

УДК 004.03

О.М. АМБРОС, В.М. БОНДАРЕНКО

Національний технічний університет України «КПІ»

ХВИЛЬОВЕ СТИСНЕННЯ В СИСТЕМАХ КОМП'ЮТЕРНОЇ ТЕЛЕФОНІЇ

Наведені результати досліджень стиснення мовного сигналу дискретним косинусним перетворенням, а також перетвореннями Уолша та Хаара. Для кожного з алгоритмів експериментальним шляхом встановлено, при якому максимальному ступені стиснення відновлена мова практично не відрізняється від вихідної та при якому максимальному ступені стиснення вона знаходиться на межі розбірливості

Ключові слова: сигнал, перетворення, стиснення, межа розбірливості

Бурхливий розвиток засобів обчислювальної техніки та зв'язку призвів до формування комп'ютерної телефонії (КТ). Системи КТ забезпечують передачу мовного сигналу по телефонній мережі, чи по IP-мережі. Сигнал по каналу зв'язку передається у цифровому вигляді і, як правило, перед передачею стискається з метою видалення надлишкових елементів [1].

До систем КТ відносяться: центри обслуговування викликів, телебанкінг, автоматичне сповіщення, телеголосування, голосова пошта та ін.

Однією з вимог до вищеперерахованих систем є здатність озвучувати для абонента те чи інше повідомлення. Тому для більшої гнучкості в системах КТ застосовується прямий синтез мовних повідомлень по тексту. Цей підхід дозволяє зекономити людські ресурси і значну частину роботи перекласти на комп'ютери. Враховуючи спектральний склад синтезованої мови, можливо здійснювати її стиснення, яке зменшує вимоги до пропускної здатності каналу зв'язку.

Постановка завдання

Роботу присвячено вирішенню питання використання хвильових алгоритмів для стиснення мовних сигналів. Окрім простоти реалізації, основними вимогами до досліджуваних алгоритмів є: добра розбірливість мови після стиснення, максимальний коефіцієнт стиснення.

Мета роботи

Мета роботи полягає у визначенні серед досліджуваних алгоритмів такого, який би мав максимальний коефіцієнт стиснення за умови що мовний сигнал є розбірливим.

Об'єкти та методи дослідження

Розглянемо існуючі види стиснення та у відповідності до вищеперерахованих вимог порівняємо між собою перетворення, що застосовуються для хвильового стиснення.

Основні види стиснення

Стиснення чи компресія даних – це алгоритмічне перетворення, яке здійснюється з метою зменшення об'єму даних. Принцип стиснення даних доволі простий.

Практично усі форми подання даних (текст, числа, зображення, відео) містять надлишкові елементи.

Отже, дані можуть бути стиснуті з однією з двох цілей: для економії місця при зберіганні чи для зниження потреб в пропускній здатності.

Існує два основні різновиди стиснення даних: *стиснення без втрат* та *стиснення з втратами*. При *стисненні без втрат* інформація не втрачається і розпаковані дані ідентичні вихідним нестисненим даним. У випадку *стиснення з втратами* розпаковані дані можуть бути прийнятним наближенням (у відповідності до прийнятого критерію точності) до вихідних нестиснених даних.

Стиснення даних без втрат дозволяє відновити вихідні нестиснені дані з точністю до біту. Це, як правило, необхідно для текстових файлів, баз даних, двійкових об'єктних файлів і т. п.

Однак існує багато типів файлів, для яких непотрібне абсолютно точне відновлення вихідних даних. Прикладами таких файлів є: аудіо- і відеокліпи, а також нерухомі зображення. В цих випадках допускається деяка кількість помилок при відновленні вихідних даних, і може бути застосоване *стиснення з втратами*.

Важливим фактором при розробці та виборі алгоритмів стиснення даних є об'єм робіт, необхідний для стиснення та розпаковки (декомпресії) даних. В загальному випадку існує компроміс між досяжним ступенем стиснення і швидкістю роботи алгоритму стиснення, а також вартістю цієї роботи [2].

Алгоритми хвильового стиснення

В останні роки збільшився інтерес до методів хвильового стиснення з втратами в високошвидкісних комп'ютерних мережах [2].

Класичним підходом є застосування дискретного косинусного перетворення.

Загальна формула для одновимірного дискретного косинусного перетворення наведена нижче:

$$S(f) = \sqrt{\frac{2}{N}} C(f) \sum_{x=0}^{N-1} p(x) \cos \left[\frac{(2x+1)\pi f}{2N} \right], \quad (1)$$

де N – кількість відліків у вихідному сигналі; $p(x)$ – значення амплітуди сигналу в залежності від відліку; f – частота, $0 \leq f < N$; $C(f)$ – коефіцієнт, що визначається наступним чином:

$$C(f) = \begin{cases} 1/\sqrt{2}, & f = 0 \\ 1, & f > 0 \end{cases} \quad (2)$$

Аналогічним підходом до хвильового стиснення даних є застосування перетворень Уолша (3) та Хаара (6).

$$W_w(z) = (-1)^{\sum_{i=0}^{m-1} w^{(m-1-i)} z^{(i)}}, \quad (3)$$

де $w^{(i)}$ та $z^{(i)}$ визначаються з двійкового розкладу w і z

$$w = \sum_{i=0}^{m-1} w^{(i)} 2^{m-1-i}, \quad (4)$$

$$z = \sum_{i=0}^{m-1} z^{(i)} 2^{m-1-i}. \quad (5)$$

$$H_l^{(n)}(\theta) = \begin{cases} 2^{l/2}, & \text{при } \frac{n-1}{2^l} \leq \theta \leq \frac{n-1/2}{2^l}, \\ -2^{l/2}, & \text{при } \frac{n-1/2}{2^l} \leq \theta \leq \frac{n}{2^l}, \\ 0, & \text{при інших } \theta \in [0,1), \end{cases} \quad (6)$$

де $0 \leq l \leq \log_2 N$, $1 \leq n \leq 2^l$.

Функції Уолша використовуються для аналізу динаміки лінійних і нелінійних систем, розробці систем оптимального керування, моделюванні процесів, ідентифікації об'єктів та для формування сигналів, які передаються по лініях радіозв'язку.

Також на базі їх використання розроблені удосконалені методи завадостійкого кодування сигналів. Окрім цього, функції Уолша застосовуються у системах з множинним кодовим розділенням каналів (Code Division Multiple Access – CDMA) в стільниковому зв'язку [3].

Функції Хаара також знаходять застосування при розробці та дослідженні засобів автоматики, при дослідженні випадкових процесів, а також при розробці конкретних типів пристроїв в системах керування та зв'язку [4].

Типове застосування двовимірні алгоритми хвильового стиснення знаходять в області обробки зображень. Однак в даній статті розглядається використання одновимірних хвильових алгоритмів для стиснення мови з метою подальшої передачі сигналу через IP-мережу.

Результати та їх обговорення

Для дослідження перетворень Уолша та Хаара, а також дискретного косинусного перетворення в програмі для синтезу мови за текстом SAPI5 TTSAPP були згенеровані мовні сигнали в аудіо форматі «wav».

Дані сигнали імпортувались в середовище Matlab, де вони являли собою набір відліків. Відліки мовного сигналу, який представляв фразу «Добрий день», а також його спектр, отриманий за допомогою дискретного перетворення Фур'є (ДПФ), представлені на рис. 1.

Подальші дослідження проводились з використанням вищеперерахованих перетворень в два етапи: на першому – компресії та декомпресії піддавалася повна фраза, а на другому – фраза розбивалась на n блоків розміром по 1024 відліки, в результаті чого компресія та декомпресія проводилась над кожним з таких блоків.

При дослідженнях на обох етапах коефіцієнт стиснення встановлювався рівним: 2, 3, 4, 5, 6, 7, 8, 9 и 10. Таке збільшення коефіцієнта стиснення звичайно відображалось на розбірливості мови для кожного з перетворень по-своєму.

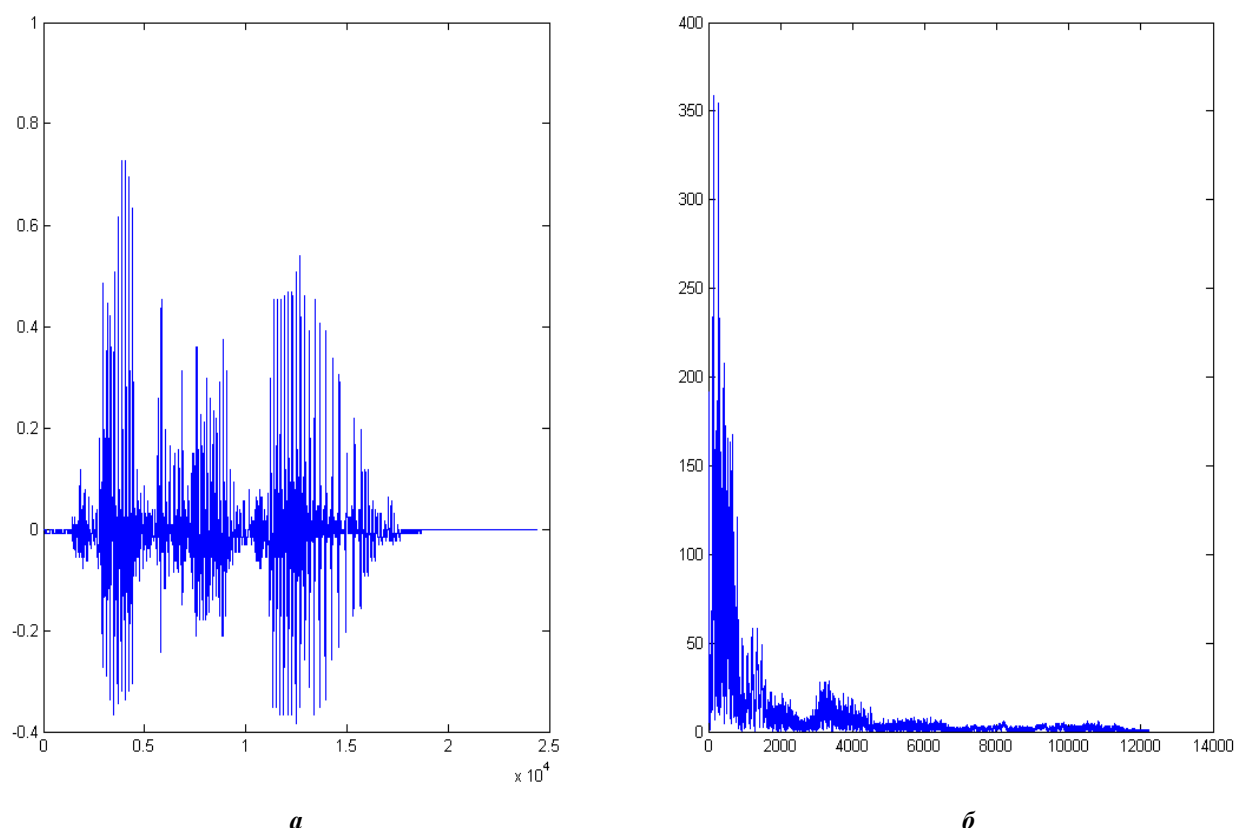


Рис. 1. Мовний сигнал: *a* – у часовій області, *b* – у частотній області

Результати досліджень по стисненню сигналу наведені в табл. 1.

На рис. 2 зображені спектри для згаданого раніше мовного сигналу після стиснення кожним з досліджуваних алгоритмів блоками по 1024 відліки при коефіцієнті стиснення рівному 8.

Спектри, отримані для мовного сигналу, стисненого за допомогою перетворень Уолша та Хаара, можна вважати схожими.

Таблиця 1. Стиснення мовного сигналу досліджуваними алгоритмами

	Коефіцієнт стиснення всієї фрази		Коефіцієнт стиснення фрази блоками	
	K_p	$K_{гр}$	K_p	$K_{гр}$
Дискретне косинусне перетворення	4	7	5	7
Перетворення Уолша	3	7	4	9
Перетворення Хаара	5	9	3	8

Позначення, прийняті у таблиці:

K_p – коефіцієнт стиснення, при якому мова є добре розбірливою;

$K_{гр}$ – коефіцієнт стиснення, при якому мова знаходиться на межі розбірливості.

Слід зазначити, що для кожного з алгоритмів, окрім різниці в значеннях максимальних коефіцієнтів стиснення, у відновленій після стиснення мові спостерігалась наявність побічних ефектів. Так, у випадку дискретного косинусного перетворення збільшення коефіцієнта стиснення призводило до збільшення вкладу фоновий шумової складової у сигналі.

Для перетворення Уолша збільшення коефіцієнта стиснення, при якому мова була ще добре розбірливою, призводило до збільшення шуму перетворення та з-за цього мова ставала все більш схожою на мову робота.

У випадку перетворення Хаара збільшення коефіцієнта стиснення на обох етапах досліджень призводило до збільшення ефекту «роботизованного» звучання мови. Однак, на другому етапі для деяких значень коефіцієнта стиснення до цього ефекту додавалась ще й «уривчастість» мови.

Для кожного з досліджуваних алгоритмів при досягненні коефіцієнтом стиснення граничної розбірливості мови в сигналі значно посилювався вклад шумової складової.

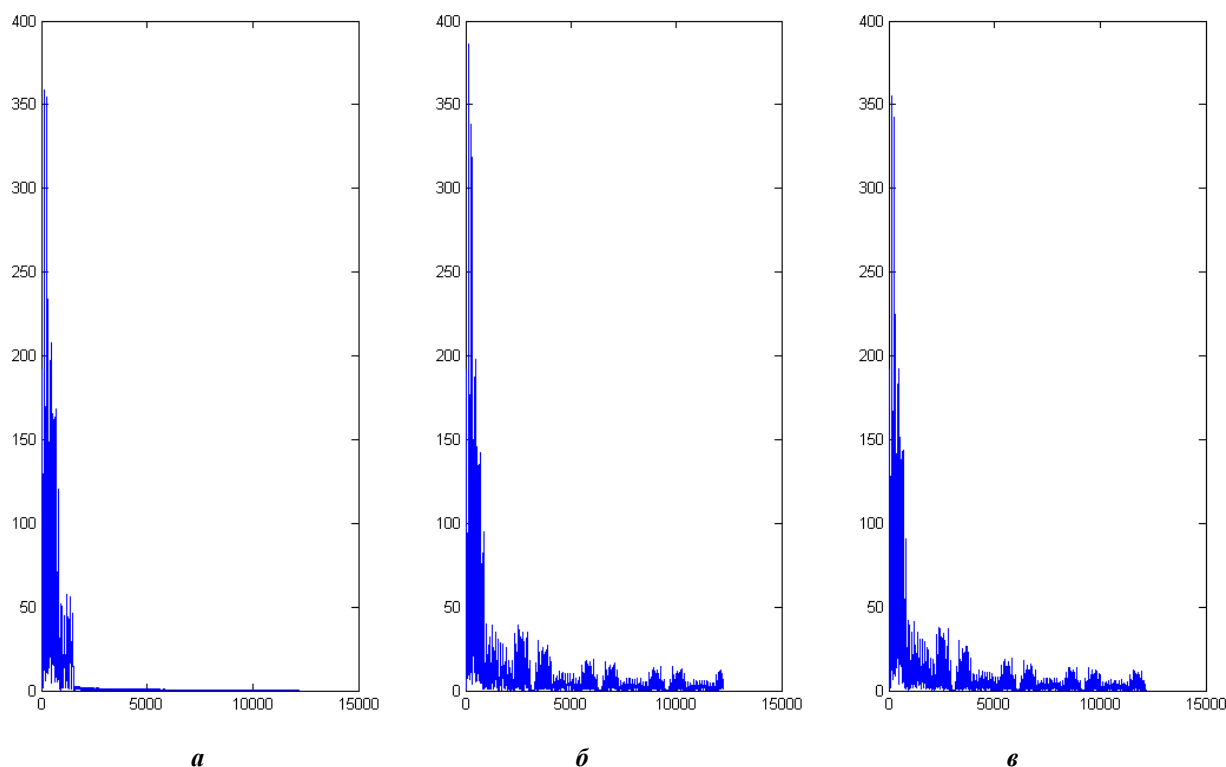


Рис. 2. Спектри ДПФ мовного сигналу при стисненні блоками за:

a – дискретним косинусним перетворенням; ***b*** – перетворенням Уолша; ***v*** – перетворенням Хаара

Висновки

Експериментальні дослідження по стисненню синтезованого мовного сигналу за допомогою різних перетворень привели до наступних результатів.

При ідентичній розбірливості мови та стисненні всієї фрази повністю, з трьох алгоритмів максимального значення коефіцієнта стиснення, рівного 9, досяг алгоритм Хаара.

Однак, слід зазначити, що при стисненні мовного сигналу блоками по 1024 відліки, було встановлено, що найбільш розбірливою при коефіцієнті стиснення, рівному 9, є мова, відновлена після стиснення перетворенням Уолша.

Оскільки на практиці у більшості випадків сигнали передаються блоками, то звідси випливає, що оптимальним для реального застосування є алгоритм Уолша, для якого, так як і для перетворення Хаара, існує ефективна реалізація у вигляді швидкого перетворення [5].

Список використаної літератури:

1. Тропченко А.Ю., Тропченко А.А. Методы сжатия изображений, аудиосигналов и видео: Учебное пособие – СПб.: СПбГУ ИТМО, – 2009. – 108 с.
2. Столлингс В. Современные компьютерные сети. 2-е изд. – СПб.: Питер, –2003. –783 с.
3. Никитин Г.И. Применение функций Уолша в сотовых системах связи с кодовым разделением каналов: Учебное пособие, – СПбГУАП, – СПб.: –2003. –86 с.
4. Гаспер Б.С., Липатов И.Н., ИВС и АСУТП: Учебное пособие, – Пермь, –1999. – 122 с.
5. Карповский М.Г., Москалев Э.С. Спектральные методы анализа и синтеза дискретных устройств. – Л.: Энергия, – 1973. – 144 с.

Стаття надійшла до редакції 12.03.2012

Волновое сжатие в системах компьютерной телефонии

Амброс А.Н., Бондаренко В.Н.

Национальный технический университет Украины «КПИ»

Приведены результаты исследований сжатия речевого сигнала дискретным косинусным преобразованием, а также преобразованиями Уолша и Хаара. Для каждого из алгоритмов экспериментальным путем установлено при какой максимальной степени сжатия восстановленная речь практически не отличается от исходной и при какой максимальной степени сжатия она находится на грани разборчивости.

Ключевые слова: сигнал, перетворення, стиснення, межа розбірливості.

Wavelet compression in computer telephony system

Ambros A.N., Bondarenko V.N.

National Technical University of Ukraine «KPI»

The results of the speech signal compression with discrete cosine transform, the Walsh and Haar transforms are presented. The maximum compression ratio at which the recovered speech is virtually identical to the original and at which it is on the verge of legibility is experimentally determined for each of the algorithms.

Keywords: signal, peretvorennya, stysnennya, meza rozbirlyvosti.