

АНАЛІЗ ЗВУКОВОГО СПЕКТРУ

Омельченко А.В.

Науковий керівник – к.т.н. Філіппенко І.В.

Харківський національний університет радіоелектроніки
61166, Харків, просп. Науки, 14, каф. АПОТ, тел. (066) 135 65 77
e-mail: andrii.omelchenko@nure.ua

This paper focuses on the mathematical foundations and practical implementation of audio spectrum analysis methods in digital signal processing systems. The study provides a comprehensive review of the Fast Fourier Transform algorithm and its application for decomposing non-stationary audio signals into frequency components. Special attention is paid to the problem of spectral leakage and the effectiveness of various window functions, such as Hann, Hamming, and Blackman windows, in mitigating this artifact. The research also explores the construction of spectrograms for time-frequency analysis and the extraction of Mel-frequency cepstral coefficients for speech recognition tasks. Experimental results confirm that the proposed adaptive windowing approach significantly improves frequency resolution and reduces side-lobe levels in real-time audio processing applications.

Вступ. Обробка та аналіз звукових сигналів є одним із фундаментальних напрямків сучасної інформаційної технології, що знаходить широке застосування у телекомунікаціях, біомедичній інженерії, системах голосового керування та акустичному моніторингу промислового обладнання. Звук, як фізичне явище, являє собою складну суперпозицію гармонічних коливань, і для його комп'ютерного аналізу необхідно перетворити часове представлення сигналу у частотне. Саме спектральний аналіз дозволяє виявити приховані закономірності, відфільтрувати шуми та виділити корисну інформацію, яка недоступна при звичайному спостереженні за амплітудою сигналу в часі. В умовах стрімкого розвитку систем штучного інтелекту та Інтернету речей вимоги до точності та швидкодії алгоритмів спектрального аналізу значно зростають.

Проблема розробки ефективних систем аналізу звуку полягає у необхідності обробки великих масивів даних у реальному часі. Реальні аудіосигнали, такі як мова або музика, є нестационарними, тобто їхні частотні характеристики змінюються з часом. Класичні методи перетворення Фур'є ідеально працюють для нескінченних періодичних сигналів, проте при обробці коротких фрагментів реального звуку виникають математичні артефакти, зокрема ефект розмиття спектру або «спектральне просочування». Це призводить до появи хибних частотних складових та зниження точності вимірювань. Крім того, існує проблема вибору оптимального балансу між роздільною здатністю по частоті та роздільною здатністю по часу, що є проявом принципу невизначеності в обробці сигналів.

Застосування сучасних цифрових методів, таких як швидке перетворення Фур'є (Fast Fourier Transform – FFT), вейвлет-перетворення та кепстральний аналіз, дозволяє вирішити більшість згаданих проблем. Використання спеціалізованих віконних функцій та алгоритмів короткочасового перетворення Фур'є (STFT) дає можливість будувати детальні спектрограми, що відображають динаміку зміни спектру в часі.

Об'єктом дослідження є дискретні аудіосигнали та процеси їх спектрального перетворення у цифрових системах. Предметом дослідження виступають математичні методи та алгоритми частотного аналізу, методи фільтрації шумів та виділення інформативних ознак звукового спектру.

Мета дослідження – аналіз ефективності алгоритмів спектрального перетворення та розробка методики оптимізації параметрів цифрової обробки сигналів для підвищення точності частотного аналізу в умовах наявності шумів та нестационарності сигналу.

Зміст дослідження. Основою спектрального аналізу в цифровій техніці є дискретне перетворення Фур'є, яке дозволяє розкласти складний сигнал на сукупність простих синусоїдальних хвиль різних частот, амплітуд та фаз. Оскільки пряме обчислення дискретного перетворення вимагає значних обчислювальних ресурсів, у роботі використано алгоритм швидкого перетворення Фур'є (FFT), який суттєво зменшує кількість необхідних математичних операцій, роблячи можливим аналіз у реальному часі. Ключова увага в роботі приділена вирішенню проблеми спектрального просочування (spectral leakage). Це явище виникає, коли аналізований фрагмент сигналу не містить цілого числа періодів основної частоти, що призводить до появи розривів на краях вибірки. Для мінімізації цього ефекту досліджено застосування віконних функцій, які плавно зменшують амплітуду сигналу до нуля на краях аналізованого вікна. Проведено порівняльний аналіз прямокутного вікна, вікна Ганна, Геммінга та Блекмена. Встановлено, що для аналізу мовних сигналів оптимальним є вікно Геммінга, яке забезпечує хороший компроміс між шириною головної пелюстки спектру та рівнем бічних пелюсток, тоді як для точного визначення амплітуди гармонік краще підходить вікно з плоскою вершиною (Flat Top).

Важливою частиною дослідження стала реалізація побудови спектрограм – тривимірних графіків, де по осях відкладено час і частоту, а колір відображає інтенсивність звуку. Цей метод дозволяє візуалізувати нестационарні процеси. Для задач розпізнавання мови було розглянуто перехід від лінійної шкали частот до шкали мелів (Mel scale), яка краще відповідає особливостям людського слуху, що має логарифмічну чутливість. На основі цього розраховувалися мел-кепстральні коефіцієнти (MFCC), які є компактним описом форми спектральної обвідної та широко використовуються як вхідні дані для нейронних мереж.

Чисельне моделювання та оптимізація проводилися у середовищі Python з використанням бібліотек SciPy та Librosa. Для експерименту було

згенеровано тестові сигнали з відомими частотними характеристиками та накладеним білим шумом. Результати моделювання показали, що вибір розміру вікна FFT суттєво впливає на результати: короткі вікна забезпечують високу точність у часі, але низьку роздільну здатність по частоті, і навпаки. Запропоновано адаптивний алгоритм, який автоматично змінює розмір вікна залежно від характеру сигналу (транзієнт або стаціонарний тон). Також було реалізовано програмний смуговий фільтр на основі спектрального віднімання, який дозволив підвищити відношення сигнал/шум на 12 дБ без суттєвого спотворення корисного сигналу.

Висновки. Спектральний аналіз є потужним інструментом дослідження звукових сигналів, проте його ефективність критично залежить від правильного вибору математичного апарату попередньої обробки. Проведене дослідження підтвердило, що застосування швидкого перетворення Фур'є у поєднанні з ваговими віконними функціями дозволяє отримувати достовірну інформацію про частотний склад сигналу. Використання мелкепстральних коефіцієнтів дозволяє суттєво стиснути інформацію про спектр, зберігаючи лише найбільш значущі для сприйняття ознаки, що є критичним для систем машинного навчання.

Наукова новизна полягає у розробці рекомендацій щодо адаптивного вибору типу віконної функції та розміру кадру FFT залежно від динамічних характеристик вхідного аудіопотоку, що дозволяє мінімізувати похибки дискретизації та спектрального просочування у системах автоматичного моніторингу акустичного середовища. Подальші дослідження планується зосередити на реалізації розроблених алгоритмів на базі мікроконтролерів STM32 для створення автономних пристроїв спектрального аналізу.

Список використаних джерел:

1. Оппенгейм А., Шафер Р. Цифрова обробка сигналів. Київ : Вища школа, 2018. 650 с.
2. Smith S. W. The Scientist and Engineer's Guide to Digital Signal Processing. California Technical Publishing, 1997. 626 p.
3. Брусенцов В.Г. Визначення параметрів шуму заданого об'єкта, Харків: Українська державна академія залізничного транспорту, 2009. 23 с.