Синхронизация нотных записей и музыки– это процесс автоматической обработки исполнения музыкальных композиций с целью отслеживания и отображения играемых нот в реальном времени.

Актуальность данной работы заключается в том, что в существующих системах синхронизации нет оценки того, сколько ошибок совершает музыкант. Целью данной работы было создание упрощенной методики оценки, способной функционировать на среднестатистическом компьютере пользователя в реальном времени.

Процесс синхронизации состоит из этапов, продемонстрированных на данном слайде. Однако наиболее сложными являются: Преобразование звукового потока в последовательность дискретных состояний какой-либо модели и сопоставления полученных временных последовательностей с расчетом позиции, темпа и оценки.

Для первого этапа были проанализированы следующие алгоритмы, работающие во временной и частотных областях, такие как пересечение с нулем, автокорреляция, Гармоническое перемножение спектров.

На основе анализа был реализован модифицированный алгоритм определения частоты основного тона на основе автокорреляции, поскольку для нее не требуется преобразование Фурье. Модификации заключались в разбиении звукового фрагмента на фрагмент низкого и высокого качества, а также в интерполяции Хермита для высоких частот. Номер ноты для последующего этапа вычисляется по данной формуле в равномернотемперированном музыкальном строе.

Для второго этапа были проанализированы алгоритмы сопоставления строк, динамической трансформации времени и, после анализа достоинств и недостатков рассмотренных алгоритмов, был выбран Алгоритм Витерби на Скрытой Марковской Модели.

СММ - статистическая модель, имитирующая работу процесса, похожего на Марковский, в которой задачей ставится разгадывание неизвестных параметров на основе наблюдаемых. В качестве состояний в данной задачи являются нотные события. Матрица A обозначают вероятности перехода из ноты i в ноту j, а матрица B – вероятность перехода из ноты i в ноту j в зависимости от высоты ноты k. pi означает матрицу старта процесса в определенной ноте. В данной работе были подобраны коэффициенты к упомянутым матрицам.

СММ в изначальном виде не полностью подходит к нотным записям, потому что в ней не учитываются длительности нот и ошибки добавления музыкантом неправильных нот. Для учета ошибок вставки были введены виртуальные состояния. Таким образом удалось предусмотреть ошибки, представленные справа: ошибки пропуска правильной ноты, проигрывания неправильной и добавления неправильной. Также виртуальные состояния оказались полезными в дальнейшем для разрабатываемой методики. Для учета различных длительностей нот, исходная запись разбивалась на множество событий одинаковой длины, которые были названы квантами. Стоит отметить, что самопереходы использовались только в виртуальных состояниях. Таким образом, общее количество состояний СММ значительно увеличилось до значения, вычисляемого по формуле.

Алгоритм Витерби в исходном виде не подходит для функционирования в реальном времени, так как история наблюдений увеличивается с каждой итерацией. Поэтому был разработан модифицированный алгоритм Витерби, учитывающий историю из нескольких последних наблюдений. Кроме того, матрица B была уменьшена до 13 состояний, в которых идентификаторы нот различных октав представлялись одним идентификатором, вычисляемого по формуле, а последнее состояние обозначало тишину. Также этот алгоритм была распараллелен. Данные модификации позволили уменьшить временную сложность алгоритма Витерби до следующего константного значения.

Для отслеживания ритма был разработан простой алгоритм, который базируется на вычислении среднего значения от последних восьми разностей между замерами времени не виртуальных квантов в возрастающей последовательности. Стоит отметить, что т.к. в данной работе рассматривался один слой виртуальных состояний, то данная последовательность будет состоять из состояний с нечетными номерами.

Несмотря на то, что существует алгоритм Прямого-обратного хода на СММ для вычисления вероятности специфической последовательности наблюдений, которыми являются входная последовательность нот, он не подошел по причине того, что в нем не учитываются виртуальные состояния, а также не учитываются нотные события. Более того, в разрабатываемой методике необходимо было учитывать темп, а также громкость звукового сигнала.

Таким образом, была разработана методика для оценки степени синхронизации, которая заключается в расчете 4 коэффициентов, вычисляющихся на каждой итерации алгоритма Витерби и зависящих от текущего и предыдущего кванта, нотного события, темпа и громкости сигнала. Данная методика позволила учитывать октавные ошибки с помощью анализа идентификатора ноты, а не только номера текущего кванта. Также эта методика позволила определять ошибки залигованных нот посредством анализа звукового сигнала при необходимости. Стоит отметить, что разработанная методика работает без расширения матрицы B, а значит не влияет на скорость работы алгоритма Витерби в реальном времени. В данной работе была решена задача подбора всех параметров в разрабатываемой методике.

Разработанная система состоит из 4 основных модулей: Модуль СММ, Модуль разбора нотных записей, модуль обработки звука и модуль отрисовки. Система разработана с учетом требований кроссплатформенности.

В качестве исходных данных для этого эксперимента моделировался идеальный синусоидальный сигнал, поэтому он не охватывает все случаи ошибок определения частоты основного тона. Из этих графиков можно сделать вывод, что модифицированный метод автокорреляции имеет наименьшую ошибку и практически постоянное время вычисления в зависимости от размера звукового фрагмента.

В качестве исходных данных в данном эксперименте являлась простая мелодия Twinkle Twinkle Little Star с темпом 120 ударов в минуту. Из графика слева можно сделать вывод, что при увеличении количества квантов в СММ, среднее время обработки, растет. Таким образом, эмпирически было выявлено оптимальное количество квантов равное 32 для темпа 120 ударов в минуту.

На графике справа представлена средняя ошибка алгоритма Витерби в реальном времени по сравнению с обычным алгоритмом. Из этого можно заключить, что при увеличении количества рассматриваемых квантов, процент ошибок алгоритма Витерби для нотных записей в реальном времени приближается к обычному алгоритму Витерби, а значит его можно использовать для данной предметной области.

Разработанная методика была протестирована на следующих мелодиях. И можно отметить, что она оказалась работоспособна для простых и коротких мелодий с небольшим количеством музыкальных орнаментов. Для сложных мелодий разработанная методика должна быть усовершенствована. Стоит отметить, что ошибки неправильного исполнения нот также влияют на штрафные баллы темпа, т.к. вычисление темпа зависит от виртуальных состояний, которые не принимаются в расчет. Однако ошибки темпа не влияют на позиционную оценку в нотах. Это может быть продемонстрировано в разработанном ПО.

В данной диссертации был реализован модифицированный алгоритм автокорреляции, реализован алгоритм Витерби в реальном времени, разработан алгоритм определения темпа, разработана методика оценка степени синхронизированности, реализовано программное обеспечение. По технологической части работы опубликованы 2 статьи. Также были проведены эксперименты и выбраны оптимальные параметры для алгоритмов. В качестве дальнейшего улучшения можно модифицировать алгоритм Витерби с помощью древовидных структур для дальнейшей минимизации ошибки, использовать алгоритм Баума-Велша для подбора оптимальных параметров матриц Скрытой Марковской Модели. Также разработанные алгоритмы и методику можно адаптировать для других областей развития человеческих способностей. Например, обучение произношения иностранных слов.