

МІНІСТЕРСТВО ОСВІТИ І НАУКИ, МОЛОДІ ТА СПОРТУ УКРАЇНИ
Вінницький національний технічний університет

ХРУЩАК СЕРГІЙ ВІКТОРОВИЧ

УДК 621.39

ВЕКТОРНЕ КВАНТУВАННЯ ПАРАМЕТРІВ СИГНАЛУ ПРИ УЩІЛЬНЕННІ
МОВЛЕННЯ В КОМП'ЮТЕРНИХ СИСТЕМАХ

Спеціальність 05.13.05 – комп'ютерні системи та компоненти

АВТОРЕФЕРАТ
дисертації на здобуття наукового ступеня
кандидата технічних наук

Вінниця – 2012

Дисертацією є рукопис.

Роботу виконано у Вінницькому національному технічному університеті Міністерства освіти і науки, молоді та спорту України.

Науковий керівник – кандидат технічних наук, доцент
Ткаченко Олександр Миколайович,
Вінницький національний технічний університет,
докторант кафедри обчислювальної техніки

Офіційні опоненти: доктор технічних наук, професор
Лужецький Володимир Андрійович,
Вінницький національний технічний університет,
завідувач кафедри захисту інформації

доктор технічних наук, професор
Саваневич Вадим Євгенійович,
Харківський національний університет
радіоелектроніки,
професор кафедри телекомунікаційних систем.

Захист відбудеться «19» травня 2012 р. о 9 годині 30 хвилин на засіданні спеціалізованої вченої ради Д 05.052.01 у Вінницькому національному технічному університеті за адресою: 21021, м. Вінниця, вул. Хмельницьке шосе, 95, ГНК, ауд. 210.

З дисертацією можна ознайомитися у бібліотеці Вінницького національного технічного університету за адресою: 21021, м. Вінниця, вул. Хмельницьке шосе, 95.

Автореферат розісланий «17» квітня 2012 р.

Учений секретар
спеціалізованої вченої ради

С. М. Захарченко

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми. Актуальним напрямком розвитку сучасних комп'ютерних систем є оброблення, зберігання та передавання мультимедійної інформації, зокрема мовленнєвих сигналів. При цьому одним з найбільш важливих питань є обсяг даних, необхідних для забезпечення якісного відтворення мовленнєвого сигналу. Ущільнення мовленнєвого сигналу зменшує обсяг даних, необхідних для його опису та відповідно апаратні витрати на його зберігання та передавання. Це дозволяє збільшувати кількість користувачів мовленнєвих сервісів мережі без збільшення апаратних витрат. Створення кодерів з високим ступенем ущільнення дозволить використовувати їх у побудові систем голосових повідомлень, автоматичних мовленнєвих довідкових служб, систем передавання мовленнєвих сигналів по низькошвидкісних комп'ютерних мережах, недорогих систем безпроводного зв'язку. Тому світовий ринок мультимедійних систем розвивається під безпосереднім технологічним впливом вчених та розробників мовленнєвих кодеків.

Слід відзначити, що основні зусилля в розробці методів і систем оброблення, передавання та зберігання цифрових мовленнєвих сигналів належать високорозвиненим країнам – США, Японії, Франції та Великобританії. Кінцеві результати досліджень формуються у вигляді національних стандартів і рекомендацій міжнародних організацій, таких як, ITU-T і групи MPEG. Значний внесок у розвиток систем ущільнення мовлення внесли Рабінер Л.Р., Атал Б.С., Палівал К.К., Ітакура Ф., Джін Ж., Макхоул Дж, Вінцюк Т.К.

Проте дослідження у цьому напрямку не втрачають своєї актуальності, оскільки залишилось чимало невирішених задач. Відсутні методи ущільнення мовленнєвих сигналів, які враховували б кореляцію між кадрами мовленнєвого сигналу, а також методи, які б використовували впорядковані кодові книги в процесі векторного квантування параметрів мовленнєвого сигналу. Таким чином актуальною є науково-технічна задача розробки програмно-апаратних засобів ущільнення даних, необхідних для подальшого передавання або відтворення мовленнєвого сигналу без втрати прийнятної якості декодованого звуку.

Зв'язок роботи з науковими програмами, планами, темами. Напрямок виконаних досліджень безпосередньо пов'язаний з науково-дослідним напрямком кафедри обчислювальних систем Вінницького національного технічного університету «Розробка методів і засобів цифрової обробки мовленнєвих сигналів».

Основні задачі дисертації відповідають державним науково-технічним програмам, що визначені Законами України «Про наукову і науково-технічну діяльність», «Про національну програму інформатизації» та планам науково-технічних робіт Міністерства освіти і науки, молоді та спорту України: 6.2.2 – Перспективні інформаційні технології і системи.

Мета і задачі дослідження. Метою дослідження є підвищення продуктивності оброблення мовленнєвих сигналів в комп'ютерній системі за рахунок розробки програмно-апаратних засобів ущільнення даних, необхідних для подальшого передавання, зберігання та відтворення мовленнєвого сигналу без втрати прийнятної якості декодованого мовлення. Для досягнення поставленої мети необхідно розв'язати такі задачі:

1. Провести аналіз існуючих методів та засобів ущільнення мовленнєвих сигналів.
2. Вдосконалити метод впорядкування векторів кодових книг та розробити алгоритм прискореного пошуку по структурованих кодових книгах.
3. Розробити метод ущільнення параметрів мовленнєвого сигналу, який би враховував кореляцію між сусідніми кадрами сигналу.
4. Вдосконалити адаптивний метод визначення голосової активності для роботи за умов низького співвідношення сигнал/шум.
5. Розробити алгоритмічне та програмне забезпечення для дослідження методів прискореного пошуку векторів у кодових книгах.
6. Розробити програмне забезпечення для ущільнення мовленнєвого сигналу на основі векторного квантування.

7. Розробити структуру комп'ютерної системи для передавання та зберігання мовленнєвих сигналів з використанням запропонованих методів.

8. Реалізувати вдосконалений метод визначення голосової активності.

Об'єктом дослідження є процес оброблення, передавання та зберігання мовленнєвих даних у комп'ютерних системах.

Предметом дослідження є методи та засоби ущільнення мовленнєвих сигналів для подальшого передавання та зберігання в комп'ютерних системах.

Методи дослідження: основні наукові результати і висновки одержані на основі: методів теорії цифрової обробки сигналів, яка використовувалась для визначення параметрів моделі лінійного прогнозування; методів теорії математичної статистики, яка використовувалась для дослідження розподілу різниці між індексами квантованих значень параметрів мовленнєвого сигналу; комп'ютерного моделювання для дослідження розроблених методів ущільнення мовленнєвих сигналів.

Наукова новизна отриманих результатів полягає в тому, що:

1. Вперше запропоновано метод дельта-ущільнення параметрів мовленнєвого сигналу, у якому, на відміну від існуючих, передається різниця між індексами поточного та прогнозованого значення векторів у структурованій кодовій книзі, що дозволило зменшити обсяги даних, які передаються та зберігаються у комп'ютерній системі.

2. Вдосконалено метод структуризації кодових книг, у якому, на відміну від існуючих, впорядкування здійснюється за принципами мажорювання векторів, що дозволило скоротити час пошуку в процесі квантування векторів у комп'ютерній системі оброблення мовленнєвих сигналів з низькою швидкістю.

3. Дістав подальшого розвитку адаптивний метод визначення голосової активності, у якому, на відміну від існуючих, використовується геометричний адаптивний поріг та згладжування спектральної обвідної, що дозволило підвищити швидкість комп'ютерної системи за умов низького співвідношення «сигнал/шум».

Практичне значення одержаних результатів.

1. Розроблено алгоритми та програмні засоби для ущільнення мовленнєвих сигналів на основі векторного квантування.

2. Розроблено методуку, алгоритми та програмні засоби для дослідження методів прискореного пошуку на основі мажорювання векторів у кодовій книзі.

3. Розроблено структуру комп'ютерної системи для оброблення мовленнєвих сигналів з підтримкою запропонованого методу дельта-ущільнення параметрів мовленнєвого сигналу.

4. Реалізовано удосконалений алгоритм визначення голосової активності.

Розроблений адаптивний метод визначення голосової активності з використанням згладжування спектральної обвідної та геометричного порогу в процесі передавання аудіоданих та рекомендації щодо особливостей застосування методів ущільнення аудіоданих впроваджено в Інституті електроніки та зв'язку Української академії наук.

Особистий внесок здобувача. Наукові положення, які містяться в дисертаційній роботі, отримані автором самостійно. У наукових працях, опублікованих у співавторстві, автору належать: у роботах [1, 2] – виділення мір для порівняння векторів та їх програмна реалізація, [3, 15] – програмна реалізація та експериментальні дослідження різних метрик для побудови кодових книг, [4] – програмна реалізація та експериментальні дослідження розподілу спектрального спотворення для різних видів розбиття векторної кодової книги, [5, 6, 10] – розробка та експериментальне дослідження методів структуризації за принципами мажорювання векторів та прискореного пошуку в структурованих кодових книгах, [7, 8, 18] – розробка та експериментальне дослідження методу дельта-ущільнення параметрів мовленнєвого сигналу, розробка програмного забезпечення для експериментального дослідження ефективності запропонованого методу, [9, 11] – вдосконалення адаптивного методу визначення голосової активності, програмна реалізація вдосконаленого методу, експериментальне дослідження ефективності методу, [16, 17] – метод прискореного пошуку

та реалізація програмного забезпечення для ущільнення мовленнєвих сигналів із застосуванням структурованих кодових книг.

Усі розробки, що містяться в дисертаційній роботі проводилися у Вінницькому національному технічному університеті.

Апробація результатів дисертації. Основні результати дисертаційної роботи доповідалися і обговорювалися на двадцяти міжнародних конференціях, зокрема, на 8-й та 9-й міжнародній конференції «Контроль і управління в складних системах» (м. Вінниця, 2005, 2008 рр.); 8-й міжнародній конференції з автоматичного управління “Автоматика-2006” (м. Вінниця, 2006); 5-й та 6-й міжнародних конференціях «Інтернет – освіта – наука» (м. Вінниця, 2006, 2008); 8-й і 9-й міжнародній науково-технічній конференції «Современные информационные и электронные технологии» (м. Одеса, 2007, 2008 рр.); 1-3-й міжнародній науково-практичній конференції “Методи та засоби кодування, захисту й ущільнення інформації” (м. Вінниця, 2007, 2009, 2011 рр.); 3-й міжнародній науково-технічній конференції «Инфокоммуникационные технологии в науке, производстве и образовании» (м. Ставрополь, 2008 р.), 9-й та 10-й Всеукраїнській міжнародній конференції з оброблення сигналів і зображень та розпізнавання образів «УкрОбраз’2008», «УкрОбраз’2010» (м. Київ, 2008, 2010 рр.), а також на 34-40-й щорічних конференціях професорсько-викладацького складу, співробітників та студентів університету з участю працівників науково-дослідних організацій та інженерно-технічних працівників підприємств м. Вінниці та області (м. Вінниця, 2005-2011 рр.).

Публікації. Основні результати дисертаційної роботи опубліковано у 29 наукових працях, серед яких 9 статей у науково-технічних фахових виданнях, що входять до переліку ВАК України, 15 робіт, опублікованих у збірниках матеріалів конференцій, 3 авторських свідоцтва на твір (комп’ютерну програму) і 2 патенти на корисну модель.

Структура та обсяг роботи. Дисертація складається зі вступу, чотирьох розділів, що містять 41 рисунок і 12 таблиць, висновків, додатків (5) і списку використаних джерел (86 найменувань). Загальний обсяг дисертації становить 176 сторінок, з яких основний зміст викладено на 117 сторінках.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** розкрито сутність і стан наукової проблеми, обґрунтовано актуальність теми дисертації та необхідність проведення досліджень, сформульовано мету та завдання дослідження, зазначено зв’язок із науково-дослідними програмами та темами. Сформульовано основні положення, які складають наукову новизну та практичну цінність отриманих результатів, наведено інформацію про впровадження та апробацію результатів роботи.

У **першому розділі** виконано аналіз існуючих систем ущільнення мовленнєвих сигналів, розглянуто основні характеристики та вимоги до таких систем. Виділено основні недоліки, що притаманні існуючим системам ущільнення мовленнєвих сигналів.

Показано, що найбільшого ступеня ущільнення дозволяють досягти параметричні методи ущільнення мовленнєвих сигналів, найбільш поширеними серед яких є методи, що базуються на методі лінійного прогнозування (LPC).

Основою даного методу є те, що на деякому відрізку мовленнєвого сигналу його наступні відліки можливо з деякою похибкою обрахувати на основі попередніх значень. Для цього сигнал розбивається на квазістаціонарні відрізки (кадри) по 20-30 мс. У середині одного відрізка коефіцієнти, що використовуються для отримання лінійної послідовності, обраховуються як прогноз сигналу $\hat{s}[n]$ на основі M попередніх відліків, виходячи з мінімізації середньоквадратичної похибки, з виразу:

$$\hat{s}[n] = \sum_{i=1}^M a_i \cdot s[n-i],$$

де a_i - коефіцієнти лінійного прогнозування LPC, M – порядок моделі прогнозування. Таким чином, прогнозування базується на основі лінійної комбінації M попередніх відліків сигналу. На практиці коефіцієнти лінійного прогнозування перетворюються в інші форми представлення, зокрема лінійні спектральні пари – LSP, які є більш стійкими до спотворень. Для подальшого ущільнення замість коефіцієнтів LSP передаються їх індекси в таблицях значень – кодових книгах. Показано, що квантування всього вектору дозволяє досягти меншого спотворення сигналу, порівняно з квантуванням окремих елементів, за однакового ступеня ущільнення, проте є обчислювально складним.

В результаті проведеного аналізу виділено основні тенденції розвитку систем ущільнення мовленнєвих сигналів, а саме: перехід від кодування форми сигналу до параметричного представлення у вигляді коефіцієнтів лінійного прогнозування та їх похідних, перехід від скалярного квантування параметрів сигналу до векторного із застосуванням кодових книг, використання кореляції як всередині кадру, так і між сусідніми кадрами для збільшення ступеня ущільнення.

Виходячи з цього, сформульовано задачу розробки методів та засобів ущільнення мовленнєвих сигналів, які б дозволили: отримати більший ступінь ущільнення за рахунок використання кореляції між кадрами сигналу, підвищити швидкодію комп'ютерної системи за рахунок використання методу визначення голосової активності, стійкого до помилок за умов низького співвідношення сигнал-шум, зменшити обчислювальну складність методів ущільнення за рахунок швидкого квантування параметрів мовленнєвого сигналу за допомогою векторних кодових книг.

У **другому розділі** розроблено теоретичні засади для побудови засобів векторного квантування параметрів мовленнєвих сигналів на основі мажорювання векторів. Вдосконалено міру оцінювання відстані між векторами параметрів мовленнєвого сигналу шляхом введення додакових вагових коефіцієнтів з метою надання більшої ваги низькочастотному спектру сигналу, спотворення якого є найбільш відчутним на слух. Вираз для зваженої міри має вигляд:

$$d(f, \hat{f}) = \sum_{i=1}^{10} [sc_i D_i (f_i - \hat{f}_i)]^2,$$

де f, \hat{f} – значення наборів коефіцієнтів LSP (лінійних спектральних пар) в Гц, sc_i – вагові коефіцієнти, що мають вигляд:

$$sc_i = \begin{cases} 0,95, & \text{для } i = 1, \\ 1,0, & \text{для } 2 \leq i \leq 9, \\ 0,8, & \text{для } i = 10, \end{cases}$$

D_i – i -й елемент **D** матриці чутливості, який визначається виходячи з виразу:

$$D_i = \frac{1}{F_B \Delta f_i^2} \int_0^{F_B} \left(10 \log_{10} \frac{P(f)}{\hat{P}(f)} \right)^2 df,$$

де F_B – частота дискретизації в герцах, $P(f), \hat{P}(f)$ - спектральні потужності сигналів, що описують дані параметри.

Важливо відзначити що для обрахунку матриці чутливості існують ефективні обчислювальні методи, які не потребують операцій ділення, піднесення до степеня чи обчислення квадратних коренів. Використання даної міри дозволяє отримати менше

спектральне спотворення при векторному квантуванні параметрів мовленнєвого сигналу, порівняно з евклідовою метрикою.

Запропоновано прискорений метод квантування векторів за допомогою кодових книг, який базується на розбитті кодових книг на рівні згідно відношення слабкого мажорювання. Кодова книга розбивається на K класів, що не перекриваються між собою. Кількість векторів у кожному класі може бути різною. Спочатку визначається приналежність вхідного вектора X до одного з класів C_k , на які розбито кодову книгу.

Для двох векторів $X = (x_1, x_2, \dots, x_n)$ та $Y = (y_1, y_2, \dots, y_n)$, $X, Y \in \mathfrak{R}^n$, компоненти яких впорядковано за незростанням говорять, що вектор Y слабо мажорює вектор X (позначається $X \prec_w Y$), якщо виконується співвідношення:

$$\sum_{i=1}^k x_i \leq \sum_{i=1}^k y_i, \quad k = 1, 2, \dots, n,$$

Оскільки структуризація кодових книг за відношенням слабкого мажорювання не гарантує знаходження найближчого вектора, перед розбиттям кодових книг відбувається перехід від оригінальних векторів до векторів відстаней від даних векторів до точок $V_0 = (0, 0, \dots, 0)$, $V_1 = (0, N_1, 0, \dots, 0)$, ..., $V_{n-1} = (0, 0, \dots, N_{n-1})$, де n – розмірність вектора параметрів. Вектори відстаней заносяться у другу кодову книгу і впорядковуються за рівнями мажорювання. Разом з ними перевпорядковуються відповідні вектори LSP-параметрів. Такий підхід вимагає вдвічі більше пам'яті для зберігання кодових книг, оскільки необхідно також зберігати в пам'яті вектори відстаней.

Основна ідея прискореного пошуку полягає у тому, що найближчі до вхідного вектора квантовані значення векторів у кодовій книзі, як правило, знаходяться на тому ж рівні мажорювання, що і вхідний вектор, або на сусідніх рівнях. За рахунок цього з'являється можливість значно скоротити число кандидатів на відбір i , відповідно, зменшити час пошуку.

Пошук найближчого вектора відбувається в два етапи. На першому етапі вхідний вектор LSP-параметрів X перетворюється у вектор відстаней до точок відліку X' . Далі у кодовій книзі відстаней відбувається пошук рівня мажорювання, якому належить вхідний вектор. При цьому послідовно, починаючи з верхнього, перевіряються всі рівні мажорювання. Якщо на рівні мажорювання L_j знаходиться вектор Y' , такий що $X' \prec_w Y'$, відбувається перехід до наступного рівня. При цьому решта векторів цього рівня не перевіряється. Пошук завершується, коли буде знайдено рівень L_k , на якому жоден вектор Y' не мажорює вхідний вектор X' , або коли досягається останній рівень мажорювання. На другому етапі відбувається пошук найближчого до вхідного квантованого вектора на знайденому рівні мажорювання та кількох сусідніх рівнях згідно заданого розміру вікна. Пошук виконується у кодовій книзі LSP-параметрів за методом повного перебору, при цьому вибирається найближчий вектор Y кодової книги, для якого зважена евклідова відстань є найменшою $D(X, Y) = \min$. Кількість векторів, що аналізуються на другому етапі, є фіксованою величиною, що залежить лише від необхідної точності пошуку і не залежить від кількості векторів на поточному рівні мажорювання та на сусідніх рівнях. Обчислювальна складність пошуку при цьому буде складати:

$$C_{maj} = \sum_{k=1}^p C_{Dmaj} \cdot L_k + \sum_{k=1}^p C_D \cdot M_k,$$

де p – кількість підвекторів у кодовій книзі, L_k – кількість рівнів для k -ї частини книги, M – розмір вікна пошуку, C_{Dmaj} , C_D – обчислювальна складність порівняння векторів за відношенням мажорювання та обрахування зваженої евклідової відстані відповідно. Для

кової книги розміром 4096 векторів та розмірністю $n = 5$ обчислювальна складність пошуку в попередньо впорядкованій векторній кодовій книзі складає приблизно 17,23% від складності повного пошуку. Отже даний прискорений пошук векторів у кодовій книзі з розбиттям на рівні мажорювання дозволяє зменшити обчислювальну складність, порівняно з повним пошуком у 5,8 разів.

Впорядкування кодової книги за відношенням мажорювання дозволяє використовувати кореляцію між сусідніми кадрами сигналу на рівні індексів векторів у кодовій книзі, що дозволяє отримати збільшення степеня ущільнення. Після впорядкування кодової книги квантоване значення наступного за даним вектору LSP-параметрів, як правило, лежить на тому ж або сусідніх рівнях, за рахунок чого є можливим представити вектори як відстані між індексами в структурованій кодовій книзі. Це проілюстровано на рис. 1, де зображено значення кореляції R для 1-го, 4-го та 7-го елементів вектора LSP ($f_i[m]$) для m послідовних кадрів. З даного рисунку видно, що кореляція коефіцієнтів для послідовних кадрів є значною.

Отже можливо здійснити перехід до передачі різниці (дельти) між індексами векторів послідовних кадрів, замість індексів самих векторів і таким чином отримати додаткове зменшення обсягів даних для опису одного кадру. Додаткове ущільнення відбувається за рахунок того, що різниця між індексами обмежується розміром вікна та є меншою за загальну кількість векторів кодової книги, що дозволяє описувати її меншою кількістю біт.

Для зменшення спектрального спотворення при дельта-ущільненні пропонується прогнозування індексу кадру на основі попередніх значень. Прогнозування відбувається шляхом створення екстраполюючої функції для індексів кількох кадрів на приймальній та передавальній стороні.

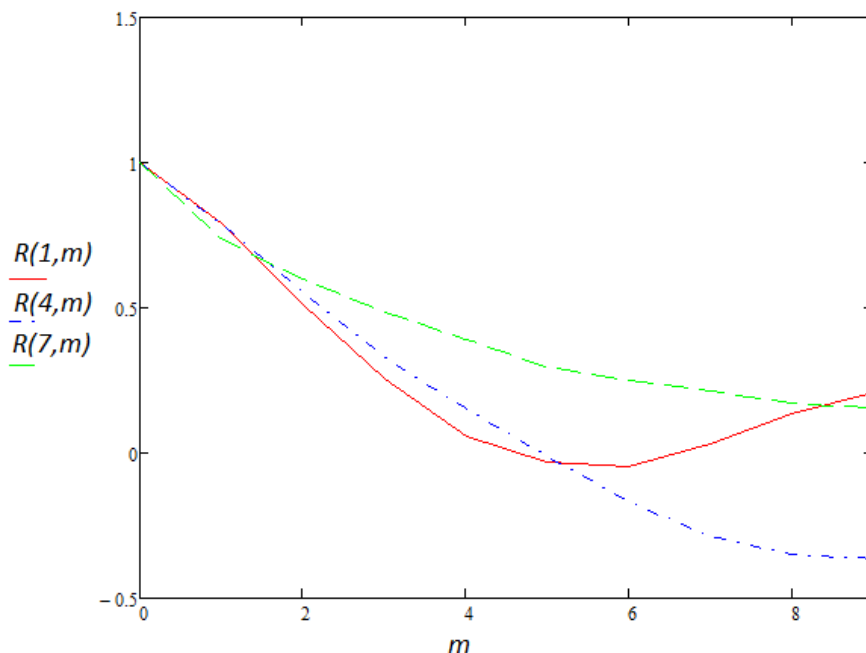


Рисунок 1 - Значення кореляції коефіцієнтів LSP для послідовних кадрів

При ущільненні на приймальну сторону передається дельта між прогнозованим та реальним значеннями. Основну ідею запропонованого методу представлено на рис. 2. Екстраполююча функція будується за методом найменших квадратів, який дає найкращу швидкодію та точність порівняно з іншими досліджуваними методами.

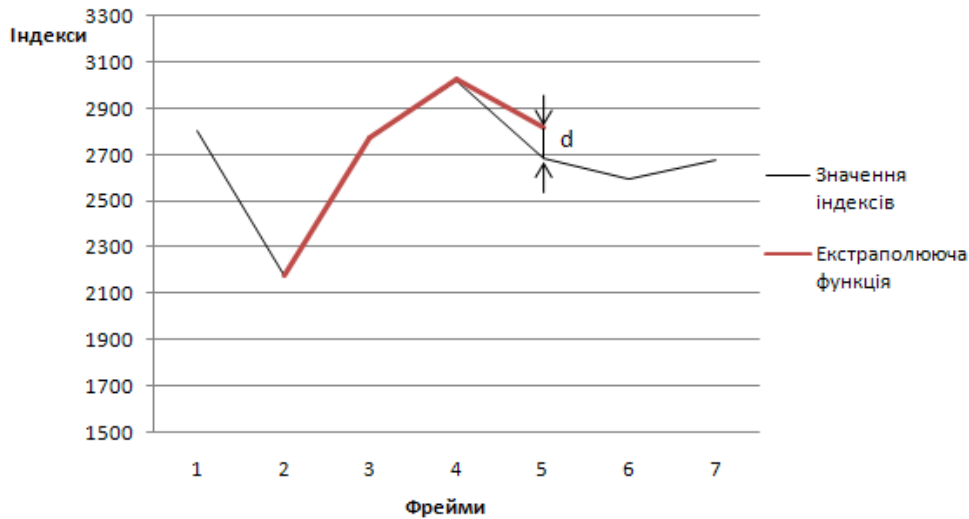


Рисунок 2 – Ілюстрація роботи методу ущільнення з екстраполяцією

Даний метод потребує введення затримки кодування, яка дозволяє корегувати значення індексів, що передаються в ситуаціях, коли прогноз виявляється невірним, при цьому відбувається корегування значень дельти за таким алгоритмом:

1. Здійснюється апроксимація значень експериментальної функції, враховуючи наступні значення індексів $i+1, i+2, \dots, i+k$, де k – величина затримки в кадрах. Для апроксимації також використовується метод найменших квадратів.

2. Здійснюється корекція значення i -го індекса, згідно до апроксимуючої функції, для того, щоб отримати кращий прогноз для наступного кадру.

3. Здійснюється прогноз $i+1$ кадру на основі попередніх кадрів $i, i-1, i-2, \dots, i-p$, де p – кількість кадрів для створення прогнозу, яка визначається порядком моделі прогнозування.

4. Якщо різниця між прогнозованим та реальним значенням не попадає в межі вікна, то корекція відбувається знову – відбувається перехід до п. 2. В крайньому випадку коли при максимально допустимому відхиленні i -го кадру прогноз всеодно є неуспішним, то всеодно відбувається перехід до п. 5.

5. Різниця між корегованим значенням та прогнозованим враховуючи значення індексів для $i-1, i-2, \dots, i-(p+1)$ кадрів, кодується та передається в канал зв'язку.

Результат роботи даного алгоритму представлено на рис. 3.

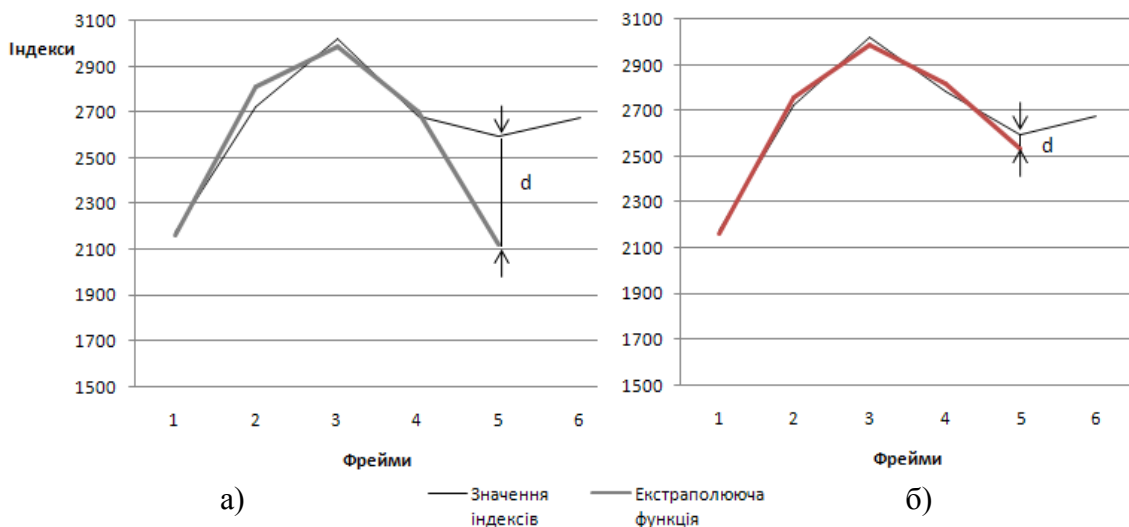


Рисунок 3 – Графіки експериментального та екстрапольованого значення індексів:

а) до виконання корекції; б) після виконання корекції

Експериментально встановлено, що для досягнення прийнятних значень спектрального спотворення достатньо ввести в сигнал затримку лише в один кадр.

Запропонований метод ущільнення мовленнєвих сигналів за рахунок внесення незначного додаткового спотворення ($\Delta SD = 0,135$ дБ) та затримки передачі сигналу дозволяє скоротити обсяг даних необхідних, для передавання параметрів мовленнєвого сигналу з 24 біт до 22 біт на кадр, тим самим зменшити швидкість, необхідну для передавання мовленнєвого сигналу по мережі з 2400 до 2178 біт/с.

У **третьому розділі** сформульовано вимоги та розроблено комп'ютерну систему передавання та зберігання мовленнєвих сигналів з використанням запропонованих методів. Розроблювана комп'ютерна система призначена для забезпечення голосового зв'язку між територіально розподіленими користувачами по низькошвидкісних IP-мережах. Для реалізації системи передавання мовленнєвих сигналів за основу взято архітектуру мережі на основі протоколу SIP. Ключовим моментом є те, що протокол SIP дозволяє використовувати сторонні мовленнєві кодеки без впливу на інші компоненти системи. Для зменшення навантаження на канали зв'язку та збільшення кількості абонентів запропоновано використання розробленого програмного клієнту SIP протоколу та SIP-проксі з опціональним перекодуванням в точках де критичною є ширина каналу. Для забезпечення ущільнення реалізовано вокодер на основі MELP, який вдосконалено за рахунок введення векторного квантизатора та дельта-ущільнення параметрів, запропонованих в розділі 2, та підсистеми визначення голосової активності, що разом дозволяє зменшити бітову швидкість необхідну для вокодера з 7.8 до 4.68 Кбіт/сек. Дані програми підтримують як нормальний режим роботи, так і додаткові режими ущільнення мовленнєвого сигналу, використовуючи вдосконалений кодек MELP. На рис. 4 наведено базову структуру комп'ютерної системи передавання мовленнєвих сигналів.

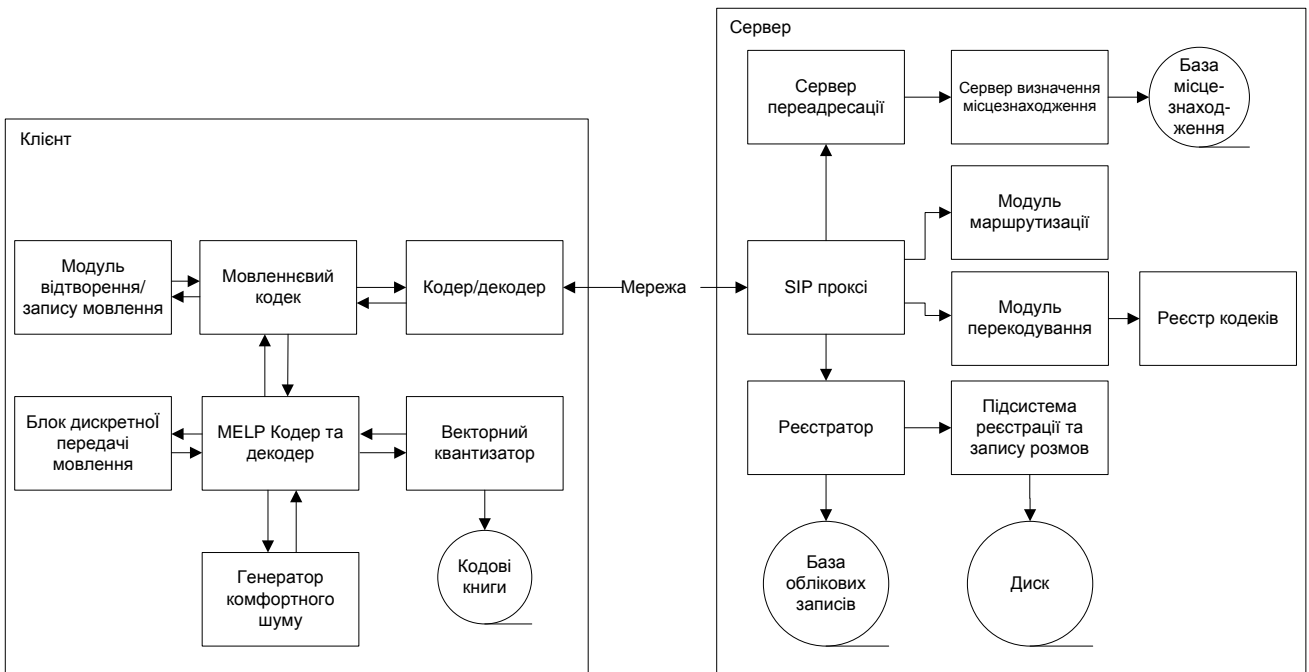


Рисунок 4 – Базова структура комп'ютерної системи передавання мовленнєвих сигналів

Для реалізації системи передавання мовленнєвих сигналів було модифіковано структуру open source реалізації SIP-проксі – Yate. При розробці вокодера з дельта-ущільненням параметрів мовленнєвого сигналу за основу було взято реалізацію вокодера MELP на мові C, в який було додано модуль векторного квантування параметрів з дельта-ущільненням та модуль визначення голосової активності та передачі комфортного шуму під

час пауз у розмові, що дозволяє зменшити результуючу швидкість, необхідну для передавання мовленнєвого сигналу з 7,8 Кбіт/с до 4,68 Кбіт/с. Також запропоновано функціональну схему пристрою для пошуку найближчого вектора параметрів сигналу у кодових книгах, що є найбільш вартісною операцією в процесі ущільнення мовленнєвого сигналу.

Створено окрему підсистему для запису та реєстрації розмов, виконану як доповнення до SIP проксі-сервера. Дана підсистема отримує інформацію про всі виклики, що проходять через сервер, та мовленнєві дані. За необхідності мовленнєві дані перекодовуються за допомогою запропонованого кодеку та архівуються на диски. Підсистему зберігання реалізовано з використанням арифметичного кодування параметрів для подальшого ущільнення і вона дозволяє записувати потоки мовленнєвих даних в ущільненому вигляді на диск для довготривалого зберігання з можливістю подальшого відтворення.

Вдосконалено метод визначення голосової активності для зменшення його обчислювальної складності та кількості неправильних спрацювань за умов значного фоновому шуму. За основу для подальшої розробки взято метод визначення голосової активності G. 729 Annex B, до якого запропоновано додати згладжування спектральної обвідної та використання геометрично адаптивного енергетичного порогу.

Рішення про вокалізованість приймаються для фреймів тривалістю 10 мс на основі таких характеристик:

1. Диференційна енергія сигналу для всього діапазону спектру сигналу – ΔE_f ;
2. Диференційна енергія сигналу для низькочастотного спектру (0-1 кГц) – ΔE_l ;
3. Диференційна частота переходів через нуль – ΔZCR ;
4. Спектральне спотворення – ΔLSP .

Вхідні параметри отримуються з мовленнєвого сигналу використовуючи авторегресійну модель першого порядку, якщо виконується умова $\Delta E_f < E_t$, де E_t – задане порогове значення. Проте дані параметри потребують значної складності обчислень, також використання частоти переходів через нуль та порогу різниці енергії, які привносять похибки при високій зашумленості сигналу.

Згладжування спектральної обвідної сигналу за допомогою апроксимуючої функції другого порядку зменшує обчислювальну складність методу. Згладжування відбувається не для всього діапазону частот, а для трьох частотних діапазонів від 0 до 1000 Гц, 1000-2000 Гц та 2000-4000 Гц. Дані діапазони обрано виходячи з того, що основна інформація мовленнєвого сигналу знаходиться в низькочастотній частині спектру. Таким чином, це дозволяє зменшити кількість обчислень, оскільки пологість спектру обраховується тільки по кількох точках, та підвищити точність алгоритму, особливо у низькочастотній смузі сигналу.

Метод геометричного адаптивного енергетичного порогу базується на припущенні про адитивність шуму, отже зашумлений сигнал може бути описаний як:

$$x(t) = s(t) + u(t),$$

де $s(t)$ – оригінальний мовленнєвий сигнал, $u(t)$ – шумовий сигнал.

Якщо дискретні значення $x(t)$ позначити як $x[n]$, тоді

$$x[n] = x(T_1 + n\Delta t), \quad \Delta t = \frac{T_2 - T_1}{N},$$

де N – кількість вибірок взята з часу T_1 по T_2 .

Тоді оцінка функції розподілу сигналу $A_x[m]$ – визначається як:

$$A_x[m] = \sum_{k=0}^m a_x[k],$$

де $a_x[m]$ – кількість відліків сигналу, які задовольняють умовам:

$$m\Delta x \leq |x[n]| \leq (m+1)\Delta x.$$

Аналогічним чином можна визначити параметри $A_s[m]$ ($a_s[m]$) та $A_u[m]$ ($a_u[m]$) використовуючи відліки де $s[n]$ та $u[n]$ відповідно. Функція щільності сигналу наблизатиметься до $f_{|x|}(x)$ $\Delta x \rightarrow 0$, $\Delta t \rightarrow 0$. Якщо розглянути функції $A_s[m]$ та $A_u[m]$ можна помітити, що сигнал і шум займають різні регіони та можуть бути частково розділені. Для чого використовується геометричний метод, який полягає в тому, що порогове значення знаходиться на перетині функції розподілу амплітуди сигналу з прямою, що проходить через координати (0,1) та точку перетину дотичних до функції розподілу в крайніх точках (рис. 5). На рисунку наведено графік розподілу ймовірності амплітуди, де OO' та MM' – дотичні до графіку розподілу, точка T – шукане порогове значення.

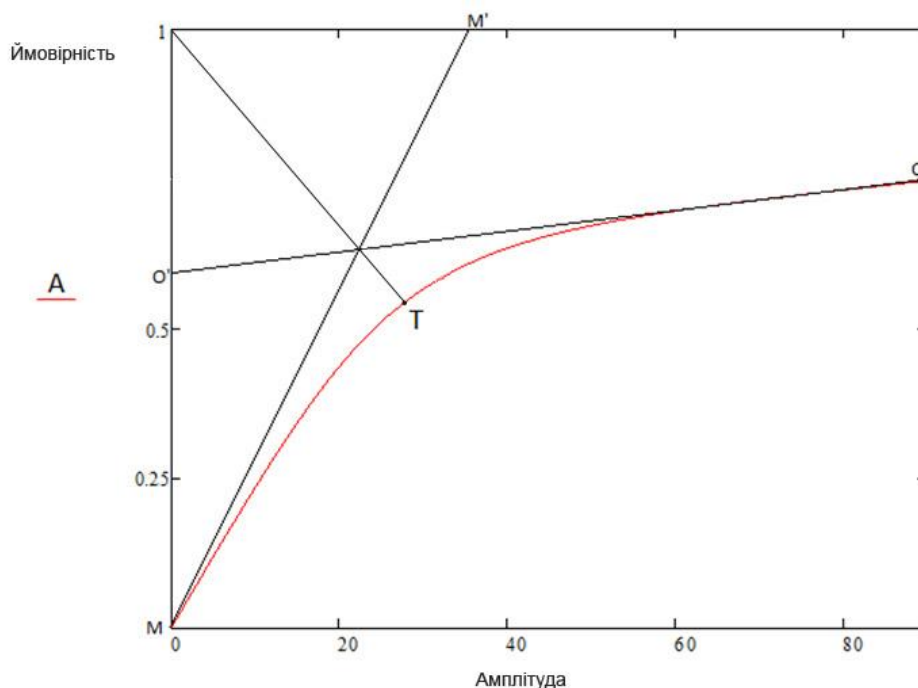


Рисунок 5 – Геометричний метод знаходження порогового значення

Даний метод є стійким до нестационарного шуму, але іноді може давати неправильний результат для різких й коротких стрибків амплітуди шуму. Завдяки простоті обчислення адаптивного енергетичного порогу також зменшується складність алгоритму визначення голосової активності.

Наведено структуру пакету для передавання мовленнєвих даних по IP-мережах. Розраховано, що максимальна кількість кадрів, які доцільно передавати в одному пакеті, становить не більше 3-х. При цьому досягається максимальна швидкодія та дотримуються умови максимальної затримки для систем передачі мовленнєвих сигналів.

У **четвертому розділі** проведено експериментальні дослідження розроблених методів та описано реалізацію розробленої системи ущільнення мовлення.

Для дослідження різних методів пошуку у векторних кодових книгах було розроблено спеціалізоване програмне забезпечення. Дане програмне забезпечення отримує дані із звукового файлу чи звукової карти, обчислює коефіцієнти лінійного прогнозування, перетворює їх до іншого представлення – лінійних спектральних пар та виконує квантування отриманих параметрів з використанням різних варіантів векторних кодових книг.

Наведено умови тестування методів ущільнення. Використовувались тренувальні вибірки з загальнодоступної англійської звукової бази ТІМІТ, та вибірки, що склалися з

двох україномовних текстів, надиктованих різними голосами. Тренувальна вибірка містила 169286 кадрів, тестова вибірка – 25781 кадр по 20 мс кожен.

Проведено експериментальне дослідження варіантів розбиття кодової книги на рівні мажорювання. Побудовано розподіл найближчих до вхідного векторів за рівнями, отриманими за результатами мажорювання. З розподілу визначено, що більше 99% векторів знаходяться у межах ± 3 рівні. Було досліджено використання методу прискореного пошуку порівняно з повним пошуком, методом пошуку з використанням класифікації векторів за середнім елементом та класифікації за зваженим середнім значенням вектора. Результати дослідження наведено в табл. 1.

Таблиця 1 – Результати дослідження прискореного пошуку векторів

Метод пошуку	Середнє спектральне спотворення, дБ	Кількість операцій порівняння	Відносна складність обчислень S , %	Кількість пропущених векторів, %
Повний пошук	0,746	16 777 216	100%	0
Класифікація за середнім елементом	0,786	6 710 886	40%	0,195
Зважена класифікація	0,771	6 710 886	40%	0,25
Пошук по впорядкованій книзі	0,7653	2 890 714	17%	0,13

Наведено результати експериментального дослідження методу визначення голосової активності порівняно з оригінальним методом та іншими часто використовуваними методами. Результати тестування показали, що покращений метод G.729 Annex B дозволяє отримати більший відсоток правильно розпізнаних фреймів мовлення на зашумлених сигналах, ніж звичайний метод (83% проти 68% у оригінальному методі), але дещо гіршу, ніж досягається в методі з використанням статистики високого порядку (87%), який, проте, є обчислювально складним.

Таблиця 2 – Результати дослідження прискореного пошуку векторів

Назва методу	Фреймів з обрізанням мовлення, %	Шумових фреймів розпізнаних як вокалізовані, %
Покращений G.729 Annex B	83%	2%
G.729 Annex B	68%	5%
AMR 2	87%	0,3%
AMR 1	76%	5,5%

Описано деталі реалізації системи для передавання та зберігання мовленнєвих сигналів у телекомунікаційній мережі «МІТРИС». Дана система дозволяє передавати телевізійні, радіо та цифрові дані в мікрохвильовому діапазоні. Ця система використовує ті ж частоти, що і супутникові системи теле- і радіомовлення, проте найважливішою перевагою системи «МІТРИС» є можливість забезпечення інформаційного обміну на значній території – десятки кілометрів при розміщенні однієї базової станції на висоті близько 100 м. Така система може надавати різноманітні телекомунікаційні послуги, зокрема багатоканальне аналогове чи цифрове телебачення, передавання цифрових даних та забезпечення цифрової телефонії. Для передавання мовленнєвих сигналів по мережі «МІТРИС» використовується

клієнтське та серверне програмне забезпечення, описане в розділі 3. Використання розробленого програмного забезпечення дозволило збільшити можливу кількість абонентів в даній мережі, що розмовляють одночасно до 260 та забезпечити одночасний запис до 127 розмов на вимогу.

Проведено оцінювання розбірливості синтезованого мовленнєвого сигналу за оцінками MOS. Для тестування використовувались артикуляційні тести, описані в стандарті ГОСТ Р 50840-95 «Передача речі по трактам зв'язи. Методи оценки качества, разборчивости и узнаваемости». В результаті проведеного тестування отримано значення розбірливості $D_{T, \text{ср дов.}} = 87,10$, що відповідає III класу розбірливості мовлення, тобто мовлення є розбірливим, з навантаженням уваги без перепитувань та повторів. Також в результаті тестів була отримана оцінка ACR, яка становить 3,36, що є припустимим для низькошвидкісних кодеків.

У додатках наведено акт впровадження результатів дисертаційної роботи, результати досліджень розроблених методів та вихідні тексти програмних засобів для дослідження методів ущільнення мовленнєвих сигналів та передавання ущільнених мовленнєвих сигналів.

ОСНОВНІ РЕЗУЛЬТАТИ РОБОТИ ТА ВИСНОВКИ

У дисертаційній роботі вирішено важливу науково-практичну задачу, пов'язану з зменшенням обсягів даних, необхідних для представлення спектральної інформації сигналу в системах передавання мовленнєвих сигналів у комп'ютерних системах. Розроблено методи та інструментальні засоби для дослідження різних методів прискореного пошуку векторів у кодовій книзі.

1. Проведено аналіз та класифікацію існуючих методів ущільнення мовленнєвих сигналів, які дозволили виявити основні тенденції їх розвитку. За основу подальших досліджень вибрано методи параметричного ущільнення мовленнєвих сигналів з векторним квантуванням LSP та урахуванням міжкадрової кореляції параметрів сигналу.

2. Вперше запропоновано метод дельта-ущільнення з прогнозуванням параметрів, суть якого полягає у передачі різниці між індексами поточного та попереднього векторів у структурованій кодовій книзі, що дозволило зменшити кількість бітів, необхідних для кодування спектральної інформації, з 24 до 22 без суттєвого збільшення спектрального спотворення.

3. Удосконалено метод пошуку найближчого вектора в кодовій книзі, який, на відміну від існуючих методів, базується на попередньому впорядкуванні векторів за принципом мажорювання, що дозволило скоротити час пошуку в 5,8 разів.

4. Дістав подальший розвиток адаптивний метод визначення голосової активності, у якому, на відміну від існуючих, використовується геометричний адаптивний поріг та згладжування спектральної обвідної, що дозволило зменшити загальну кількість помилок на 15% та підвищити швидкодію методу на 12,5%.

5. Розроблено методику та програмні засоби для дослідження методів ущільнення мовленнєвих сигналів з використанням векторного квантування, а також алгоритмічне та програмне забезпечення для ущільнення мовленнєвих сигналів для їх зберігання та передавання в комп'ютерних системах.

6. Застосування запропонованих методів під час кодування мовленнєвих сигналів дозволяє додатково зменшити обсяги мовленнєвих даних у 1,6 разів. Це дозволяє підвищити швидкість оброблення мовленнєвих даних у комп'ютерних системах, призначених для роботи з мультимедійною інформацією.

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

За темою досліджень опубліковано 29 наукових праць, основними з них є:

1. Ткаченко О. М. Вокодер LSF зі швидкістю 1600 біт/с / О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак // Вісник Вінницького політехнічного інституту. – 2006. – № 1. – С. 78 – 86.
2. Ткаченко О. М. Розробка кодових книг для вокодера на основі LSP / О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак // Оптико-електронні інформаційно-енергетичні технології. – 2006. – № 2. – С. 219 - 223. – ISSN 1681-7893.
3. LSF-вокодер на основі векторного квантування / Н. О. Біліченко, О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак // Реєстрація, зберігання і обробка даних. – 2007. – № 1. – С. 35 – 41. – ISSN 1560-9189.
4. Ткаченко О. М. Ефективне векторне квантування LSF-параметрів при ущільненні мовних сигналів / О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак // Інформаційні технології та комп'ютерна інженерія. – 2007. – № 1. – С. 124-129.
5. Швидкий пошук при векторному квантуванні лінійних спектральних частот / Н. О. Біліченко, О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак // Реєстрація, зберігання і обробка даних. – 2008. – № 2. – С. 37 – 47. – ISBN 1560-9189.
6. Ткаченко О. М. Ефективний пошук векторів у кодовій книзі / О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак // Проблеми інформатизації та управління. – 2007. – № 22. – С. 121 - 127. – ISBN 5-7763-4415-8.
7. Ткаченко О. М. Метод дельта-ущільнення мовних сигналів / О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак // Інформаційні технології та комп'ютерна інженерія. – 2008. – № 1(11). – С. 8-13.
8. Ткаченко О. М. Дельта-ущільнення мовленнєвих сигналів з прогнозуванням індексів кодової книги / О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак // Вісник Вінницького політехнічного інституту. – 2009. – № 1. – С. 85–88.
9. Ткаченко О. М. Адаптивне визначення голосової активності при дельта-ущільненні мовленнєвих сигналів [Електронний ресурс] / Ткаченко О. М., Крупельницький Л. В., Хрущак С. В. // Наукові праці Вінницького національного технічного університету. – № 3. – 2011. Режим доступу: http://www.nbu.gov.ua/e-journals/vntu/2011_3/2011-3.files/uk/11omtoss_ua.pdf.
10. Ущільнення мови із застосуванням структурованих кодових книг / О. Д. Азаров, О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак // Труды девятой международной научно-практической конференции «Современные информационные и электронные технологии». – Одесса, 2008. – С. 28.
11. Ткаченко О. М. Дельта-ущільнення мовленнєвих сигналів із застосуванням методу визначення голосової активності / О. М. Ткаченко, О. Д. Феферман, С. В. Хрущак // Матеріали Міжнародної науково-практичної конференції. Укробраз-2010. – Київ, 2010. – С. 193-198.
12. Хрущак С. В. Дельта ущільнення мовлення з використанням інтервального прогнозування // Збірник тез доповідей Міжнародної науково-практичної конференції «Інформаційні технології та комп'ютерна інженерія». – Вінниця, 2011. – С.172-173.
13. Хрущак С. В. Визначення голосової активності з використанням адаптивного порогу // Збірник тез доповідей III МНТК «Методи та засоби кодування, захисту й ущільнення інформації». – Вінниця, 2011. – С.178-179. – ISBN 978-966-641-406-2.
14. А. с. 12396. Комп'ютерна програма аналізу алгоритмів ущільнення мови / Ткаченко А. Н., Феферман О. Д., Хрущак С. В. - № 14842; заявл. 10.01.05; опубл. 03.03.2005. – 5 с.
15. А. с. 24757. Комп'ютерна програма «Програма структуризації векторних кодових книг для застосування в системах цифрового зв'язку» / Ткаченко А. Н., Феферман О. Д., Хрущак С. В. - № 25094; заявл. 18.03.08; опубл. 17.06.2008. – 5 с.
16. А.с. 26051. Комп'ютерна програма «Програма ущільнення мовних сигналів із застосуванням структурованих кодових книг» / Ткаченко А. Н., Феферман О. Д., Хрущак С. В. - № 26051; заявл. 02.07.08; опубл. 13.10.2008. – 5 с.

17. Пат. 32410 Спосіб ущільнення мовних сигналів. МПК (2006) G10L 21/00 G10L 19/00 / Ткаченко О.М., Феферман О.Д., Хрущак С. В. власник патенту вінницький національний технічний університет. — № 200800956; заявл. 28.01.08 ; опубл. 12.05.08, Бюл. № 9.

18. Пат. 38505 Спосіб дельта-ущільнення мовних сигналів друк. Патент України на корисну модель / Ткаченко О. М., Феферман О. Д., Хрущак С. В. власник патенту вінницький національний технічний університет / 2009, заявл. 12.01.09, Бюл. №1.

АНОТАЦІЯ

Хрущак С. В. Векторне квантування параметрів сигналу при ущільненні мовлення в комп'ютерних системах. – На правах рукопису.

Дисертація на здобуття наукового ступеня кандидата технічних наук зі спеціальності 05.13.05 – комп'ютерні системи та компоненти. – Вінницький національний технічний університет, Вінниця, 2012.

Дисертацію присвячено розробленню комп'ютерної системи ущільнення мовленнєвих сигналів на основі векторного квантування параметрів сигналу. У роботі розв'язано науково-технічну задачу підвищення продуктивності комп'ютерних систем оброблення мовленнєвих сигналів за рахунок розробки програмно-апаратних засобів ущільнення даних, які необхідні для подальшого передавання або відтворення мовленнєвого сигналу без втрати прийнятної якості декодованого мовлення. Вдосконалено метод структуризації векторних кодових книг, який, на відміну від існуючих, базується на принципі мажорювання векторів, що дозволило скоротити час квантування набору параметрів. Запропоновано метод дельта-ущільнення з прогнозуванням параметрів мовленнєвого сигналу, що дозволило зменшити обсяги даних, що передаються. Вдосконалено метод визначення голосової активності для зменшення складності обчислень та підвищення точності за умов значної зашумленості сигналу. Розроблено методику та програмні засоби для дослідження методів ущільнення мовлення на основі векторного квантування. Експериментально показано, що застосування розроблених методів, дозволяє зменшити обчислювальну складність та обсяги даних, що передаються в системах ущільнення мовлення.

Ключові слова: параметричне ущільнення мовлення, структуризація кодових книг, теорія мажорювання, дельта-ущільнення мовленнєвих сигналів, визначення голосової активності, метод лінійного прогнозування.

АННОТАЦИЯ

Хрущак С. В. Векторное квантование параметров сигнала при сжатии речи в компьютерных системах. – На правах рукописи.

Диссертация на соискание ученой степени кандидата технических наук по специальности 05.13.05 – компьютерные системы и компоненты. – Винницкий национальный технический университет, Винница, 2012.

Диссертация посвящена разработке компьютерной системы сжатия речевых сигналов на основе векторного квантования параметров сигнала. В работе решена научно-техническая задача повышения производительности компьютерных систем обработки речевых сигналов за счет разработки программно-аппаратных средств сжатия данных, необходимых для дальнейшей передачи или воспроизведения речевого сигнала без потери приемлемого качества декодированной речи.

Проанализированы существующие подходы к построению систем параметрического сжатия речи, выделены основные недостатки таких систем. Сформулированы основные требования к системам сжатия речи. Проведенный анализ позволил определить основные тенденции развития таких систем и сформулировать задачи исследований.

Разработаны теоретические основы для построения средств векторного квантования параметров речевых сигналов на основе мажорирования векторов. Усовершенствован метод структуризации векторных кодовых книг, который, в отличие от существующих, базируется

на принципе мажорирования векторов, что позволило сократить время квантования набора параметров. На основе разработанного метода структуризации кодовых книг предложен ускоренный метод квантизации параметров речевого сигнала. Основная идея ускоренного поиска заключается в том, что ближайшие к входному вектору квантованное значение вектора в кодовой книге, как правило, находятся на том же или соседних уровнях мажоризации, что и входной вектор. За счет этого появляется возможность значительно сократить число кандидатов на отбор и, соответственно, уменьшить время поиска.

Структуризация кодовой книги за счет мажорирования позволяет использовать корреляцию между соседними кадрами сигнала на уровне индексов векторов в кодовой книге, что позволяет получить увеличение степени уплотнения. Так что возможно осуществить переход к передаче разницы (дельты) между индексами векторов последовательных кадров, вместо индексов самих векторов и таким образом получить дополнительное уменьшение объемов данных, необходимых для описания одного кадра. Дополнительное уплотнение происходит за счет того, что разница между индексами ограничивается размером окна и меньше за общее количество векторов кодовой книги, что позволяет описывать ее меньшим количеством бит. Для уменьшения спектрального искажения при дельта-уплотнении предлагается прогнозирование индекса кадра на основе предыдущих значений. Прогнозирование происходит путем создания экстраполирующей функции для индексов нескольких кадров на приемной и передающей стороне.

Усовершенствован метод определения голосовой активности для уменьшения сложности вычислений и повышения точности в условиях значительной зашумленности сигнала. За основу для дальнейшей разработки взят метод определения голосовой активности описанный в G. 729 Annex B. В данный метод добавлено сглаживание спектральной огибающей сигнала с помощью аппроксимирующей функции второго порядка, которое позволяет уменьшить вычислительную сложность метода, за счет того, что пологость спектра сигнала рассчитывается только по нескольким точкам. А также метод геометрического адаптивного порога, который есть более стойким к нестационарному шуму и уменьшает вычислительную сложность метода.

Сформулированы требования и разработана компьютерная система передачи и хранения речевых сигналов с использованием предложенных методов. Разработанная компьютерная система предназначена для обеспечения голосовой связи между территориально распределенными пользователями по низкоскоростным IP-сетям.

Разработана методика и программные средства для исследования методов сжатия речи на основе векторного квантования. Экспериментально показано, что применение разработанных методов, позволяет уменьшить вычислительную сложность и объемы передаваемых данных в системах сжатия речи. Проведено оценивание разборчивости синтезированного речевого сигнала за оценками MOS и тестами описанными в стандарте ГОСТ Р 50840-95 «Передача речи по трактам связи». Методы оценки качества, разборчивости и узнаваемости», значение ACR составляет 3.36, что есть нормальным для низкоскоростных методов сжатия речи.

Результаты диссертационных исследований подтверждаются актом внедрения в телекоммуникационной системе «МИТРИС».

Ключевые слова: параметрическое сжатие речи, структуризация кодовых книг, теория мажоризации, дельта-сжатие речевых сигналов, определение голосовой активности, метод линейного прогнозирования.

ABSTRACT

Khruschak S. V. Vector quantization of signal parameters for speech compression in computer systems. – A manuscript.

The thesis for a Ph.D. science degree on a specialty 05.13.05 – Computer Systems and Components. Vinnytsia National Technical University. Vinnytsia, 2012.

The thesis is devoted to development of computer system ащк speech signals compression based on vector quantization of signal parameters. In this research the important scientific and technical problem of computer speech signals processing systems productivity is solved by developing software and hardware means for data compression. These means allows to reduce amount of data necessary for further transmission or speech signal reproduction without loss of acceptable quality of decoded speech. Method for vector code books structurization is improved, which, unlike the existing ones, based on the vectors majorization principle and allows to reduce the time necessary for parameters quantization.

The method of speech parameters delta-compression with prediction is proposed, which allows to reduce the amount of data transmitted. Voice activity detection method is improved by reducing its complexity and increasing accuracy under high signal to noise ratios. The methodology and software for speech compression methods investigation is developed. Experimentally shown that the use of the developed methods can reduce computational complexity and amount of transmitted data in speech compression systems.

Keywords: parametric speech compression, codebooks structurization, majorization theory, delta-compression of speech signals, voice activity detection, linear predictive coding.

Підписано до друку 09.04.2012 р. Формат 29.7×42 ¼

Наклад 100 прим. Зам. № 2012-048

Віддруковано в комп'ютерному інформаційно-видавничому центрі

Вінницького національного технічного університету.

м. Вінниця, вул. Хмельницьке шосе, 95. Тел.: 59-87-38