

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ ИМЕНИ Н.Э. БАУМАНА
(НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ)

Классификация методов определения ритмического рисунка и темпа цифровой музыкальной записи

СТУДЕНТ: ПЕТРОВА АННА АЛЕКСЕЕВНА

ГРУППА: ИУ7-76Б

РУКОВОДИТЕЛЬ: КИВВА КИРИЛЛ АНДРЕЕВИЧ

Цель и задачи

Цель — изучить основные существующие методы определения ритмического рисунка и темпа цифровой музыкальной записи.

Задачи:

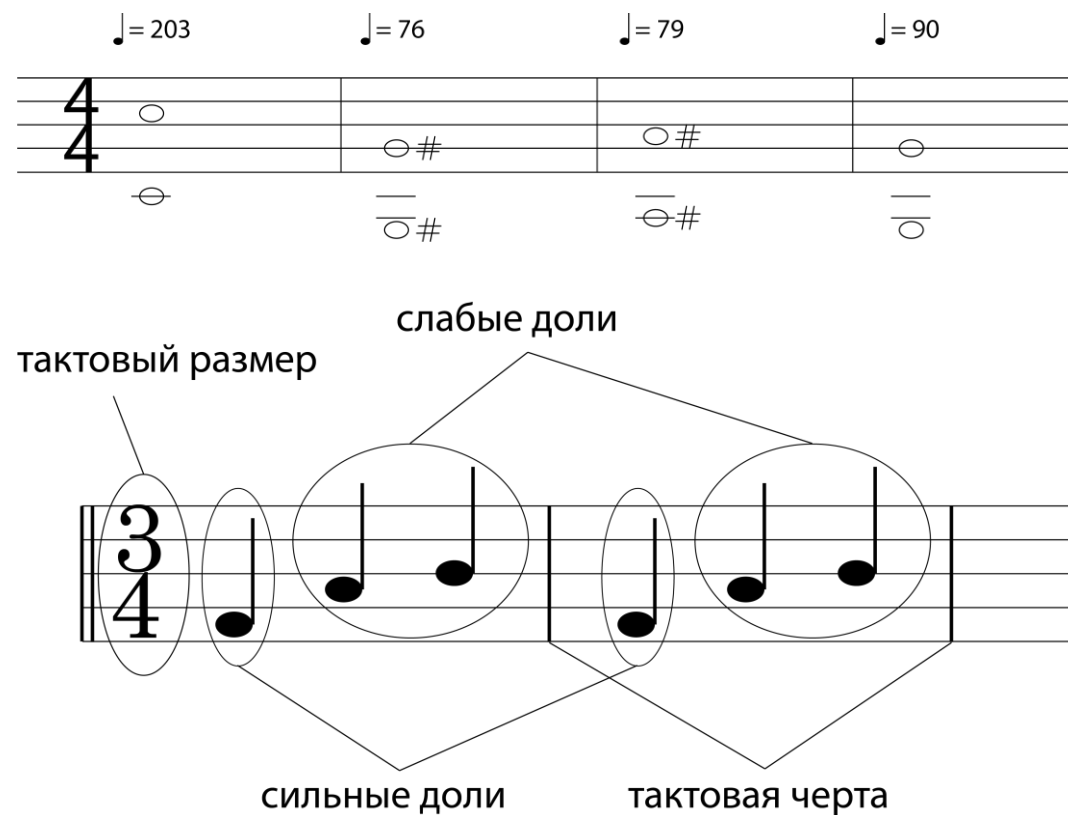
- провести анализ предметной области и сформулировать проблему;
- сформулировать критерии сравнения методов выделения информации о ритме и темпе музыки;
- классифицировать основные существующие методы.

Основные понятия

Темп – мера времени в музыке, упрощенно – «скорость исполнения музыки». Измеряется в bpm (число ударов в минуту).

Ритм – регулярная, периодическая последовательность акцентов. Такое определение ритма фактически идентично метру.

Метр – чередование сильных и слабых долей в определенном темпе. Численно фиксируется с помощью тактового размера.



Проблема определения ритма и темпа

- нечеткое попадание в ритм и темп на живых записях;
- переменный ритм и темп.

Критерии сравнения методов:

- точность результатов;
- определение переменного темпа и ритма;
- ограничения на формат входного аудиофайла;
- размеры датасетов (если обучение необходимо).

Дискретное вейвлет-преобразование

Преобразование Фурье (ПФ) \Rightarrow оконное ПФ \Rightarrow вейвлет-преобразование

Основная идея — разделение сигнала на высокие и низкие частоты с помощью фильтров.

$$y_{high}[k] = \sum_{n=-\infty}^{\infty} x[n]g[2k - n], \quad (1)$$

$$y_{low}[k] = \sum_{n=-\infty}^{\infty} x[n]h[2k - n]. \quad (2)$$

Скрытые модели Маркова

- n-граммная модель (длина ноты предсказывается исходя из предыдущих n-1 нот в вероятностном смысле);
- «ритмический словарь» (состоит из всех известных ритмических рисунков за единицу времени).

$Q = \{q_1, q_2, \dots, q_N\}$ – идеальные длительности нот, $X = \{x_1, x_2, \dots, x_N\}$ – наблюдаемые длительности нот.

$$P(Q) = p_{q_0} \prod_{t=1}^N a_{q_{t-1}q_t}, \quad (3)$$

$$P(Q|X) = \frac{P(X|Q)P(Q)}{P(X)}. \quad (4)$$

Байесовское иерархическое моделирование

- языковая модель;

- модификация нотных паттернов (добавление синкоп в модель);
- процесс Дирихле;

$\pi_{kk'}$ - вероятность перехода от паттерна k к паттерну k' . $\pi \sim Dir(\alpha\omega)$.

- модель представления;

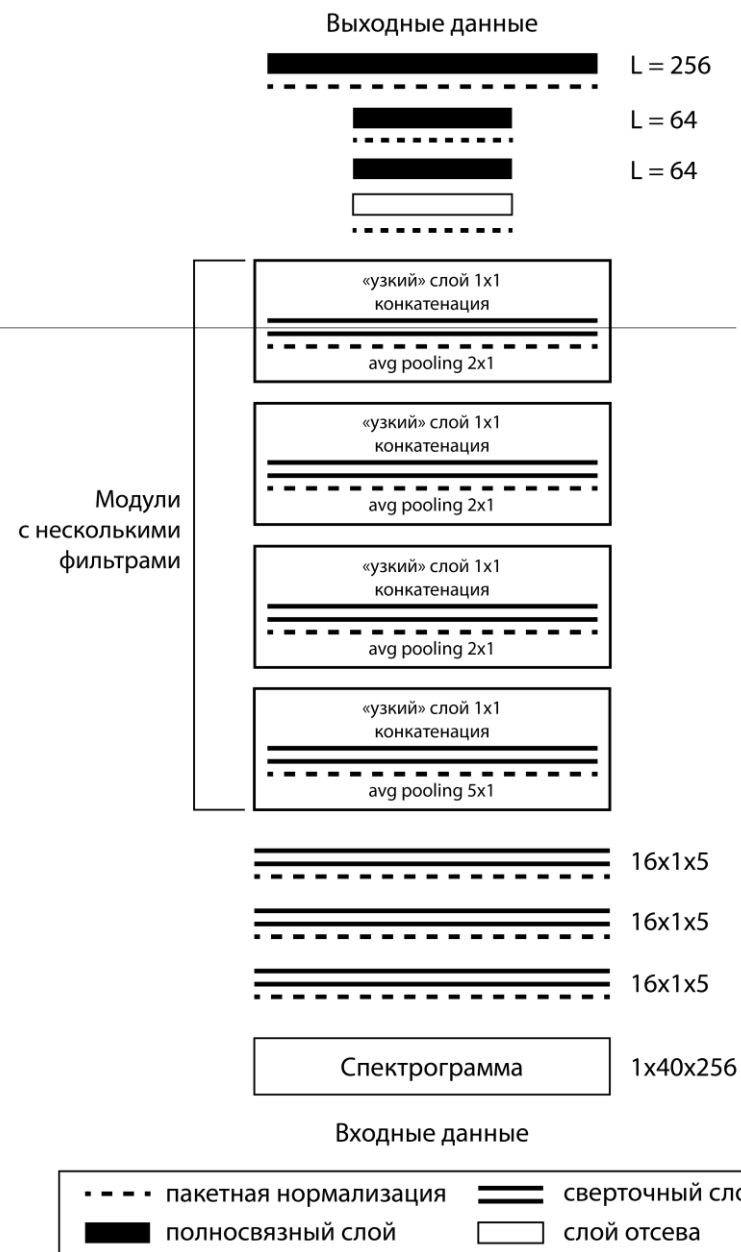
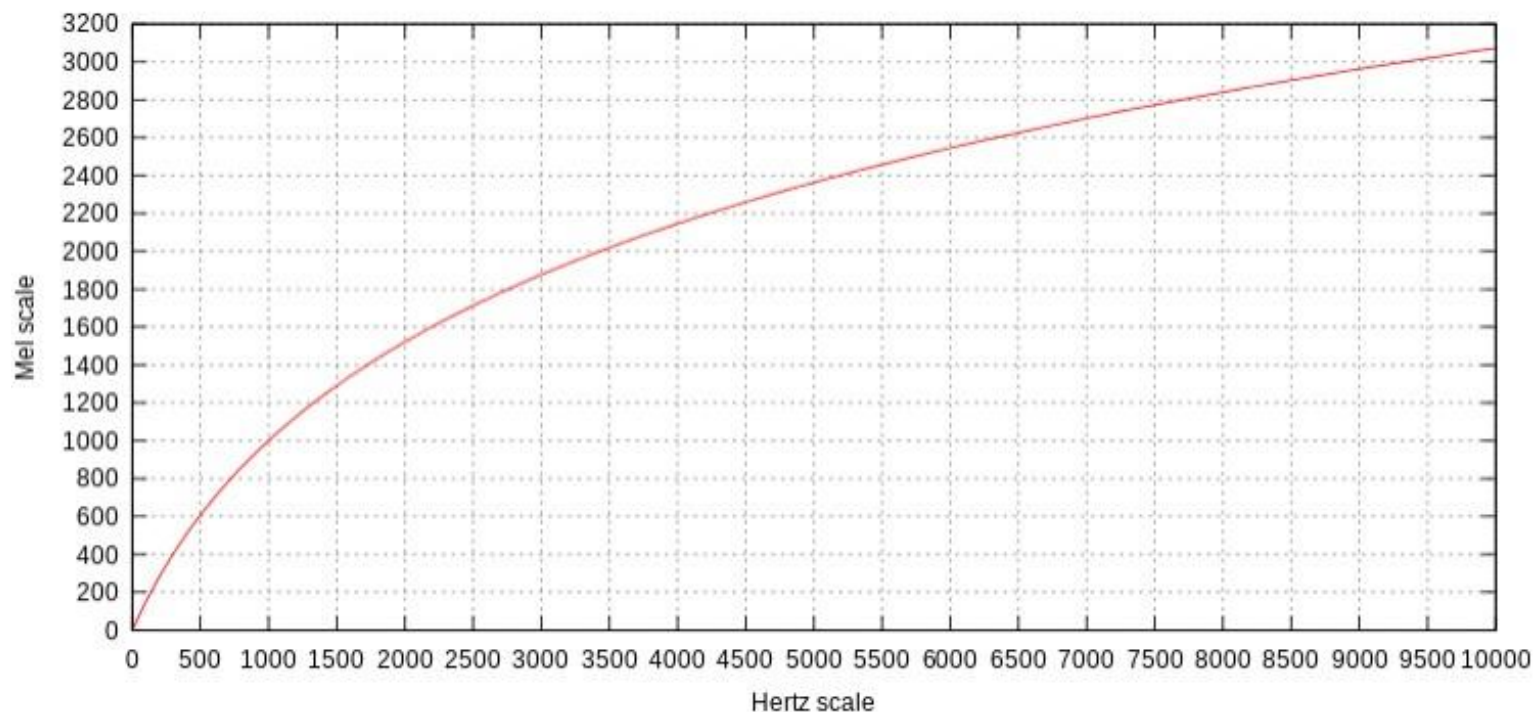
- колебания (неточности) темпа;

$$v_n | v_{n-1} \sim N(v_{n-1}, \sigma_v^2) \quad (5)$$

- колебания ритма.

$$d_n | v_n, x_n \sim N(v_n x_n, \sigma_t^2) \quad (6)$$

Сверточные нейросети



Сравнение методов

Метод	Точность результатов	Переменный темп и ритм	Формат входного аудиофайла	Размер обучающего датасета*
ДВП	Сильно зависит от жанра	Не определяются	Нет ограничений	Обучение не нужно
Скрытые марковские модели	Разный темп может определяться как одинаковый	Могут определяться при модификации метода	MIDI	88
Байес	Выше марковских примерно на 2%	Не определяются	MIDI	100
Сверточная нейросеть	До 98%	Не определяются	Нет ограничений	3611

**На основе данных из исследований*

Выводы

- рассмотрены понятия предметной области, такие как темп и ритм музыки, и проанализирована проблема автоматического определения темпа и ритма;
- определены критерии сравнения методов;
- рассмотрены основные методы автоматического определения темпа и ритма музыки: дискретное вейвлет-преобразование, скрытые марковские модели, байесовское иерархическое моделирование и сверточные нейронные сети;
- произведено сравнение изученных методов по выделенным ранее критериям.