

# Комп'ютерне моделювання алгоритмів діагностики голосової системи людини

Денисенко О. І., ORCID [0000-0001-5148-922X](https://orcid.org/0000-0001-5148-922X)

Національний технічний університет України

"Київський політехнічний інститут імені Ігоря Сікорського" ROR [00syn5v21](https://ror.org/00syn5v21)

Київ, Україна

**Анотація**—Серед методів діагностування патологій, найпопулярнішими з яких є неінвазійні, що є найбільш комфортними для пацієнтів. Одним з таких методів є спектральний аналіз формантного складу голосних звуків. Для цього необхідно розробити зручний та актуальний в сенсі наявних рішень по виявленню патологій апаратно-програмний комплекс. У даній роботі виконано порівняння методів спектрального аналізу голосних звуків та вибір найкращого варіанту, а також наведено вигляд графічного інтерфейсу отриманого програмного забезпечення для допомоги лікарям фониатрам.

**Ключові слова** — дискретне перетворення Фур'є; швидке перетворення Фур'є; обертон; форманта.

## I. ВСТУП

Фониатрія, як розділ оториноларингології, займається дослідженнями характеристик голосу, та, в разі появи патологій, його відновлення. Як і у інших галузях медицини, дуже велике значення має те, коли було діагностовано захворювання. В даній роботі обмежимося розглядом захворювань гортані. В цьому випадку використовується ціла низка методів діагностики - від перегляду скарг пацієнтів з перцептивною оцінкою голосу лікарем (суб'єктивна оцінка), до рентгенографії, комп'ютерної та магнітно-резонансної томографії гортані (об'єктивна оцінка).

Досить поширеним серед дослідників є спектральний аналіз голосового сигналу за допомогою персонального комп'ютеру. Власне, такий аналіз є елементом об'єктивного підходу. Спектрографія дає змогу визначати такі важливі акустичні параметри голосу, як амплітуда та частота основного тону, дослідження зон обертонів та зон високої або низької співочих формант, що є основою якісних характеристик голосу [1].

Відомо, що кожна фонема голосних звуків має на графіку спектру характерні посилення, які є постійними за їх центральними частотами і не залежать від частоти основного тону. Такі зони посилення обертонів отримали назву формант. Відомо, що кожна голосна фонема має дві основні формантні області. Перша обумовлена резонансом об'єму повітря в глотці, друга – резонансом об'єму повітря в ротовій порожнині.

Окрім зазначених основних формант у спектрі голосу можуть міститися ще 1-3 ділянки підсиленних обертонів, що є відчутними у процесі формування тембру, але не використовуються при діагностиці патологій голосового тракту.

На думку І. М. Литвака та Л. А. Варшавського [2], голосні фонеми «у», «о», «а», та «і» є одноформантними, «е» та «и» - двоформантними. М. А. Сапожков та Г. Фант [3] наводять дані із зазначенням трьох формантних областей.

В роботі [4] наведені результати досліджень акустичних параметрів голосового апарату вокалістів-професіоналів зі стажем роботи від 3 до 40 років, що мали скарги на якість голосу, але при звичайному фониатричному огляді не було виявлено значних відхилень від нормальних показників. Отримані результати дозволили зробити наступні висновки [4]:

- у групі без скарг мають місце стійкі співочі форманти, щільно насичені обертонами;
- у контрольній групі зі скаргами виявлено меншу стійкість формант та менш щільне насичення спектру обертонами;
- у співаків не завжди узгоджувалась максимальна кількість формант із серединою їхнього робочого діапазону, як було в контрольній групі.
- у низки вокалістів з задовільною суб'єктивною оцінкою голосу і відсутністю проявів патологій в гортані, сонограми вказували на те, що для підтримки необхідної інтенсивності звуку використовуються не резонаторні порожнини, а близька до граничної напруга голосових зв'язок;
- регулярність обертонових піків у спектрі майже завжди відповідає його якості.

Автори [4] відзначають, що вивчення голосового спектру може бути основним методом діагностики у визначенні стійких порушень голосу. Поряд з дос-

лідженням голосового поля та віброметрії, спектральний аналіз дає змогу визначити вид порушень голосового апарату професіоналів. Акустичне тестування на наявність високих і низьких співочих формант у голосі співака може мати велике діагностичне значення у визначенні його працездатності та професійного прогнозу, може служити критерієм діагностики стійких порушень голосу та застосування результатів діагностики на ранніх стадіях професійних захворювань гортані.

В даній роботі продемонстроване створення власного програмного забезпечення, кінцевою метою якого є оцінювання спектральних дігностичних ознак стану голосового апарату співака. Для розв'язання даної задачі таке програмне забезпечення має дозволити робити записи голосних вокальних фонем та демонструвати їхній спектр, кепстр та спектрограму, а також вимірювати час максимальної фонації.

## II. МЕТОДИ СПЕКТРАЛЬНОГО АНАЛІЗУ ГОЛОСНИХ ЗВУКІВ ВОКАЛЬНИХ СИГНАЛІВ

При комп'ютерному моделюванні алгоритмів діагностики голосової системи людини використано MATLAB [5] – технічне й обчислювальне середовище, розроблене для вирішення широкого кола питань розробки програмного забезпечення та наукових проблем будь-якої складності практично у всіх областях науки та техніки.

Пакет MATLAB [5] містить велику кількість наборів інструментів для аналізу аудіо даних. Одним з таких пакетів є Signal Processing Toolbox. Він містить готові функції та програми для роботи з дискретизованими сигналами, їх аналізу та обробки. Також він містить певний «генератор» цифрових фільтрів, що спрощує їх аналіз та проектування, надає інструменти для зміни частоти дискретизації сигналу, визначення та видалення математичного очікування та власне оцінки спектру сигналу. Нарешті, є можливість визначати обвідні сигналів, знаходити їх піки, проводити оцінку схожості, а також визначати відношення сигнал/шум та коефіцієнт спотворень.

Що стосується отримання графіків спектрів головних звуків вокальних сигналів, даний продукт надає змогу використовувати різні методи. В першу чергу вони базуються на дискретному перетворенні Фур'є (ДПФ) та швидкому перетворенні Фур'є (ШПФ).

При цьому всі методи отримання спектру поділяються на два класи:

- непараметричні (періодограма та метод Уелча);
- параметричні (авторегресійна модель, MUSIC та EV).

### A. Непараметричний періодограмний метод

Непараметричним [6] є періодограмний метод, котрий являє собою оцінку спектральної щільності потужності, що розраховується по відлікам єдиної реалізації випадкового процесу, тобто одного запису аудіо. Розрахунок періодограми відбувається за наступною формулою [7]:

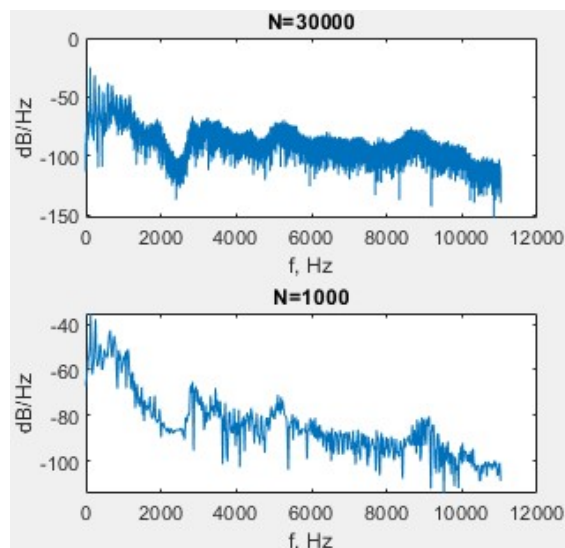


Рис. 1 Вигляд оцінки спектру методом періодограми при різних довжинах сегменту

$$\hat{W}(\omega) = \frac{1}{Nf_d} \left| \sum_{k=0}^{N-1} x(k) e^{-j\omega kT} \right|^2,$$

де  $x(k)$  - це послідовність відліків сигналу,  $f_d$  - частота дискретизації.

З ростом кількості відліків сегменту, що досліджується, значення періодограми все більше флюктують таким чином, що відповідний графік такої оцінки стає більш «поріганим». Вплив описаного ефекту видно на рис. 1.

Щоб зменшити цей ефект, необхідно вдатися до певного усереднення результату. Зазвичай пропонується згладити швидкі флюктуації спектру, отримані шляхом усереднення по сусіднім частотам. Цей метод реалізується шляхом обчислення згортки періодограми з віконною функцією.

Також відомий метод Бартлетта. За цим методом усереднення аналізований запис поділяється на сегменти без перекриття. Для кожного сегменту вираховується періодограма, й отримані періодограми усереднюються.

### B. Непараметричний метод Уелча

Метод Уелча є вдосконаленням методу Бартлетта. До вищевказаної процедури додається застосування до сегментів віконної функції та перекриття сегментів. Віконна функція дозволяє зменшити «розтікання» спектру. Завдяки введенню перекриття вдається компенсувати втрати інформації, спричинені застосуванням віконної функції.

Алгоритм дій для оцінки спектру методом Уелча має наступний вигляд:

- дискретний сигнал поділяється на сегменти з перекриттям;
- усі сегменти множаться на обрану вагову функцію;

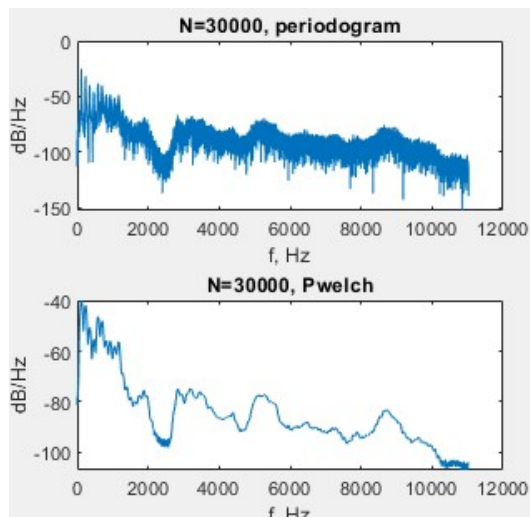


Рис. 2 Оцінка спектру методом періодограми та Уелча

- для отриманих сегментів обраховуються періодограми;
- отримані модифіковані періодограми усереднюються.

Порівняння звичайної періодограми та оцінки за методом Уелча представлено на рис.2.

#### С. Вагові функції-вікна в методі Уелча

Розглянемо вплив різних віконних функцій на графік спектру сигналу. Для створення вікон Бартлетта, Бартлетта-Ханна, Хеммінга, Блекмена, Гаусса-Віннера та Кайзера. Оберемо для всіх вікон розмір у 512 відліків (рис. 3) при частоті дискретизації 22050 Гц, що забезпечує роздільну здатність за частотою 43 Гц.

З рис. 3 видно, що вікно Кайзера ( $\beta = ?$ ) не дає достатньої інформації про посилені спектральні області (форманти), що свідчить про недоцільність його використання. У той же час, для вікон Бартлетта, Бартлетта-Ханна, Блекмана, Гаусса-Віннера та Хеммінга зони існування формант є достатньо виразними.

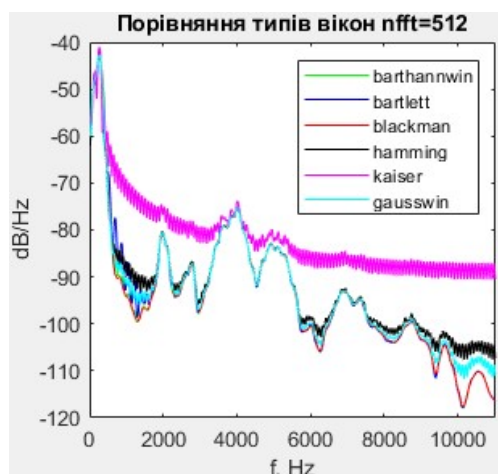


Рис. 3 Порівняння типів вікон для метода Уелча

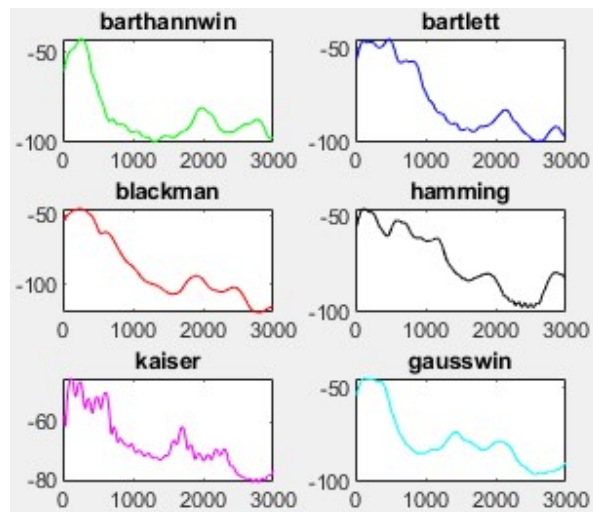


Рис. 4 Порівняння типів вікон для меншого розміру

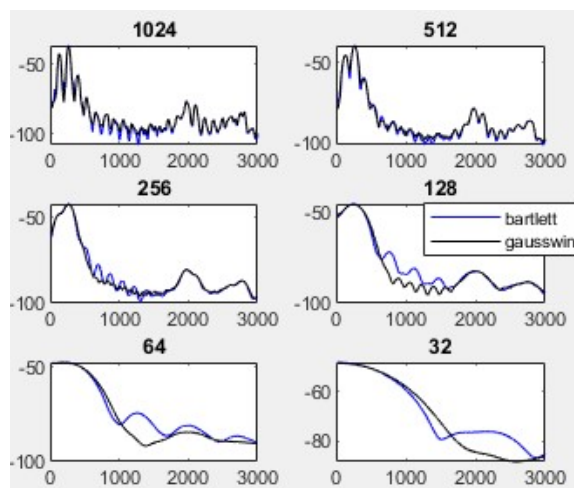


Рис. 5 Порівняння розмірів вікон для метода Уелча

Щоб наочніше розглянути це питання, масштабуємо графік до ділянки розміщення формантних областей – 20-3000 Гц. В той же час зафіксуємо розмір вікон у 256 відліків та переберемо різні голосні фонему (рис. 4).

З графіку видно, що при використанні вікон Кайзера та Хеммінга з'являються флуктуації. Кращий результат дають вікна Гаусса-Віннера та Бартлетта.

Для розгляду впливу розміру вагового вікна на графік спектру, створено набори вікон, які виділено на минулому етапі, розмірами від 32 до 1024 відліків (рис. 5). З цих графіків видно, що користувачеві необхідно залишити можливість вибору як типу вікна, так і його розміру, щоб при потребі обирати між ступенями гладкості форми спектру.

#### Д. Параметричний метод

Іншою групою методів спектрального аналізу є параметричні оцінки спектру. Відповідно сама оцінка спектру полягає в вирішенні задачі оптимізації параметрів створеної моделі спектру.

Серед існуючих параметричних методів найбільш поширеними є методи, що засновані на авторегресійній моделі формування сигналу. Це пов'язано з тим, що такі моделі є простими, зручними для розрахунків і добре відповідають реальним параметрам сигналів.

Відповідно до регресійної моделі, відліки  $x(k)$  утворюються шляхом пропускання відліків дискретного білого шуму  $n(k)$  через рекурсивний фільтр  $N$ -го порядку. В результаті спектральна щільність потужності такого сигналу розраховується за формулою:

$$W(\omega) = \frac{\sigma_n^2}{f_\omega \left| 1 - a_1 e^{-j\omega T} - a_2 e^{-j2\omega T} - \dots - a_N e^{-jN\omega T} \right|^2}$$

Отже, наведений метод зводиться до розрахунку коефіцієнтів регресійної моделі  $a_i$  заданого порядку  $N$ , оцінці потужності білого шуму  $\sigma_n^2$  та власне розрахунку спектральної щільності потужності.

Авторегресійні моделі переважно використовуються для сигналів, що подібні до авторегресійного процесу. Якщо очікуваний спектр сигналу, що досліджується, має чітко виражені сплески у спектральній області, авторегресійна оцінка спектру дає гарний результат в тому сенсі, що адекватно відображає місцеположення та інтенсивність таких сплесків.

Використання наданих інструментарієм MATLAB методів Берга та Юла-Уолкера дає змогу отримати досить гладку оцінку спектру при збереженні інформації про формантний склад.

Так само, як і для непараметричного методу Уелча, для реалізації поставленої задачі доцільно прослідкувати вплив порядку регресійної моделі та обрати оптимальні параметри для оцінки спектру мови авторегресійним методом Берга або Юла-Уолкера (рис. 6).

З графіків рис. 6 видно, що низький порядок моделі не дає змогу прослідкувати окремі спектральні обертони, тоді як результатом вибору зависокого порядку моделі є розщеплення спектральних компонент.

Таким чином, задача створення програмного забезпечення для спектрального аналізу голосових сигналів зводиться до необхідності створення  $m$ -функції для спектрального аналізу з можливістю вибору:

- діапазону аналізованих частот;
- початку та кінця досліджуваного сигналу;
- типу оцінки;
- порядку авторегресійної моделі у випадку параметричної оцінки;
- розміру вікна у випадку непараметричної оцінки;

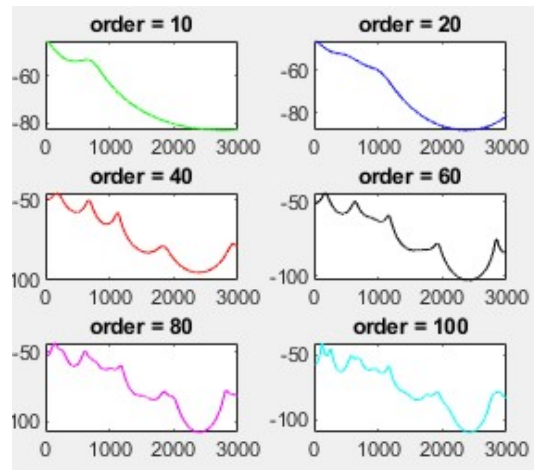


Рис. 6 Оцінка спектру голосної фонемі методом Берга при різному порядку моделі

- вибору масштабу (логарифмічного або лінійного).

До наведеного переліку доцільно додати видалення постійної складової та відсікання тиші від сигналів.

### III. ВИГЛЯД ГРАФІЧНОГО ІНТЕРФЕЙСУ

При розробці графічного інтерфейсу [8] було вирішено розділити робоче поле на дві частини – часову та спектральну. Це зроблено з метою контролю того, які саме дані є вхідними для програми. Це добре видно на графіку сигналу (рис. 7). Якщо запустити програму та активізувати позицію меню «Файл», отримуємо доступ до кнопки Open. Ця кнопка, як і кнопка «Записати», є основною, оскільки дає можливість ввести до програми сигнал, що буде піддано подальшому аналізу.

Інтерфейс спектральної частини програми показано на рис. 8.

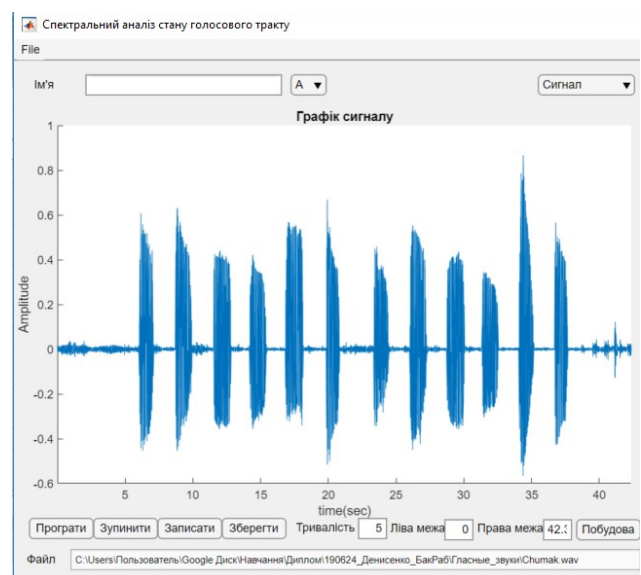


Рис. 7 Інтерфейс часової частини програми

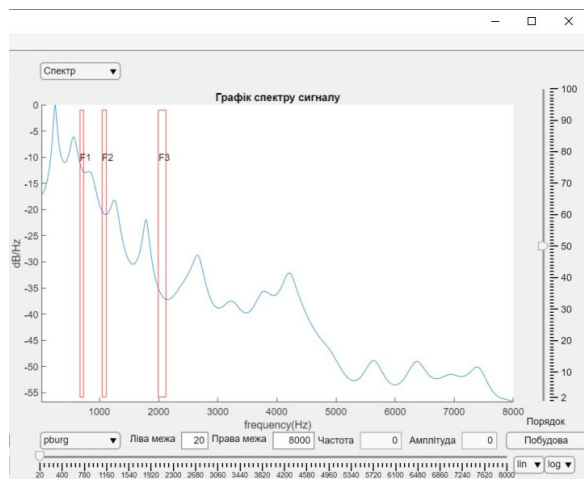


Рис. 8 Інтерфейс спектральної частини програми

Після цього потрібно обрати межі розташування однієї фонемі з їх набору. Для цього в поля «Права межа» та «Ліва межа» підбирається час кінця та початку запису фонемі.

Щоб створити новий запис фонемі наведена кнопка «Записати».

Щоб зупинити програвання запису, або зупинити створення нового запису з мікрофону, існує кнопка «Зупинити».

Створений або обрізаний аудіо файл можна зберегти на диску. Для цього потрібно вписати ім'я запису та обрати фонему з випадючого списку. Водночас цей список відповідає за побудову на графіку спектру формантних областей.

За необхідності кнопкою «Побудова» є можливість перебудувати графік запису сигналу відповідно до новообраних лівої та правої меж, а також заново нормалізувати графік для нових меж.

Як було вказано вище, є необхідність дати користувачеві вибір методу оцінки та його параметрів (розмір вікна або порядок моделі). Для цього є випадючий список та слайдер. Зміна одного з цих елементів інтерфейсу приводить до перебудови графіку спектру.

Подібно до графіку часової області є кнопка «Побудова» для вирівнювання графіку спектру по центру. Крім того є випадючі списки «lin» та «log», що відповідають за перемикання режимів шкал області значень та області визначення.

Порівняння комп'ютерного макету запропонованої діагностичної системи із прототипом, який наразі експлуатується в Інституті отоларингології ім. проф.

О.С. Коломійченка НАМН України [1], дозволяє зробити наступні висновки про переваги розробленої системи в частині спектрального аналізу голосних звуків вокальних сигналів:

- 1) до класу непараметричних оцінок спектру додано клас параметричних оцінок спектру, що дозволить полегшити прийняття рішення про формантний склад голосних звуків вокальних сигналів, що піддаються аналізу;
- 2) підвищено гнучкість спектрального аналізу вокальних сигналів наданням можливості вибору параметрів спектрального аналізу (розмір вікна та вид вікна даних – для непараметричних оцінок спектру, порядок авторегресійної моделі – для параметричних оцінок), а також можливості вибору лінійного або логарифмічного масштабу представлення графіків спектрів.

В подальшому доцільно оптимізувати розроблений графічний інтерфейс із врахуванням рекомендацій професійних фоніатрів [9], [10].

#### ВИСНОВКИ

Для розробки системи алгоритмів спектрального аналізу вокальних звуків виконано аналіз варіантів реалізації такої системи. Показано, що за основу доцільно обрати як параметричні, так і непараметричні методи спектрального аналізу. Зокрема, продемонстровано доцільність надання можливості користувачу обирати між параметричним методом Берга та непараметричним методом Уелча з вікнами Бартьєта або Гаусса-Віннера. Крім того, доцільною є можливість керування користувачем розміром обраних вікон або порядком регресійної моделі.

Показало доцільність керування користувачем розміром обраних віконних функцій при використанні непараметричних оцінок Уелча для забезпечення роздільної здатності за частотою від 20 Гц до 700 Гц. Також показано доцільність керування користувачем порядком авторегресійної оцінки Берга або Юла-Уолкера в інтервалі від 10 до 100.

За допомогою програмного середовища MATLAB розроблено програмну частину апаратно-програмного комплексу для аналізу стану голосового тракту. При цьому вдалося створити зручне та гнучке у виборі налаштувань програмне рішення, що готове для використання лікарем фоніатром у режимі тестування.

Варто зауважити, що безкоштовна бібліотека MATLAB Runtime дає принципову можливість скомпілювати та поширити отриманий продукт для використання фоніатрами.

#### ПЕРЕЛІК ПОСИЛАНЬ

- [1] Shydlovs'ka T. A. *Funktsional'ni porushennya holosu [Functional disorders of the voice]*, 2011
- [2] Pokrovs'kiy N. B. *Raschet i izmereniye razborchivosti rechi [Calculation and measurement of speech intelligibility]*. Moskva, Svyaz'izdat, 1962
- [3] Sapozhkov M. A. *Rechevoy signal v kibernetike i svyazi [Speech signal in cybernetics and communication]*. Moskva, Svyaz'izdat, 1963.



- [4] O. G. Pavlikhin, A. P. Meshcherkin "Diagnosticheskoye znacheniyе komp'yuternogo spektral'nogo analiza golosa u vokalistov [Diagnostic value of computer spectral analysis of the voice of vocalists]" *Sbornik nauchnykh trudov. Pervyy mezhdunarodnyy mezhdistsiplinarnyy kongress "Golos"*, 2007, pp. 86–90.
- [5] MathWorks - Makers of MATLAB and Simulink - MATLAB & Simulink URL: <https://www.mathworks.com/>.
- [6] Sergiyenko A. B. *Tsifrovaya obrabotka signalov [Digital signal processing]*. 2011. 768 p
- [7] Didkovs'kyi V. S., Didkovs'ka M. V., Prodeus A. M. "Komp'yuterna obrobka akustychnykh syhnaliv [Computer processing of acoustic signals]". Navchal'nyy posibnyk. Kyiv, «Imeks-LTD», 2010. 420 p.
- [8] Denysenko, O. I. *Spektral'nyy analiz vokal'nykh zvukiv [Spectral analysis of vocal sounds] : diplomna robota bakalavra* : 171 Elektronika. Kyiv, 2020. 75 s. URI : <https://ela.kpi.ua/handle/123456789/34773>
- [9] Pediatric otolaryngology. Edited by Charles D. Bluestone. 4th ed., 2003. 2214 p.
- [10] Poorjam A., Jensen J., Christensen M. Dominant Distortion Classification for Pre-Processing of Vowels in Remote Biomedical Voice Analysis. URL: <https://www.researchgate.net/publication/317064232>

Надійшла до редакції 27 березня 2022 р.

Прийнята до друку 21 квітня 2022 р.

UDC 534.4

# Computer Modeling of Human Voice System Diagnostic Algorithms

O. I. Denysenko, ORCID [0000-0001-5148-922X](https://orcid.org/0000-0001-5148-922X)

National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute» ROR [00syn5v21](https://ror.org/00syn5v21)

Kyiv, Ukraine

**Abstract** — Every year, people are faced with an increasing number of diseases that require timely detection and diagnosis without causing discomfort to the patient. Thus arose the systems of objective and subjective assessment of the quality of hearing, as well as the first systems of analysis of the state of the vocal tract. The problem of voice pathology is inherent in a fairly large risk group, which includes teachers, artists, call center operators. The specifics of the development of such systems is that the hardware and software complex developed by the engineer should be convenient for use by a doctor and relevant in terms of existing solutions to identify certain acoustic parameters of the voice.

The study of the voice spectrum can be the main diagnostic test in the diagnosis of persistent voice disorders and, along with the study of the voice field and vibrometry, allows you to determine the form of voice disorders in professionals. Acoustic tests for the presence of high and low singing forms in the singing voice can be highly important in determining the singer's performance and professional prognosis, can serve as a criterion in the diagnosis of persistent voice disorders, and their use in early stages of occupational laryngeal diseases will help prevent preventive measures.

MATLAB software comes with a large number of tools, which facilitates the implementation of many engineering, mathematical, computational issues of development and research of various processes related to any field of research. Based on these tools, there are a large number of basic functions of digital signal analysis, including audio signals.

FFT - fast Fourier transform algorithms are chosen as a basis, which with a certain modification are divided into parametric and nonparametric. In this case, the nonparametric Welch method and the parametric Berg method are chosen. The user is given access to choose between the parameters required for their operation. For the first it is the dimension of the weight function, for the second it is the order of the autoregressive model. All this gives opportunity to analyze the spectrum of vowel phonemes.

The AppDesigner package provides great opportunities for creating interfaces in software development. By manipulating the functions of Callback, you can bring the program to the finest settings, which at first may seem invisible, but generally create comfort when working. It is important to build a certain algorithm of action of each component. It often happens that it is necessary to take into account such details for which a certain component in itself is not responsible. For example, this could be changing the signatures of other components. To simplify code writing, it is important to create m-functions. However, this must be taken into account when editing them after entering them in the main script.

Using the above toolbox, a software interface was developed, which is divided into two working areas: time and spectral parts. In addition, the interface is filled with controls for input data and spectrum analyzer parameters, as well as spectrum analysis tools.

**Keywords** — discrete Fourier transform; rapid Fourier transform; overtone; the formant.

