October 10, 2006.

3. Аксельрод Б.М. Проблемно-орієнтована пошук по дії з використанням патентних баз даних: новий пошуково-вирішальної інструмент / Б.М. Аксельрод // МА ТРИЗ, Праці міжнародної конференції "Три покоління ТРИЗ". – Санкт-Петербург, 13-18 жовтня 2006 р.
4. Колчанов С.А. Концепція автоматизованої системи функціонально-орієнтованого пошуку / С.А. Колчанов, М.С. Рубін, Є.Л. Соколов // МА ТРИЗ, Праці міжнародної конференції "ТРИЗ Фест - 2007". – Москва, 9-10 липня 2007р.
5. Андрейчико А.В. Интеллектуальные информационные ситсемы / А.В. Андрейчиков, О.Н. Андрейчикова. – М.: Финансы и статистика, 2004. – 424с.
6. Ащинянц Р.А. Логические методы в искусственном интеллекте / Р.А. Ашинянц. – М.: МГАГИ, 1996 – 201с.
7. Ащинянц Р.А. Логические методы в автоматизации обучения / Р.А. Ашинянц. – М.: МГАГИ, 1996 – 201с.
8. Бондарев В.Н. Искусственный интеллект. Учебное пособие для вузов / В.Н. Бондарев, Ф.Г. Аде. – Севастополь: СевНТУ, 2002. – 615с.
9. Юхимчук С. В. Математичні моделі ризику для систем підтримки прийняття рішень / С.В.Юхимчук, А.О. Азарова. – Монографія, 2003. - 188 с.

Надійшла 26.1.2011 р.

УДК 621

А.В. КУДРЯШОВ, К.Л. ГОРЯЩЕНКО
Хмельницький національний університет

СПЕКТРАЛЬНИЙ АНАЛІЗ МОВНИХ СИГНАЛІВ В СЕРЕДОВИЩІ MATLAB

Передача мовних сигналів є задачею передачі звукових фонем. Кожна з фонем є сукупністю ряду звукових сигналів різної частоти та амплітуди. Спотворення фонем призводить до спотворення прийнятого мовного повідомлення. В статті описано принципи аналізу мовних сигналів.

A transmission of linguistic signals is the task of transmission of voice phonemes. Each of phonemes is the aggregate of row of voice signals of different frequency and amplitude. Distortion of phonemes results in distortion of the accepted linguistic report. Principles of analysis of linguistic signals are described in the article.

Ключові слова: фонема, спектральний аналіз.

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 [2, 4, 9]:

- 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 особливостям, 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 [4].

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 самоналаштовується, 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 морфеми, слова, їхні форми в мовному потоці. "Фонема" є поняттям близьким до поняття "звук", однак має суттєву відмінність. Зміна у звукових складових фонеми веде до утворення іншої фонеми, 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:

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

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]:

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

де: $\varphi_l(t)$ - масштабуючі або скейлинг-фукнції (phi-функції) з одиничним значення 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-фунції $\varphi_l(t)$ 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 переконатись, 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 $t \to nT$, отримаємо дискретний аналог формули [1]:

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

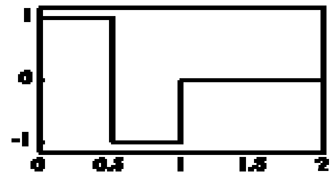
Природно припустити, що часовий зсув *b* також дискредитувати з кроком *T*. У даному випадку *b* буде містити стільки ж відліків, скільки містить цифровий сигнал $S(nT)$. Дискретний діапазон зміни параметру масштабування a задається самим дослідником. Таким чином, безперервна функція $C(a,b)$ перетворюється 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 [9]:

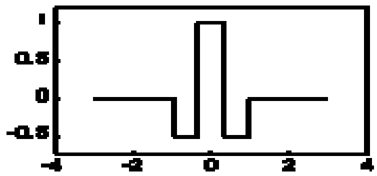
HAAR – вейвлет:

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



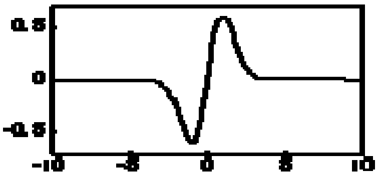
FHAT - вейвлет («Французький капелюх» - French hat):

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



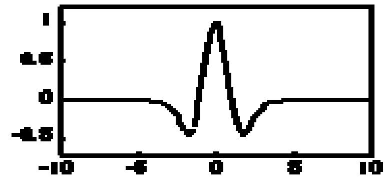
Wave – вейвлет:

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



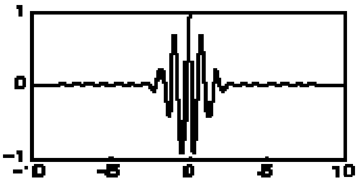
MHAT - вейвлет ("Мексиканський капелюх" - Mexican hat):

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



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 дискредитації 44100Гц (рис.1).

Для цього використаємо середовище MATLAB:

```
>> [y,Fs,bits]=wavread('sample1',1)  % Визначається частота дискредитації 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
```

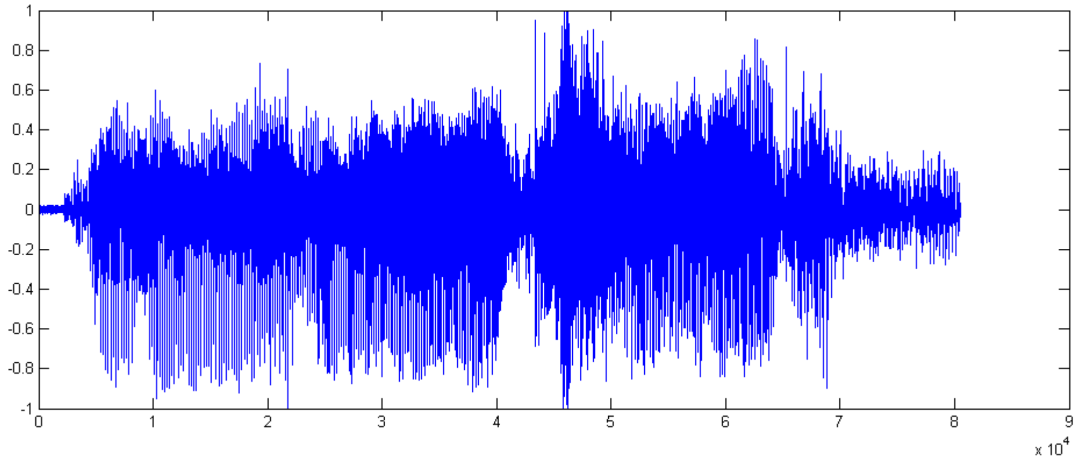


Рис. 1. Амплітудно-часова характеристика початкового сигналу

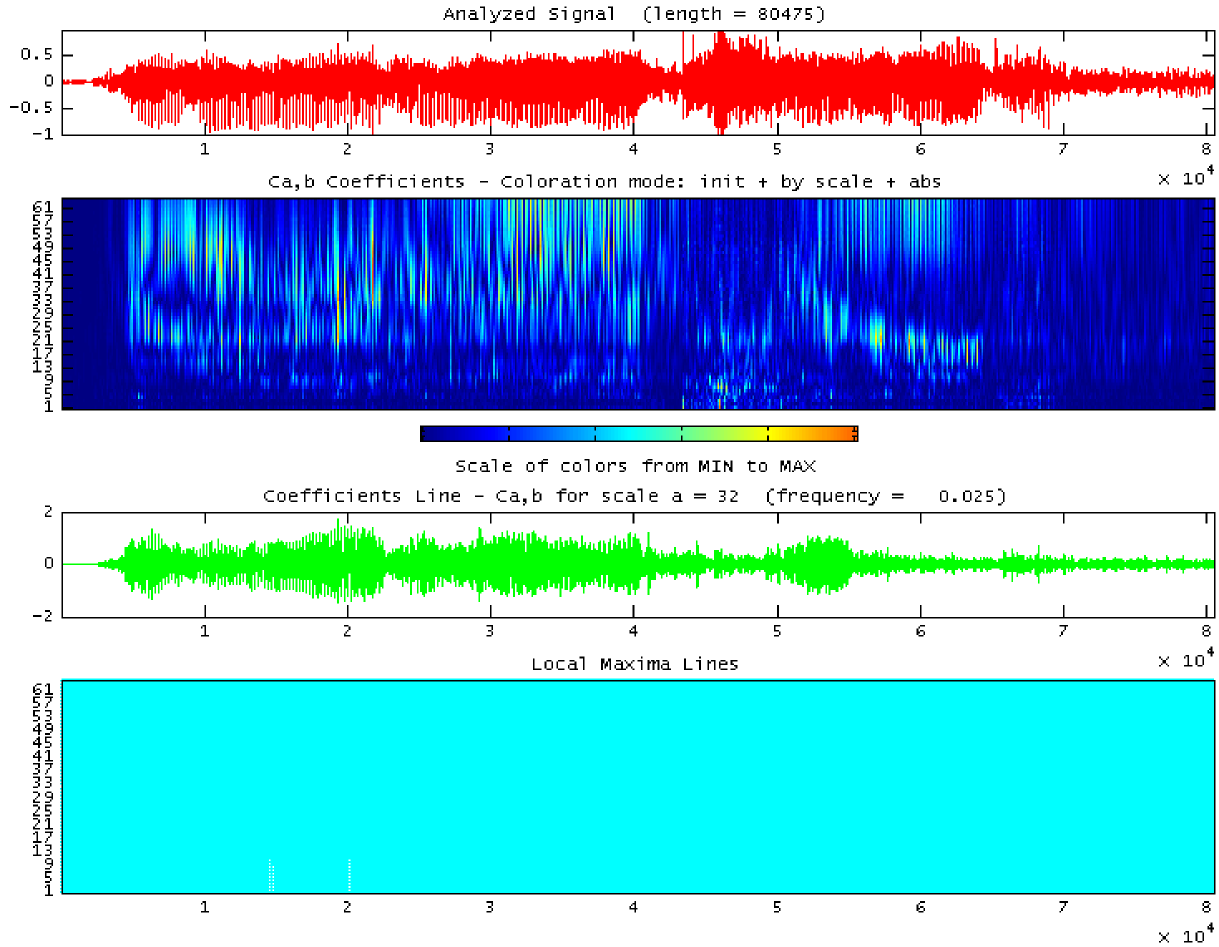


Рис. 2. Безперервне вейвлет-перетворення початкового сигналу

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 дискредитації(**Sampling**): $T = \frac{1}{Fs} \approx 0.00002267573697\ \text{сек}$; базисною функцією обрано *вейвлет Морле* (**morl**); дискретний діапазон зміни параметру масштабування (**Scale Setings**) 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.

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 вивчені робить його незамінним у аналізі та обробці речових повідомлень в трактах звукового мовлення.

Література

1. Солонина А.И. Цифровая обработка сигналов. Моделирование в MATLAB : учебное пособие / Солонина А.И. – СПб. : БХВ-Петербугр, 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 [електронний ресурс] . – Режим доступу до документу : http://www.math.uic.edu/~jan/mcs320s07/matlec7.pdf

4. Отнес Р., Эноксон Л. Прикладной анализ временных рядов : основные методы / Р. Отнес, Л. Эноксон / [пер. с англ.]. – М. : изд-во "Мир". – 1982. – 430 с.

5. Davis Yen Pan. Digital audio compression / Digital Technical Journal. – Vol. 5. – No. 2, Spring 1993. – p. 1-13.

6. K. Brandenburg and G. Stoll. The ISO/MPEGAudio Codec: A Generic Standard for Coding of High Quality Digital Audio / Preprint 3336, 92nd Audio Engineering Society Convention, Vienna. – 1992.

7. 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.

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

9. Amara Graps. An introduction to Wavelets / IEEE Computational Science and Engineering, Summer 1995, vol. 2, num. 2

Надійшла 16.1.2011 р.

УДК 004.89

Т.О. САВЧУК, А.В. САКАЛЮК
Вінницький національний технічний університет

ЗАСТОСУВАННЯ КЛАСТЕРНОГО АНАЛІЗУ ДЛЯ КОЛАБОРАТИВНОЇ ФІЛЬТРАЦІЇ

В статті описано основні проблеми, які виникають при розробці рекомендаційних система на основі колаборативної фільтрації та запропоновано підходи до їх вирішення, а саме використання кластерного аналізу та згладжування за допомогою нейронної мережі на основі радіальної базисної функції.

In this paper is described the main problems that arise when developing a recommendation system based on collaborative filtering and approaches for their solution, namely the use of cluster analysis and smoothing using a neural network based on radial basis functions

Ключові слова: колаборативна фільтрація, кластеризація, радіальна базисна функція, функція активації, алгоритм згладжування, розрідженість, масштабованість, самоорганізаційна карта Кохонена.

Вступ

Світ сучасної економіки наповнений товарами, властивості яких важко оцінити завчасно, виходячи