

міжперіодної обробки сигналів дає змогу підвищити когерентність сигналів відбитих від поверхні однієї і тієїж цілі.

Література

1. Вопросы перспективной радиолокации. Коллективная монография / Под редакцией А.В.Соколова. – М.: Радиотехника, 2003 – 512с.
2. Кук Ч. Радиолокационные сигналы / Кук Ч., Бернфельд М.: Пер. с английского под ред. В.С. Кельзона. – М. "Советское радио", 1971. – 568 с.
3. Помехозащищенность систем радиосвязи с расширением спектра сигналов методом псевдослучайной перестройки рабочей частоты / Борисов В.И., Зинчук В.М., Лимарев А.Е., Мухин Н.П., Шестопапов В.И. - М.: Радио и связь, 2000. – 384 с.
4. Свистов В.М. Радиолокационные сигналы и их обработка / Свистов В.М. - М.: "Сов. радио", 1977. – 448 с.
5. Справочник по радиолокации. Под ред. М.Сколника. С 74 Нью-Йорк. 1970. Пер. с англ. (в четырех томах) под общей ред. К.Н.Трофимова. Том 1. Основы радиолокации. Под ред. Я.С.Иццоки. - М. Сов.радио, 1976. – 456 с.
6. Теоретические основы радиолокации. Под ред. Ширмана Я.Д. Учебное пособие для вузов. М., изд-во "Советское радио", 1970, 560 с.
7. Теоретические основы радиолокации: Учебное пособие для вузов / А.А. Коростелев, Н.Ф. Ключев, Ю.А. Мельник. Под ред. В.Е. Дулевича. - М.: "Сов. радио", 1978. – 608 с.
8. Финкельштейн М.И. Основы радиолокации: Учебник для вузов. Изд. 2-е, перераб. и доп. / Финкельштейн М.И. – М.: Радио и связь, 1983. – 536 с.

Надійшла 22.4.2011 р.

УДК 621

А.В. КУДРЯШОВ, К.Л. ГОРЯЩЕНКО

Хмельницький національний університет

ВИКОРИСТАННЯ ВЕЙВЛЕТ-ПЕРЕТВОРЕНЬ У АНАЛІЗІ ТА ВІДНОВЛЕННІ МОВНИХ СИГНАЛІВ ТА МОДЕЛЮВАННЯ В СЕРЕДОВИЩІ МАТЛАВ

Розглянуто задачу усунення спотворення звукових складових мовних повідомлень при застосуванні операції кодування та декодування із застосуванням алгоритмів із втратами, на прикладі MPEG Layer-3.

The task of removal of distortion of voice constituents of linguistic reports is considered at application of operation of code and decoding with application of algorithms with losses, on the example of MPEG Layer-3.

Ключові слова: вейвлет-перетворення, мовний сигнал.

Вступ

Для забезпечення відновлення сигналів, що несуть у собі акустичне повідомлення, потрібно в першу чергу звернути увагу на методи їх аналізу та представлення їх у багатовимірному просторі ознак, які містять потрібну для розпізнавання інформацію.

Серед існуючих методів побудови простору ознак (спектрального аналізу) можна виділити основні, такі як:

- Перетворення Фур'є;
- Лінійне передбачення мови;
- Кепстральний аналіз;
- Вейвлет-аналіз.

В останні роки стало очевидним, що традиційний апарат представлення довільних функцій та сигналів у вигляді рядів Фур'є виявляється малоефективним для функцій із локальними особливостями, частково, для імпульсних та цифрових сигналів та зображень, що отримали широке застосування [1, 2]. Це пов'язано з тим, що базисна функція рядів Фур'є – синусоїда ($y = U_0 \sin(\omega \cdot t)$) – визначена у просторі від $t = -\infty$ до $t = +\infty$ та за своєю природою є рівною та строго періодичною.

Таким чином, відомі методи аналізу сигналів та функцій постійно наштовхувались на принципові теоретичні обмеження, що не дозволяють серйозно говорити про кардинальне рішення проблеми однозначного їх представлення способами, створеними на основі розкладання у ряди Фур'є.

Із відкриттям вейвлетів ця складна і актуальна наукова проблема була вирішена. Основою стали розробки Гросмана та Морле, у середині 80-х років, як альтернатива перетворенню Фур'є для дослідження часових (просторових) рядів із вираженою неоднорідністю.

На відмінну від перетворення Фур'є, що локалізує частоти, але не дає часове розширення процесу, та від апарату d-функцій, що локалізує моменти часу, але не має частотного розширення, вейвлет-перетворення, яке наділене рухомим частотно-часовим вікном, що самоналаштується, однаково добре

виявляє як низькочастотні, так і високочастотні характеристики сигналу на різних часових масштабах.

Вейвлети мають вигляд коротких хвильових пакетів з нульовим інтегральним значенням та з тією чи іншою, досить часто дуже складною, формою, локалізованих на осі незалежною перемінною t , під якою розуміють незмінний час, здатних до зсуву по ній та масштабуванню (стисненню/розтягненню).

Мовне повідомлення, що використовується у людській мові, складається з окремих фонем, що складені послідовно між собою. Фонема – найменша (неподільна) структурно-семантична звукова одиниця, що здатна виконувати деякі функції у мовленні, так фонема творить, розділяє і розпізнає морфемі, слова, їхні форми в мовному потоці. "Фонема" є поняттям близьким до поняття "звук", однак має суттєву відмінність. Зміна у звукових складових фонемі веде до утворення іншої фонемі, а отже – зміни змістового навантаження на речове повідомлення в цілому. Тому стає зрозуміло, що питання передачі звукового повідомлення із мінімальними спотвореннями є принципово важливою задачею. Адже ціна питання – чіткість передачі змісту від одного до іншого користувача.

З технічної точки зору, мовне повідомлення, що складається з ряду фонем є послідовністю окремих звуків, а тому мовний сигнал – це є складний сигнал, що складається з багатьох сигналів різної амплітуди та частоти. При чому має місце постійна зміна як частотної так і амплітудної складової сигналу.

Модель звукового сигналу можна описати таким виразом:

$$S(t) = \sum_{i=20}^{20000} U_i(t) \cdot \sin(\omega_i \cdot t) \quad (1)$$

де $U_i(t)$ – амплітудна складова;

ω_i – частотна складова.

Межі виразу (1) від 20 до 20000 обумовлені загальним звуковим діапазоном, який людина здатна сприймати, а також тим звуковим діапазоном, що здатна створювати звукогенеруюча техніка.

Мовний сигнал $S(t)$ (1) відповідно до вейвлет перетворення буде мати вигляд [1]:

$$S(t) = \sum_l a_l \varphi_l(t) + \sum_k d_k \psi_k(t), \quad (2)$$

де: $\varphi_l(t)$ - масштабуючі або скейлінг-функції (phi-функції) з одиничним значенням інтегралу ($\int_{-\infty}^{\infty} \varphi(t) dt = 1$), що визначають грубе приближення сигналу і породжують коефіцієнти апроксимації a_l .

$\psi_k(t)$ - вейвлет-функції (psi-функції) з нульовим значенням інтегралу ($\int_{-\infty}^{\infty} \psi(t) dt = 0$), що

визначають деталі сигналу і породжують коефіцієнти деталізації d_k .

Phi-функції $\varphi_l(t)$ притаманні далеко не усім вейвлетам, а тільки тим, котрі відносяться до ортогональних. Ортогональність, як прийнято в функціональному аналізі, вводиться через скалярний добуток, який позначається $\langle \rangle$ [1]:

$$\langle \varphi_l, \varphi_i \rangle = 0 \text{ при } l \neq i;$$

$$\langle \psi_k, \psi_j \rangle = 0 \text{ при } k \neq j;$$

$$\langle \varphi_l, \psi_k \rangle = 0 \text{ при } l \neq k.$$

Psi-функції $\psi_k(t)$ створюються на основі тієї чи іншої базисної функції $\psi_0(t)$, яка задовольняється властивостями зміщення по осі часу і масштабування та визначає тим вейвлета.

Можна перекопатись, що обидві властивості об'єднуються виразом [1]:

$$\psi(t) = a^{-1/2} \psi_0\left(\frac{t-b}{a}\right),$$

де числовий параметр a задає ширину вейвлета; а числовий параметр b задає його положення на осі часу.

Вейвлети, які задаються частотним представленням задаються Фур'є-образами. У цьому випадку мова йде про огинаючу спектра вейвлета, зосередженого на осі частот. Накладаєма на функцію $\psi(t)$ умова (нульове значення інтегралу) вказує на зміщення Фур'є-образа по осі частот у сторону нуля із зосередженням навколо деякої частоти ω_0 , яку можна розглядати як *середню кругову частоту* вейвлета. У частотній області спектри багатьох вейвлетів нагадують виплески, які швидко прямують до нуля на краях. У цьому випадку говорять про *ширину смуги частот*, яку займає вейвлет.

Для сигналів обмеженої енергії $S(t)$ *пряме неперервне вейвлет перетворення* задається формулою [1]:

$$C(a, b) = \frac{1}{\sqrt{a}} \int_V s(t) \psi\left(\frac{t-b}{a}\right) dt,$$

де інтервал V відповідає області зміни безперервного часу t .

Результатом перетворення буде безперервна функція двох аргументів $C(a, b)$. Звичайно, в комп'ютерних технологіях мови про безперервність речового сигналу бути не може. Здійснюючи заміну безперервного часу на дискретне $t \rightarrow nT$, отримаємо дискретний аналог формули [1]:

$$C(a, b) = \frac{1}{\sqrt{a}} \sum_n s(nT) \psi\left(\frac{nT - b}{a}\right).$$

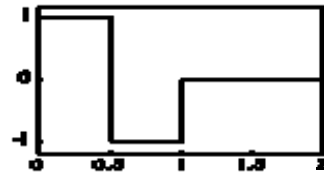
Природно припустити, що часовий зсув b також дискретизувати з кроком T . У даному випадку b буде містити стільки ж відліків, скільки містить цифровий сигнал $S(nT)$. Дискретний діапазон зміни параметру масштабування a задається самим дослідником. Таким чином, безперервна функція $C(a, b)$ перетворюється в матрицю коефіцієнтів вейвлет-розкладання розміром $l_a \times l_b$ (l - кількість відліків), що містить інформацію про особливості сигналу.

Всі вейвлети прийнято класифікувати за видом і особливостями утворюючої функції $\psi_0(t)$ та за іменем вченого, що вперше запропонував дану структуру.

Приклади часто використовуваних вейвлетів:

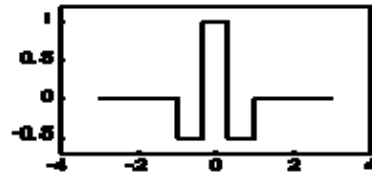
HAAR – вейвлет:

$$\psi(t) = \begin{cases} 1, & 0 \leq t < 1/2 \\ -1, & 1/2 \leq t < 1 \\ 0 & t < 0, t \geq 1 \end{cases}$$



FHAT - вейвлет («Французький капелюх» - French hat):

$$\psi(t) = \begin{cases} 1, & |t| \leq 1/3 \\ -1/2, & 1/3 < |t| \leq 1 \\ 0 & |t| > 1 \end{cases}$$



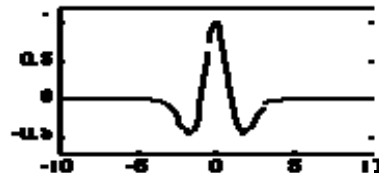
Wave – вейвлет:

$$\psi(t) = t \exp\left(-\frac{t^2}{2}\right)$$



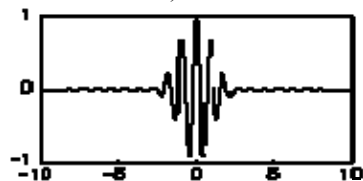
MHAT - вейвлет ("Мексиканський капелюх" - Mexican hat):

$$\psi(t) = (1 - t^2) \exp\left(-\frac{t^2}{2}\right)$$



Вейвлет Морле (утворює комплексний базис):

$$\psi(r) = \exp\left(ik_0 r - \frac{r^2}{2}\right)$$



Зручним інструментом для первинного аналізу та подальшої обробки звукових повідомлень є середовище MATLAB. Для прикладу можливості опрацювання існуючих даних, проведемо вейвлет-перетворення мовного повідомлення записаного у студії звукозапису і оцифрованого у 16-бітний сигнал із частотою дискретизації 44100Гц (рис.1).

Для цього використаємо середовище MATLAB:

```
>> [y,Fs,bits]=wavread('sample1',1) % Визначається частота дискретизації та кількість біт на відлік
y = -0.0064
Fs = 44100
bits = 16

>> siz=wavread('sample1', 'size') % Визначається кількість відліків у файлі та кількість каналів запису
siz = 80475 1

>>y=wavread('sample1'); % Будується часовий графік сигналу
>>plot(y); % у каналі
```

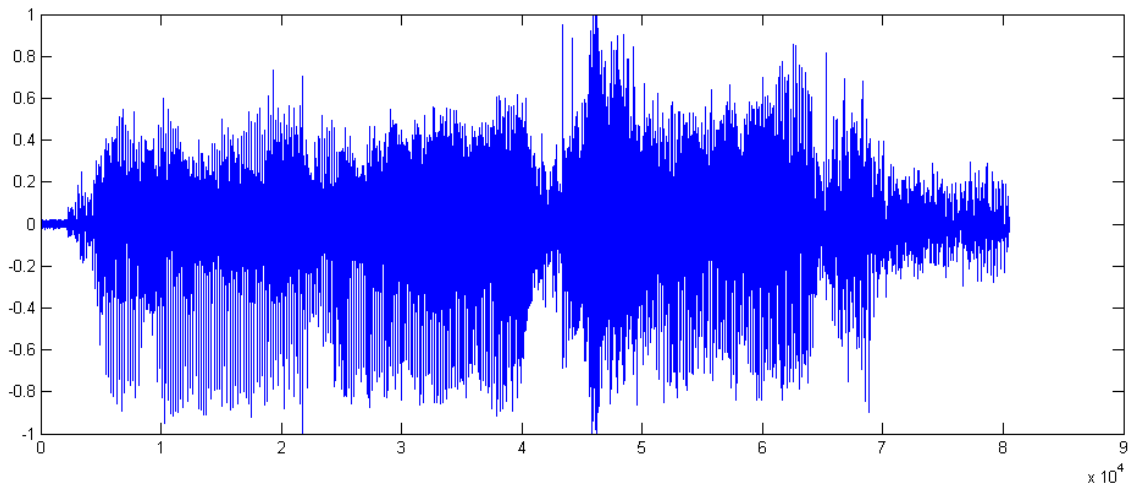


Рис. 1. Амплітудно-часова характеристика початкового сигналу

MATLAB здійснює роботу із вейвлетами у рамках пакета розширення **Wavelet Toolbox**, яке має графічний інтерфейс користувача (GUI).

Виклик інтерфейсу відбувається за допомогою команди:

```
>> wavemenu
```

Для проведення *одновимірного неперервного вейвлет-перетворення (Continuous Wavelet 1-D)* був обраний період дискретизації (**Sampling**): $T = \frac{1}{Fs} \approx 0.00002267573697$ сек; базисною функцією обрано *вейвлет Морле (morl)*; дискретний діапазон зміни параметру масштабування (**Scale Settings**) a був обраний за схемою «крок за кроком» (**Step by Step Mode**): min=1, max=64, step=1.

Результатом виконання буде зображення сигналу, що аналізується, спектрограми отриманої із матриці коефіцієнтів $C(a,b)$, самих коефіцієнтів та зображення ліній локальних максимумів:

Для більш детального перегляду спектрограми можливе її збільшення та зміна кольорового режиму за бажанням дослідника. Також можна забирати вікна коефіцієнтів та локальних максимумів, що призводить до зайняття спектрограмою їхнього місця у робочому просторі і як наслідок її збільшення.

Приведений механізм дослідження звукових повідомлень застосуємо для аналізу роботи алгоритмів кодування-декодування із втратами, що характерні для застосування у телекомунікаційних мережах для передачі аудіо повідомлень. Алгоритми кодування-декодування із втратами дозволяють суттєво зменшити бітовий потік даних, що передається у мережі, у порівнянні із аналогічним у так званому "чистому" або не стисненому форматі (RAW). Ця різниця досягає від 1:1 до 1:64 за бажанням оператора і може бути як постійна, так і змінна у часі.

Застосування алгоритмів із втратами обумовлює втрату частини інформації, а отже постає питання – яка саме частина інформації є втрачена, і чи можливо відновити цю інформацію на виході декодера, якщо відомий алгоритм та його "особливості".

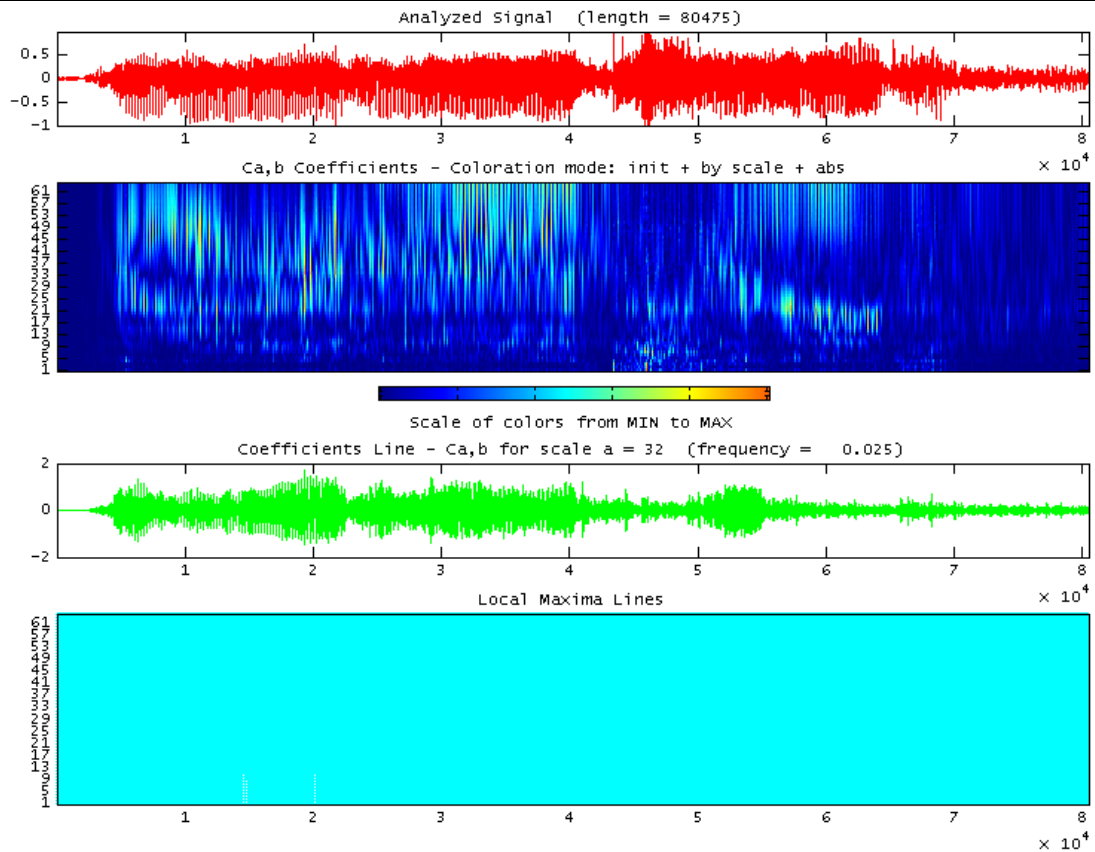


Рис. 2. Безперервне вейвлет-перетворення початкового сигналу

Для прикладу дослідження, початковий сигнал у форматі 44100 Гц, 16-біт, моно, без стиснення закодуємо за допомогою кодека MPEG Layer-3 у аудіо формат із потоковою швидкістю 64 КБ/сек, 44100 Гц, 16-біт, моно і декодуємо назад у WAV (44100 Гц, 16-біт, моно, без стиснення). Отриманий сигнал прийме вигляд:

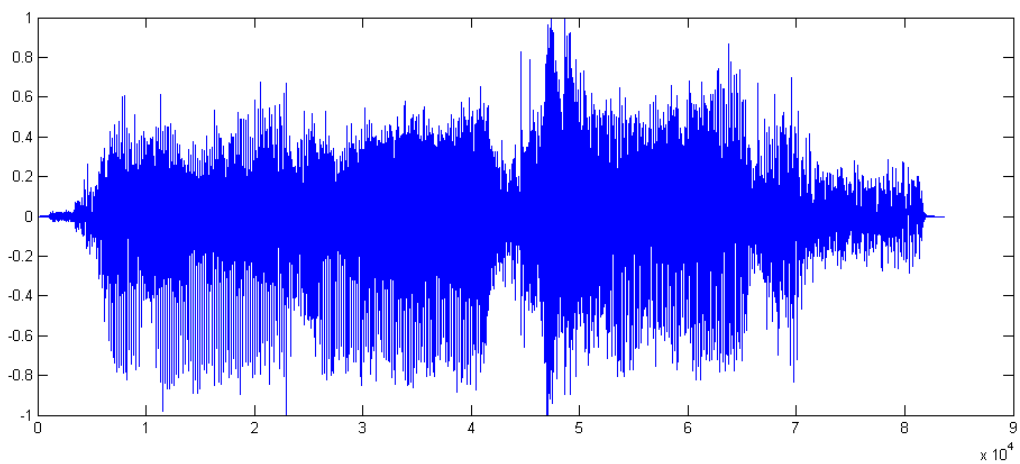


Рис. 3. Амплітудно-часова характеристика сигналу після кодування-декодування MPEG Layer-3

Проведемо спектральний аналіз за допомогою вейвлет-перетворення із такими ж параметрами як і для початкового сигналу.

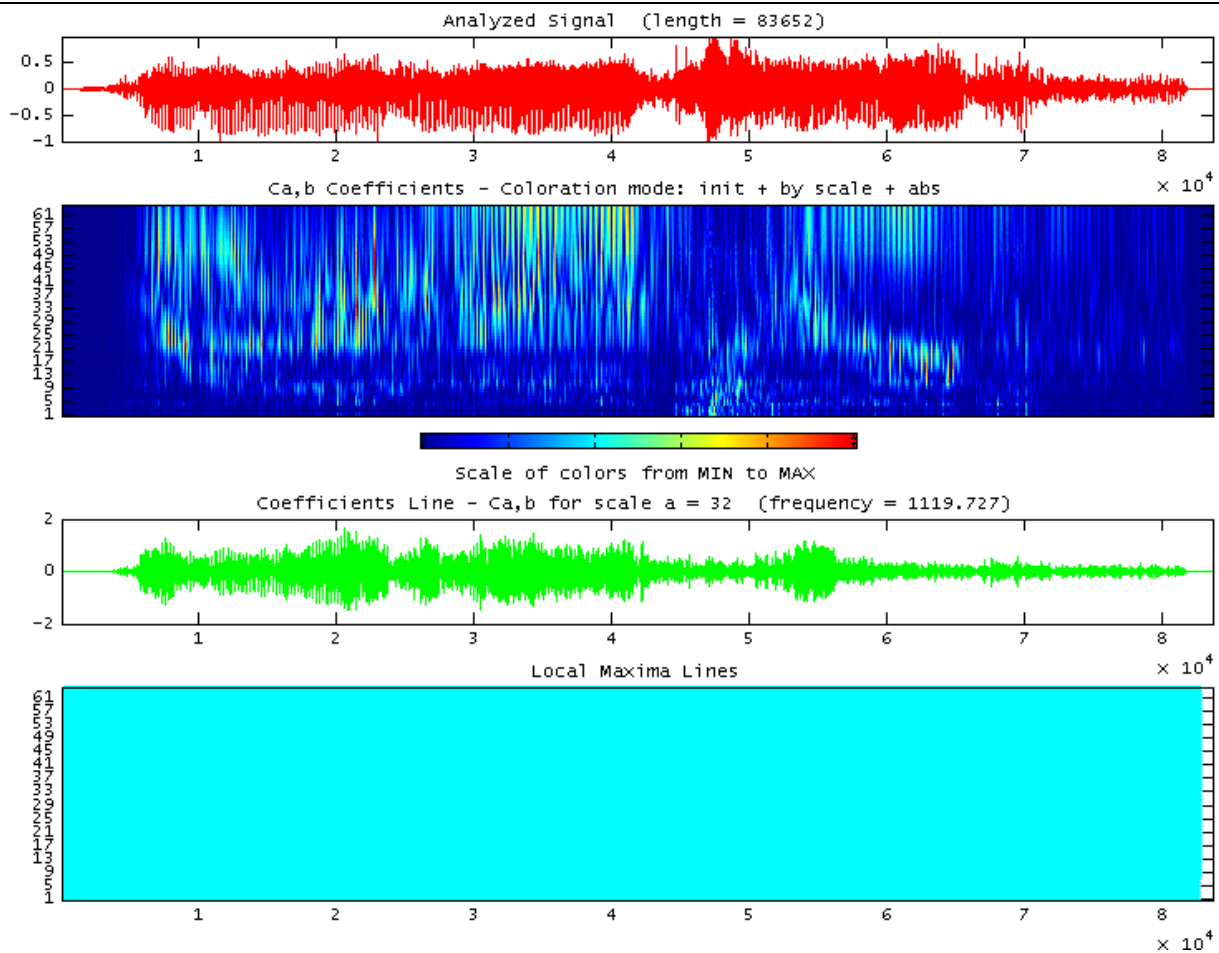


Рис. 4. Безперервне вейвлет-перетворення сигналу після кодування-декодування MPEG Layer-3

Деякі зміни у спектральній картині, а особливо у лініях, що зображують локальні максимуми, помітні відразу без використання додаткових обрахунків.

Канали зв'язку та інші реальні фактори додають до результуючої функції випадковий шум (рис. 5). Додамо до нашого сигналу випадковий шум (рис. 5).

```
>>noise=randn (1,size(y,2));           % Генерується випадкове значення одновимірного шуму
>>ey=y+noise;                          % Додається шум до початкового сигналу
>>plot(ey);                             % Будується часовий графік отриманого сигналу
```

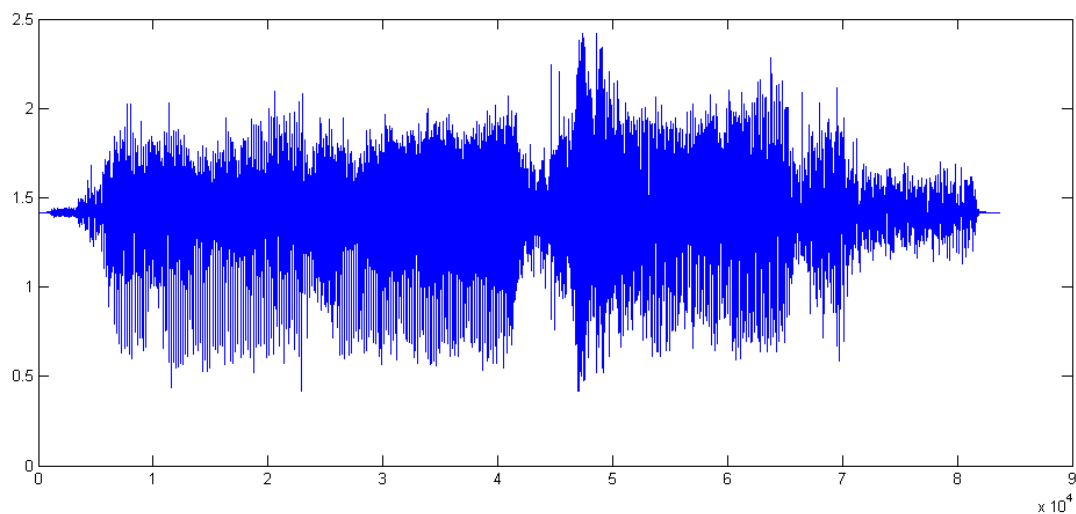


Рис. 5. Амплітудно-часова характеристика сигналу після кодування-декодування MPEG Layer-3 із випадковим шумом

Видалення шуму зі стаціонарного сигналу. Статистичні властивості стаціонарних та нестаціонарних сигналів значно відрізняються. Для їх моделювання та аналізу використовують різноманітні методи. В середовищі MATLAB у середині пакета роботи із вейвлетами **Wavelet Toolbox** існує графічне середовище, що реалізує можливість видалення шуму із стаціонарного одновимірного сигналу. Щоб його

викликати потрібно натиснути у головному меню кнопку **SWT De-noising 1-D**.

Для нашого сигналу обираємо базисом симплет (sym8), використовуючи 2 рівня розкладу. Далі обираємо стратегію порогової обробки **Fixed form threshold** та структуру шуму **Scaled white noise**. Натиснення кнопки **De-noise** дозволяє отримати сигнал очищений від шуму. Він зображається на одному графіку разом із початковим (рис 6).

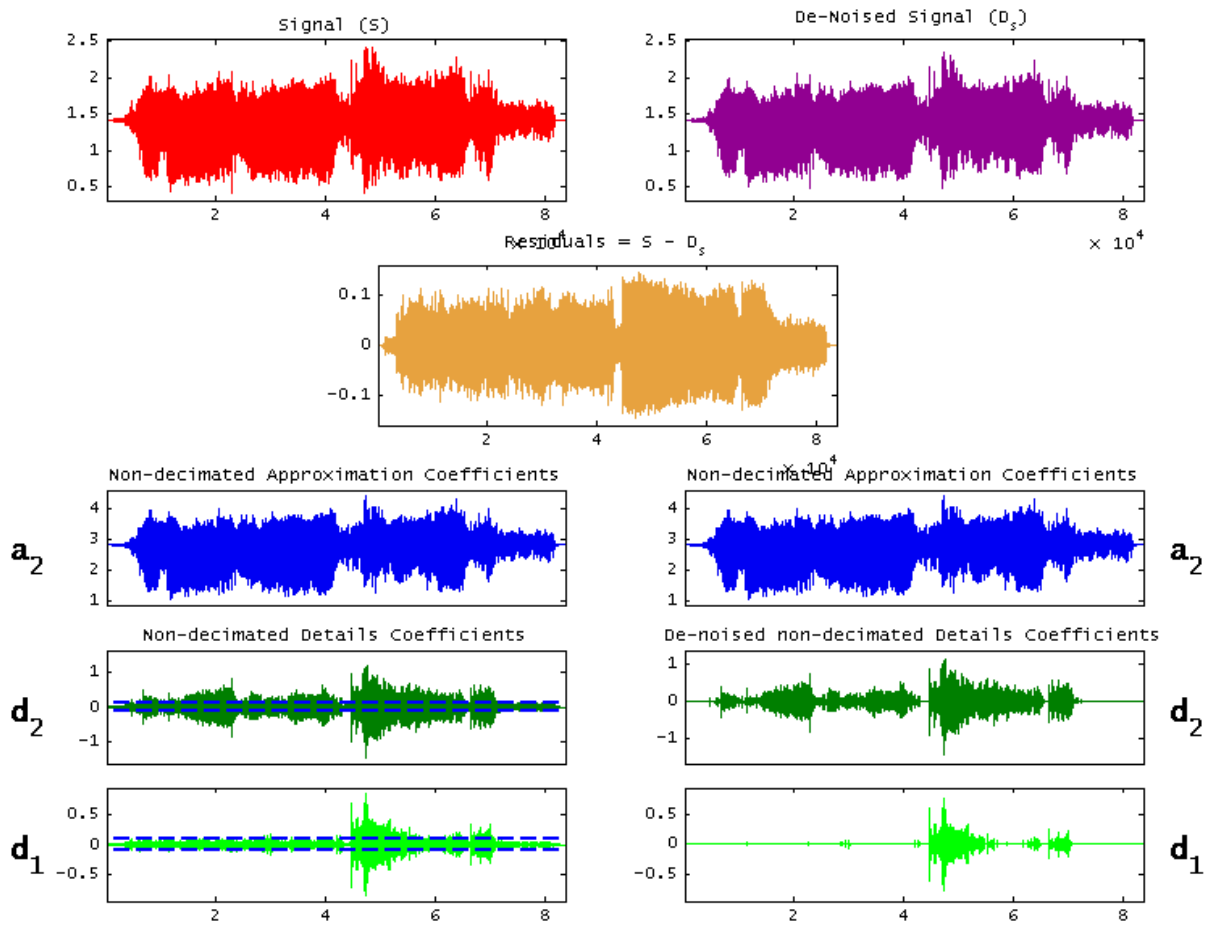


Рис. 6. Очищення сигналу від шуму

Очищення від шуму можна проводити або вручну окремо для кожного рівня, використовуючи повзунки, що розміщені з правого боку, або лівою кнопкою миші, переміщуючи курсором горизонтальні жовті пунктирні лінії порогів у лівих вікнах із зображення вейвлет-коефіцієнтів. Кнопка **Residuals** дає можливість оцінки характеристик різниці між початковим та очищеним сигналами (рис 7).

Наступним етапом аналізу сигналів є **оцінка густини розподілу**. Методи вейвлет-аналізу дозволяють оцінити густину розподілу ймовірностей значень відліків цифрового сигналу. В основі алгоритму лежать наступні положення:

1. Відліки сигналу розглядаються як вибірка значень випадкової величини із густиною розподілу $p(x)$. Довжина вибірки дорівнює N . Сигнал спостерігається у наявності шуму. За вибірковими значеннями буде отримана оцінка густини $\hat{p}(x)$.

2. Вибирається число m рівних інтервалів розбиття області значень відліків сигналу. Мінімально допустима величина $m = 64$. За замовчення рахують $m = N/4$.

3. Підраховується кількість відліків, що попали в кожний i -й інтервал N_i ($i = 1, 2, \dots, m$). Кожне N_i рахується новим сигналом.

4. Проводиться розкладання кожного N_i вибраним вейвлетом до потрібного рівня.

5. Реалізується порогова обробка вейвлет-коефіцієнтів одним із методів, що використовуються у процедурах видалення шуму.

6. Відновлюється кожний сигнал \hat{N}_i та гістограмним способом будується оцінка густини розподілу $\hat{p}(x)$.

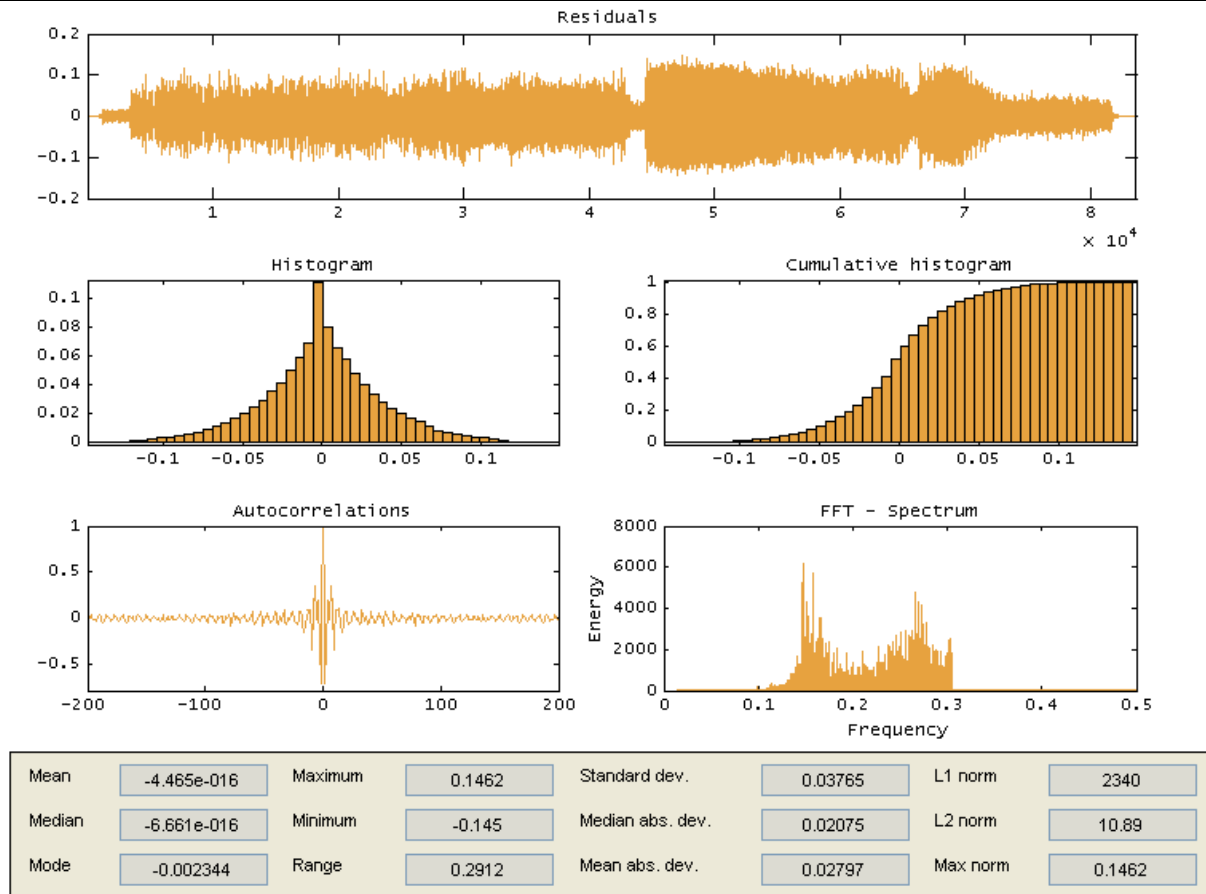


Рис. 7 Оцінка характеристик між початковим та очищеним сигналом

Застосуємо алгоритм до отриманого нами раніше сигналу із доданим випадковим шумом. Для цього потрібно у графічному інтерфейсі пакету роботи із вейвлет-перетвореннями обрати пункт **Density Estimation 1-D**. Загрузивши попередньо збережений нами сигнал у форматі ***.mat** оберемо вейвлет Дебози (**db3**) та 3 рівня розкладу. Далі обирається кількість **Nb bins** розбиття проміжку значень для побудови гістограми, вибираємо його значення – 256. Після цього при натисканні кнопки **Decompose** відбувається вейвлет-розкладання, результати якого відображаються в основному полі графічного вікна зліва (рис. 8).

Натискання кнопки **Estimate** дозволяє отримати оцінку густини розподілу. Ця густина складається з компонентів сигналу, отриманих окремо із оброблених коефіцієнтів на кожному рівні. Досліднику надається можливість експериментувати із різними стратегіями визначення порогового значення, перебираючи відповідні варіанти у списку **Select thresholding method** або безпосередньо переміщуючи пунктирні лінії порогів на графіках коефіцієнтів.

Оцінка регресії сигналу. Під регресією розуміється залежність середнього значення деякої випадкової величини від однієї чи декількох інших величин, які можуть бути детермінованими чи випадковими. На відміну від залежності $y = f(x)$, коли кожному значенню незалежної змінної x відповідає визначене значення y , при регресійному зв'язку одному значенню x можуть відповідати різноманітні значення y_1, y_2, \dots, y_k , що залежать від випадкових факторів. Таким чином, *рівняння регресії* $Y = F(X)$ визначається наступним чином: якщо величина X приймає конкретне значення x_0 , то $F(X)$ являє собою математичне очікування випадкової величини Y .

Реалізована в MATLAB оцінка регресії шукає залежність $Y = F(X)$ із незаданим раніше типом функції. Тому йде мова про видалення шуму e в моделі $Y = F(X) + e$, що робить використання вейвлетів ефективним.

В залежності від характеру змінної X розрізняють дві схеми регресії:

1. Постійна схема регресії (Fixed Design Regression), коли значення X детерміновані.
2. Стохастична схема регресії (Stochastic Design Regression), коли значення X з'являються випадково. Цей варіант є більш загальним.

Алгоритм оцінки функції регресії F на основі вейвлетів співпадає з викладеним вище алгоритмом оцінки густини розподілу ймовірностей.

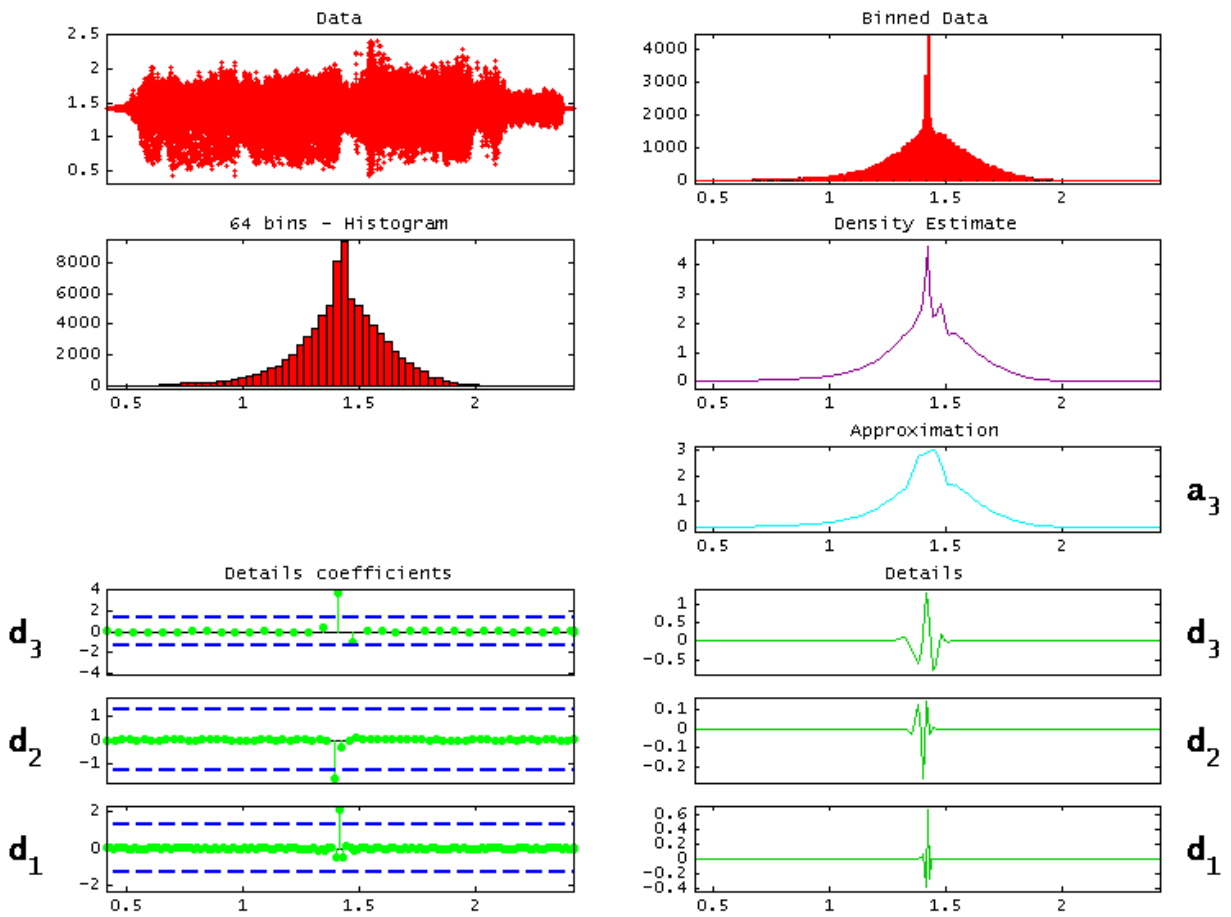


Рис. 8. Оцінка густини сигналу

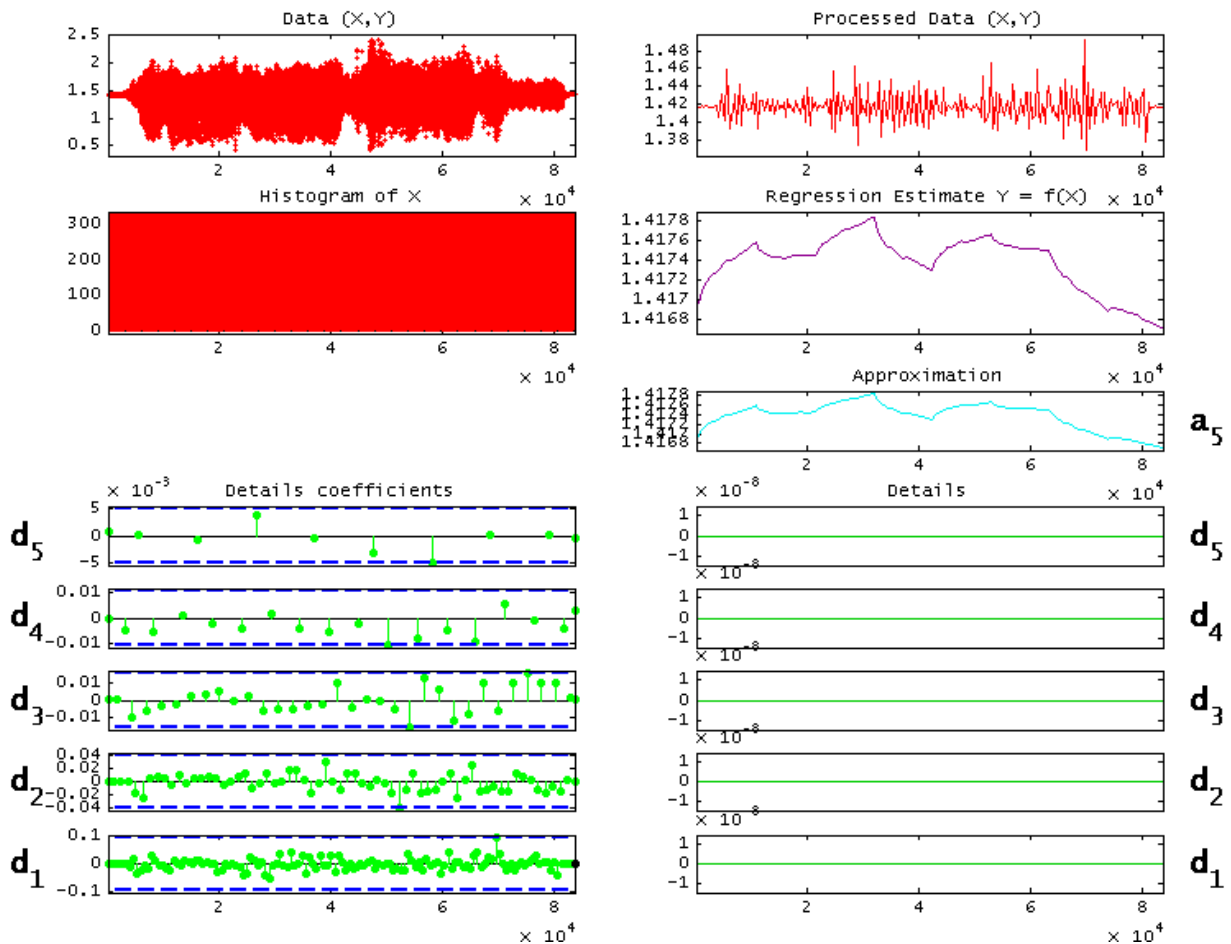


Рис. 9. Оцінка регресії сигналу

Щоб застосувати до нашого сигналу потрібно у головному меню пакету вибрати **Regression Estimation 1-D** та обрати тип даних із запропонованих (Data for Fixed Design Regression, Data for Stochastic Design Regression). У нашому випадку оберемо другий варіант та загрузимо потрібний сигнал. Було обрано вейвлет Дебози (db2) та 5 рівнів розкладу. Значення **Nb bins** залишимо, як у попередньому варіанті (256). Натиснувши **Decompose** проводимо вейвлет-розкладання у 5 рівнів, результати якого видні на рис. 9 зліва внизу. Після чого дозволяється стандартним шляхом провести видалення шумів. Вибравши усі параметри натискається кнопка **Estimate** та з'являється зображення апроксимуючої кривої разом із коефіцієнтами розкладання.

Після видалення шуму для дослідження його статистичних властивостей на кожному рівні розкладу можливе викликання додаткового вікна, натисканням на кнопку **Residuals** (рис 10).

Висновок. Наведений вище аналіз показав, що представлення мовних повідомлень у вигляді вейвлет-базисних функцій буде мати ряд переваг, серед яких: локалізація у часовій та частотній області, можливість масштабного перетворення і зсувів, розроблений математичний апарат для локалізації та класифікації особливих точок сигналу, відмінне відображення динаміки зміни сигналу уздовж масштабування. А середовище моделювання MATLAB наділене потужним математичним апаратом, який у поєднанні із зручністю у використанні, гнучким налаштуванням та простотою у вивченні робить його незамінним у аналізі та обробці речових повідомлень в трактах звукового мовлення.

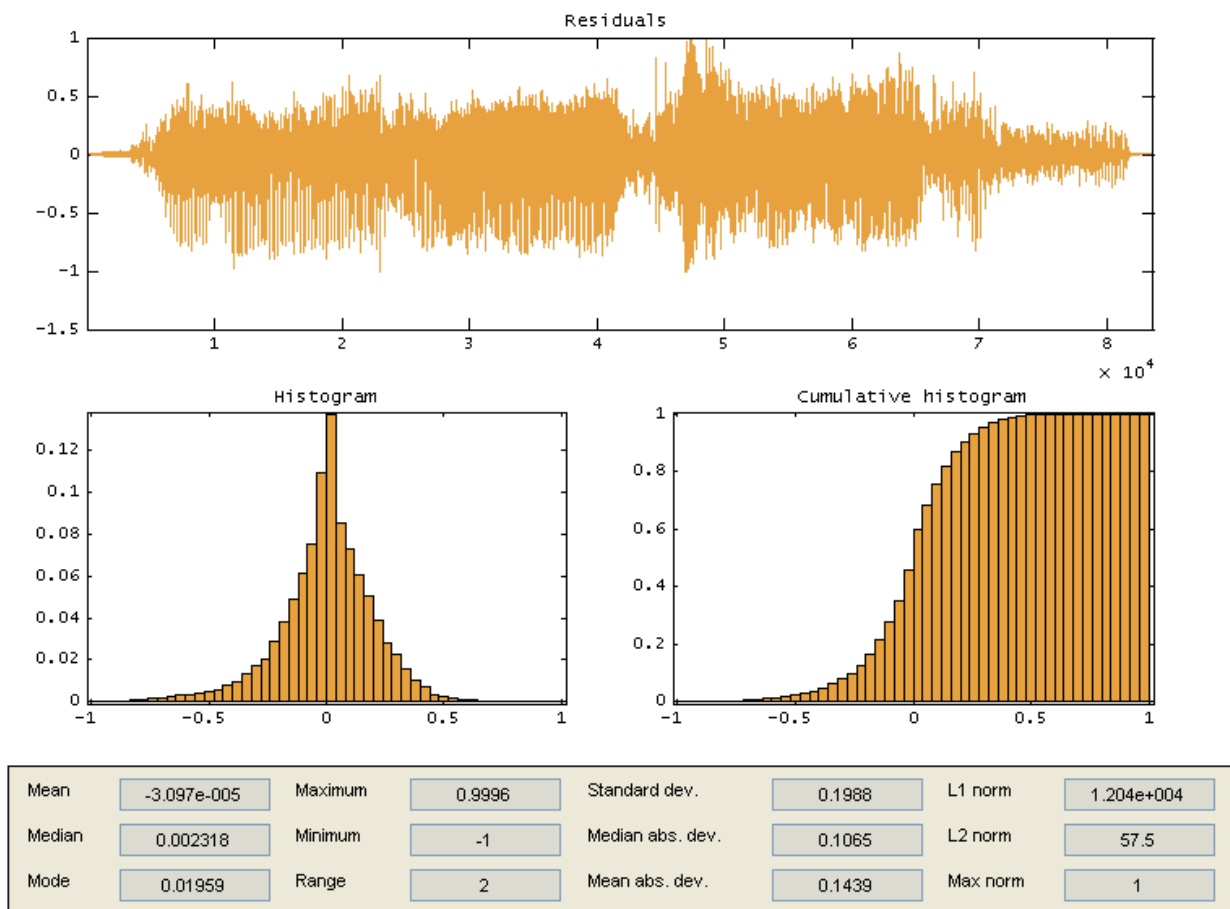


Рис. 10. Дослідження статистичних властивостей сигналів окремого рівні

Література

1. Солонина А.И. Цифровая обработка сигналов. Моделирование в MATLAB : учебное пособие / Солонина А.И. – СПб. : БХВ-Петербург, 2008 – 816с.: ил.
2. Попов О.Б., Рихтер С.Г. Цифровая обработка сигналов в трактах звукового вещания : Учебное пособие для вузов / О.Б. Попов, С.Г. Рихтер. – М.: Горячая линия – Телеком, 2007. – 314с.: ил.
3. MATLAB Lecture 7. Signal Processing in MATLAB [электронный ресурс] . – Режим доступа до документа : <http://www.math.uic.edu/~jan/mcs320s07/matlec7.pdf>
4. Davis Yen Pan. Digital audio compression / Digital Technical Journal. – Vol. 5. – No. 2, Spring 1993. – p. 1-13.
5. K. Brandenburg and G. Stoll. The ISO/MPEG Audio Codec: A Generic Standard for Coding of High Quality Digital Audio / Preprint 3336, 92nd Audio Engineering Society Convention, Vienna. – 1992.
6. J. D. Johnston. Transform Coding of Audio Signals Using Perceptual Noise Criteria / IEEE Journal on

Selected Areas in Communications. – vol. 6. – February 1988. – p. 314-323.

7. J. Ward and B. Stanier, "Fast Discrete Cosine Transform Algorithm for Systolic Arrays / Electronics Letters. – vol. 19. – No. 2. – January 1983.

Надійшла 5.4.2011 р.