часто є досить тонкими, тобто такими, коли незначні зміни фактичної траєкторії руху призведуть до зміни семантики руху. Тому жести не можуть бути жорстко закодовані в програмі, вони мають

запам'ятовуватися в результаті процесу навчання програми.

### **Постановка задачі**

Таким чином, необхідно розробити математичну модель, яка дозволить виділяти в рухомому образі окремі команди та аналогові величини, має можливість навчання на основі готових прикладів та може виконуватися на персональних ПК в режимі реального часу.

### **Виділення рухів**

У зв'язку з тим, що конкретний вигляд предметів, які рухаються, не важливий у контексті поставленого завдання, та враховуючи вимоги обчислювальної оптимізації, доцільним буде не ставити завдання розпізнавання нерухомих образів, а зосередитися на аналізі руху зображення. Першим завданням стає виділення руху із серії зображень. Існуючі алгоритми для цієї задачі можна поділити на дві категорії: ті, що виділяють окремі дискретні величини із даних (наприклад, характеризують рух одним із чотирьох напрямків) та ті, що використовують всі наявні дані. Перші можуть ефективно застосовуватися до таких задач, як т.зв. “жести миші”, коли рухам курсора ставиться у відповідність деяка команда. Друга категорія, як правило, на виході дає карту переміщення, що задає відповідність між точками двох сусідніх кадрів. Знаходиться ця карта виходячи з методу найменших квадратів шляхом мінімізації відмінності реального наступного кадра та попереднього кадра, модифікованого згідно з картою [1]. В нашому випадку доцільним є використання другого алгоритму, оскільки перший є надто грубим і може призвести до втрати багатьох нюансів руху, які в межах нашої задачі є важливими.

### **Розпізнавання рухів за допомогою прихованих марківських моделей**

Для визначення семантики руху його найперше слід віднести до однієї з категорій рухів (одиничних команд чи рухів, що задають аналогові величини). Оскільки такі випадкові процеси, як жести, можуть гарно бути наближені марківськими моделями, то для даного підзавдання найбільш підходить метод прихованих марківських моделей, який дозволяє проводити таку класифікацію без моделювання внутрішнього стану ланцюга Маркова [1]. Існує також узагальнення цього алгоритму на багатовимірні випадки, яке якраз відповідає нашому завданню[2]. Для обчислення використовується метод прямого ходу для прихованої

марківської моделі. Після обчислення ймовірностей отримання фактично отриманого жесту з моделі кожного жесту, виконується Байєсова класифікація жесту, в результаті якої він відноситься до одного із відомих рухів.

Нехай маємо наступні позначення:  $V = \{v_1, \dots, v_M\}$  — множина можливих переміщень;  $S = \{S_1, \dots, S_N\}$  — множина внутрішніх станів моделі;  $q_t$  — стан, в якому знаходиться модель у момент часу  $t$  (тобто  $q_t$  — одне из  $S_i$ );  $o_t$  — переміщення, що спостерігається в момент  $t$ ;  $O = \{o_1, \dots, o_T\}$  — послідовність, що спостерігається;  $\pi = \{\pi_i\}$  — розподіл ймовірностей вибору початкового стану, тобто  $\pi_i = P(q_1 = S_i)$  — ймовірність того, що в початковий момент  $t = 1$  система буде перебувати в стані  $S_i$ ;  $a_{ij}$  — умовна ймовірність  $a_{ij} = P(q_t = S_j | q_{t-1} = S_i)$ ;  $A = \{a_{ij}\}$  — матриця ймовірностей переходу — квадратна матриця розміру  $N \times N$ ;  $b_j(k)$  — ймовірність того, що в стані  $S_j$  спостерігається об'єкт  $v_k$ , тобто  $b_j(k) = P(o_t = v_k | q_t = S_j)$  — матриця  $N \times M$ . Прихованою Марківською моделлю  $\lambda$  будемо називати набір  $\lambda = (A, B, \pi)$ . Ця модель може породжувати послідовність  $O = o_1, o_2, \dots, o_T$ , яку будемо називати послідовністю, що спостерігається. Введемо пряму змінну  $\alpha_t(i)$ , котру визначимо як значення ймовірності того, що на момент часу  $t$  спостерігалась послідовність  $o_1 o_2 \dots o_t$ , і в момент  $t$  система знаходиться в стані  $S_i$ :  $\alpha_t(i) = P(o_1 o_2 \dots o_t, q_t = S_i | \lambda)$ .

Значення прямої змінної обчислюються згідно з наступним алгоритмом:

1.  $\alpha_1(i) = \pi_i b_i(o_1), \quad 1 \leq i \leq N$
2. Для всіх  $t = 1, 2, \dots, T-1$ ;  $1 \leq j \leq N$ ,  $\alpha_{t+1}(j) = \left[ \sum_{i=1}^N \alpha_t(i) a_{ij} \right] b_j(o_{t+1})$
3. Обчислення ймовірності, що шукається:  $P(O | \lambda) = \sum_{i=1}^N \alpha_T(i)$ . [1]

### Виділення аналогових величин

Для виділення аналогових величин згаданий алгоритм потребує модифікації. А саме, аналогова величина представляється у вигляді окремих жестів, що характерні для різних значень цієї величини. Після проведення класифікації обчислюється зважена сума цих значень відповідно до отриманих ймовірностей [3].

## Визначення кінця жесту. Граничні випадки

На відміну від інших моделей розпізнавання жестів, в даному випадку отримання сумнівних станів, тобто станів, які не можна однозначно класифікувати як певний жест, є вкрай небажаним. Тому у випадках, які описаний вище класифікатор не може ідентифікувати однозначно, використовується або зважена сума наслідків виконання обох жестів, або вибирається один, більш ймовірний.

## Керування виконанням твору

Вихід моделі складається з двох потоків: потік подій та потік величин. Популярний музичний цифровий інтерфейс MIDI надає підтримку для обох видів потоків і може бути використаний як вихідний формат моделі. За рахунок цього можна скористатися вже готовими рішеннями з синтезу звуків на підставі отриманих команд. Налаштування виводу полягає в відображенні вихідних потоків на конкретні MIDI-канали та події.

## Висновки

В роботі описана модель розпізнавання та інтерпретації рухомих образів з метою виділення з них інформації, важливої для виконання музичного твору. За основний було обрано метод, заснований на моделюванні прихованих мереж Маркова. Результати роботи показують можливість ефективно вирішувати завдання розпізнавання жестів.

Розробка може застосовуватися для вбудови в музичну апаратуру, розробки тренажерів для диригентів та ін.

## Література

1. *Olivier Cappé, Eric Moulines, Tobias Rydén* (2005). Inference in Hidden Markov Models. Springer. pp. 347-356.
2. *Li J, Najmi A, Gray RM* (February 2000). "Image classification by a two dimensional hidden Markov model". IEEE Transactions on Signal Processing 48 (2): 517-533. pp. 2-6
3. *C. D. Manning, H. Schütze*. Foundations of Statistical Natural Language Processing. MIT Press, 1999, 680 p. pp.125-127
4. *Нуммелин Э.*, Общие неприводимые цепи Маркова и неотрицательные операторы. — М.: Мир. с. 201-210.