

ИСПОЛЬЗОВАНИЕ WAVELET-ПРЕОБРАЗОВАНИЯ ДЛЯ РЕШЕНИЯ ЗАДАЧИ РАСПОЗНАВАНИЯ ПТИЦ

Петровец В.Н., Мискевич П.Л., Деменковец Д.В.

*Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь*

Батура М.П. – доктор технических наук, профессор

В этом исследовании описывается подход к определению видов птиц по их пениям с применением wavelet-преобразования. Основное внимание уделено разработке и обучению модели, способной точно и эффективно идентифицировать вид птицы по звуковой записи её пения. Методика включает этапы подготовки данных и последующее тестирование модели. Предложенный подход может стать важной частью комплексного алгоритма для распознавания птиц.

Обычное преобразование Фурье (FFT) является мощным инструментом для анализа частотного содержания сигнала и широко используется в обработке звука. Однако для анализа акустических сигналов, таких как пения птиц, иногда оно может быть недостаточным из-за некоторых причин. Главной из них является ограничение в разрешении во временной и частотной областях. Преобразование Фурье предоставляет информацию только о частотах, присутствующих в сигнале, и не даёт подробной информации о временной динамике сигнала [1]. В случае пения птиц, которое часто характеризуется сложными временными и частотными модуляциями, это ограничение может быть проблематичным. Также следует учитывать, что использование FFT предполагает, что сигнал стационарен во времени, что может не выполняться для пения птиц, которое может быть динамичным и изменчивым [1]. Ввиду этого факта, преобразование Фурье является неустойчивым к динамичным изменениям в сигнале.

Для решения вышеперечисленных недостатков было решено использовать механизм wavelet-преобразования. Во-первых, wavelet-коэффициенты позволяют анализировать акустические сигналы, учитывая как временные, так и частотные характеристики [2]. Это особенно важно для пения птиц, которое может быть коротким и содержать различные частотные модуляции. Кроме того, полученные коэффициенты адаптивны к характеристикам сигнала, позволяя эффективно анализировать пения птиц с разной длительностью и спектральным составом. Использование wavelet-преобразования также обеспечивает высокое разрешение во временной и частотной областях, что позволяет более точно охарактеризовать сложные временные и частотные модуляции, присутствующие в пениях птиц.

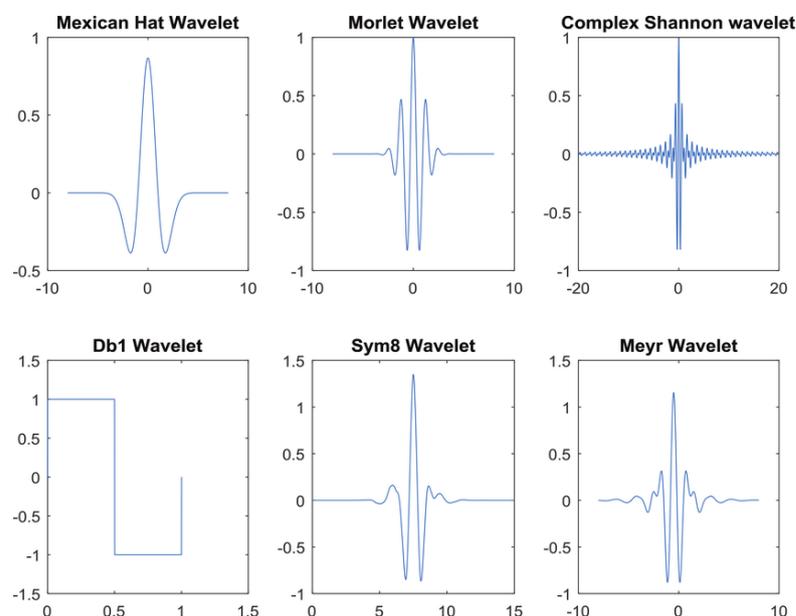


Рисунок 1 – Виды базовых функций wavelet-преобразования

Для исследования применимости данного механизма был разработан следующий алгоритм. Для начала исходный набор данных должен быть очищен от шумов и разделен на короткие сегменты или кадры, обычно длительностью от нескольких миллисекунд до нескольких секунд. Необходимо, чтобы каждый сегмент захватывал отдельный звуковой сигнал пения птицы или элемент вокализации. Также было предпринято решение по усилению высоких частот каждого сэмпла, но по итогам тестирования данное преобразование не принесло ощутимого эффекта. Далее необходимо применить непосредственно wavelet-преобразование к каждому аудио-сегменту, чтобы получить временно-частотное представление сигнала. Wavelet-преобразование разбивает сигнал на различные частотные

полосы на различных масштабах, захватывая как информацию о времени, так и о частоте. Результатом работы wavelet-преобразования является трехмерная матрица коэффициентов, отображающая спектральные составляющие звукового сигнала во временном измерении. Полученная матрица коэффициентов представлена на рисунке 1. В качестве базовой функции wavelet-преобразования была выбрана функция Морле [2]. Другие вариации wavelet-функций предоставлены на рисунке 2.

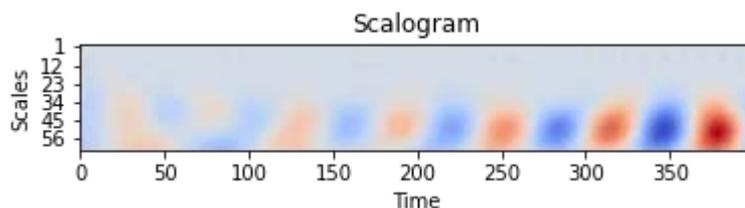


Рисунок 2 – Результат работы wavelet-преобразования, представленный в виде скалограммы

Следующим этапом необходимо обозначить характеристики из коэффициентов wavelet-преобразования. Эти характеристики могут включать статистические показатели (например, среднее, дисперсия, асимметрия, эксцесс) коэффициентов в каждом частотном диапазоне и масштабе. Кроме того, есть возможность извлечь такие характеристики, как энергия, энтропия или моменты, из коэффициентов преобразования для описания спектральных и временных свойств звуковых сигналов пения птиц.

Извлеченные wavelet-коэффициенты могут быть использованы в качестве входных данных для классификаторов машинного обучения, таких как машины опорных векторов (SVM), случайные леса или нейронные сети глубокого обучения. В рамках данного исследования было предпринято использовать машины опорных векторов. После обучения классификатора, его производительность была оценена при помощи тестового набора данных и перекрестной проверки.

После тестирования машины опорных векторов (SVM) на наборе данных аудиозаписей пения птиц была достигнута точность классификации на уровне 57%. Несмотря на то, что эта точность далека от идеальной, она всё равно указывает на потенциал использования wavelet-преобразования в качестве одного из признаков для распознавания видов птиц по их голосам. Этот результат демонстрирует, что акустические характеристики, извлеченные из аудиозаписей с помощью wavelet-преобразования, могут быть полезными для машинного обучения и классификации птиц по их пению.

Список использованных источников:

1. Ричард Лайонс. Цифровая обработка сигналов: Второе издание. Пер. с англ. / Ричард Лайонс – М.: ООО "Бином-пресс", 2006 г. – 656 с.
2. Wavelet theory and applications / Merry, R. J. E. // Technische Universiteit Eindhoven., 2005. – P. 7-10, 29.