

МІНІСТЕРСТВО ТРАНСПОРТУ ТА ЗВ'ЯЗКУ УКРАЇНИ
ДЕРЖАВНА АДМІНІСТРАЦІЯ ЗВ'ЯЗКУ
ОДЕСЬКА НАЦІОНАЛЬНА АКАДЕМІЯ ЗВ'ЯЗКУ ім. О. С. ПОПОВА

Рудий Є.М., Русляченко О.Ю., Рябуха О.М.

ТЕХНОЛОГІЇ ПЕРЕДАЧІ ДИСКРЕТНИХ ПОВІДОМЛЕНЬ

МОДУЛЬ 1. Стиснення телефонних сигналів
МОДУЛЬ 2. Стиснення даних, сигналів зображень
та звукового мовлення

Методичні вказівки
до лабораторних занять

для студентів напряму бакалаврської підготовки «Телекомунікації»

Одеса 2011

УДК 534.78:621.39

План НМВ 2011 р.

Рудий Є.М. Технології передачі дискретних повідомлень: метод. вказівки для лаб. занять. Модуль 1, Модуль 2 / Рудий Є.М., Русляченко О.Ю., Рябуха О.М. – Одеса: ВЦ ОНАЗ, 2011. – 84 с.

Містить методичні вказівки для проведення лабораторних занять за модулями 1 та 2 з дисципліни Технології передачі дискретних повідомлень.

Відпов. редактор – проф. *Захарченко М.В.*

СХВАЛЕНО
на засіданні кафедри
Інформаційної безпеки
та передачі даних
і рекомендовано до друку.
Протокол №9
від 10 червня 2011 р.

Комп'ютерне редагування та макетування – *Кірдогло Т.В.*

ЗМІСТ

С.

Лабораторна робота 1.1	
ДОСЛІДЖЕННЯ КАНАЛУ ТОНАЛЬНОЇ ЧАСТОТИ.....	4
Лабораторная работа 1.1	
ИССЛЕДОВАНИЕ КАНАЛА ТОНАЛЬНОЙ ЧАСТОТЫ.....	8
Лабораторна робота 1.2	
МОДЕЛЮВАННЯ ГАРМОНІЧНОГО ВОКОДЕРА.....	12
Лабораторная работа 1.2	
МОДЕЛИРОВАНИЕ ГАРМОНИЧЕСКОГО ВОКОДЕРА.....	18
Лабораторна робота 1.3	
ДОСЛІДЖЕННЯ МОДЕМА З ЕХОКОМПЕНСАТОРОМ.....	24
Лабораторная работа 1.3	
ИССЛЕДОВАНИЕ МОДЕМА С ЭХОКОМПЕНСАТОРОМ.....	31
Лабораторна робота 1.4	
МОДЕЛЮВАННЯ СИСТЕМИ СТИСНЕННЯ СИГНАЛІВ	
ЗВУКОВОГО МОВЛЕННЯ	37
Лабораторная работа 1.4	
МОДЕЛИРОВАНИЕ СИСТЕМЫ СЖАТИЯ СИГНАЛОВ	
ЗВУКОВОГО ВЕЩАНИЯ.....	43
Лабораторна робота 2.1	
СТИСНЕННЯ ДАНИХ БЕЗ ВТРАТ.....	50
Лабораторна робота 2.1	
СЖАТИЕ ДАННЫХ БЕЗ ПОТЕРЬ.....	57
Лабораторна робота 2.2	
АЛГОРИТМ СТИСНЕННЯ ДАНИХ JPEG	65
Лабораторная работа 2.2	
АЛГОРИТМ СЖАТИЯ ДАННЫХ JPEG.....	75
Лабораторна робота 2.3	
ПОРІВНЯННЯ РІЗНИХ ВЕЙВЛЕТ-ПЕРЕТВОРЕНЬ ПРИ СТИСНЕННІ....	84
Лабораторная работа 2.3	
СРАВНЕНИЕ РАЗЛИЧНЫХ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЙ	
ПРИ СЖАТИИ	91
Лабораторна робота 2.4	
ОЧИЩЕННЯ СИГНАЛІВ ВІД ШУМІВ ЗА ДОПОМОГОЮ ВЕЙВЛЕТІВ..	99
Лабораторная работа 2.4	
ОЧИСТКА СИГНАЛОВ ОТ ШУМОВ С ПОМОЩЬЮ ВЕЙВЛЕТОВ.....	104

Лабораторна робота 1.1

ДОСЛІДЖЕННЯ КАНАЛУ ТОНАЛЬНОЇ ЧАСТОТИ

1. Мета роботи

Моделювання і дослідження характеристик каналу тональної частоти (КТЧ).

2. Ключові положення

КТЧ називають комплекс апаратури зв'язку, який призначено для передачі сигналів в смузі частот (0,3 ...3,4) кГц. КТЧ використовують для передачі сигналів телефонного зв'язку та передачі даних.

При проектуванні, настройці та експлуатації якість КТЧ оцінюється електричними характеристиками. Норми на характеристики задаються як на всю магістраль, так і на окремі її частини. Електричні характеристики кабельних та радіорелейних магістралей нормуються для гіпотетичного кола, що має довжину 2500 км.

У даній лабораторній роботі моделюється тільки дві електричні характеристики КТЧ: амплітудно-частотна характеристика та фазочастотна характеристика.

Амплітудно-частотною характеристикою (АЧХ) називають залежність модуля коефіцієнта передачі від частоти передаваного сигналу. Оцінюють нерівномірність АЧХ відносним коефіцієнтом передачі Y , що представляє собою відношення коефіцієнта передачі $K(f)$ на даній частоті до коефіцієнта передачі K_0 на середній частоті. Коефіцієнт передачі часто подають в децибелах

$$Y = 20 \lg \frac{K(f)}{K_0}. \quad (1)$$

Величина Y не повинна виходити за певні межі поля припусків АЧХ.

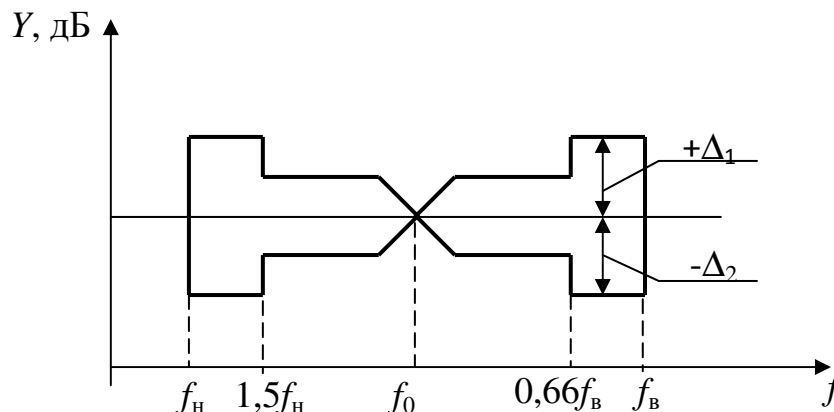


Рисунок 2.1 – Поле припусків амплітудно-частотної характеристики:

f_H, f_B – нижня і відповідно верхня межі номінального діапазону частот;

f_0 – середня частота діапазону

Як правило

$$|+\Delta_1| < |-\Delta_2|. \quad (2)$$

Це зумовлено тим, що людина краще помічає позитивні відхилення АЧХ, ніж негативні.

Якщо характеристика пристрою вміщується у межах поля припусків АЧХ, то такий пристрій відповідає нормам.

Коректування характеристик зводиться до формування АЧХ, які мають обернений вид. Сумарна ідеальна характеристика (пристрій плюс ланка корекції) повинна мати вигляд прямої лінії, паралельної осі частот.

Додаткові спотворення сигналів в системах зв'язку можуть виникати за рахунок нелінійності фазочастотної характеристики (ФЧХ). Практично важко вимірювати ФЧХ систем зв'язку. Тому ФЧХ оцінюють груповим часом проходження, який являє собою похідну від ФЧХ за частотою

$$t_r = \frac{d\phi}{d\omega}. \quad (3)$$

Величина t_r має розмірність часу. Наявність фазочастотних спотворень свідчить про те, що окремі частотні складові сигналу проходять через канал зв'язку через різні проміжки часу. На рис. 2.2 подано нелінійну фазочастотну характеристику.

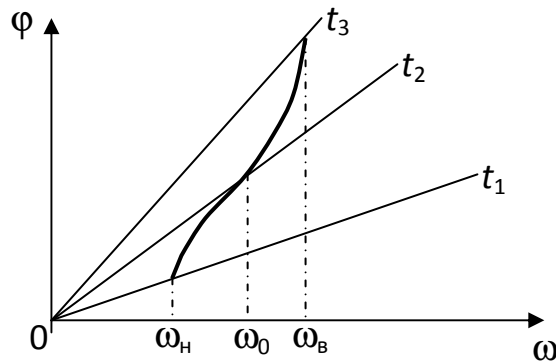


Рисунок 2.2 – Фазочастотна характеристика:
 ω_n , ω_0 , ω_b – нижня, середня та верхня частоти сигналу

Причому

$$t_3 > t_2 > t_1. \quad (4)$$

Нерівність (4) свідчить про те, що нижня частота сигналу проходить каналом зв'язку через найменший проміжок часу, а верхня частота – через найбільший проміжок часу. Форма сигналу на виході каналу зв'язку при цьому змінюється.

Фазові спотворення нормуються для нижніх та верхніх частот номінального діапазону:

$$\left. \begin{aligned} \Delta t_{\text{H}} &= \left| \frac{d\varphi}{d\omega} \right|_{\omega \rightarrow \omega_{\text{H}}} - \left| \frac{d\varphi}{d\omega} \right|_{\omega \rightarrow \omega_0}, \\ \Delta t_{\text{B}} &= \left| \frac{d\varphi}{d\omega} \right|_{\omega \rightarrow \omega_{\text{B}}} - \left| \frac{d\varphi}{d\omega} \right|_{\omega \rightarrow \omega_0} \end{aligned} \right\} \quad (5)$$

Малі фазові спотворення не помічаються слухом людини. Якщо $\Delta t_{\text{B}} > 8$ мс, то фазові спотворення призводять до розмитості мови та зниження розбірливості. Помітність фазових спотворень в області нижніх частот є менша й може сягати величини 80 мс.

В багатоканальних системах передачі фазові спотворення виявляються у вигляді збільшення групового часу проходження на краях діапазону передаваних частот. Якщо канали тональної частоти використовуються для передачі телефонних передач, то фазові коректори при цьому не використовують. Передача дискретних сигналів каналами тональної частоти потребує вмикання фазових коректорів. Оскільки фазочастотні характеристики різних каналів тональної частоти можуть суттєво відрізнятися одна від одної, то це вимагає використання змінних фазових коректорів, характеристики яких можуть бути підлаштовані під характеристику якого завгодно каналу.

Математичне моделювання за допомогою ЕОМ дозволяє проаналізувати параметри окремих вузлів апаратури і їх характеристики без виготовлення макетів апаратури, що значно спрощує процес розробки і дозволяє вибрати оптимальний варіант.

3. Ключові питання

1. Для чого призначений КТЧ?
2. Яка довжина гіпотетичного кола?
3. Що називають амплітудно-частотною характеристикою?
4. Дайте визначення поля припусків.
5. Як проводиться корекція амплітудно-частотних характеристик?
6. Як оцінюють фазочастотні спотворення?
7. Для чого використовують фазові коректори?

4. Домашнє завдання

Підготувати відповіді на ключові питання.

5. Лабораторне завдання

1. Запустити файл CANAL.exe, який є на робочому столі.
2. Ввести дані для роботи програми згідно таблиці 5.1.

Таблица 5.1 – Вхідні дані для однієї та п'яти переприйомних ділянок

№ п/п	Текст запроса программы	Данные, вводимые по запросу программы	
1	Введите тактовую частоту, кГц	8	8
2	Введите номер АЧХ: 1 – хорошая, 2 – плохая	1	1
3	Введите число переприемов	1	5
4	Есть ли колебания АЧХ в полосе пропускания? Да – 1; Нет – 0	0	0
5	Используется ли сплайн-функция? Да – 1; Нет – 0	0	0
6	Используется ли согласованный фильтр? Да – 1; Нет – 0	0	0
7	Введите номер ФЧХ: 1 – хорошая, 2 – плохая, 3 – нулевая	1	1
8	Введите число переприемов по ФЧХ в диапазоне частот ниже 1,9 кГц	1	5
9	Введите число переприемов по ФЧХ в диапазоне частот ниже 1,9 кГц	1	5
10	Записывать на диск частотную характеристику? Нет – 0	0	0
11	Нужно ли усечение отклика? – Да – 1	1	1
12	Введите порог, в дБ	40	40
13	Нужно ли вычисление свертки с косинусным фильтром? Нет – 0	0	0
14	Нужен ли вывод АЧХ и ГВП? Нет – 0	0	0
15	Нужен ли вывод отклика? Нет – 0	0	0
16	Записывать на диск АЧХ? Нет – 0	0	0
17	Записывать на диск ГВП? Нет – 0	0	0
18	Записывать на диск отклик? Да – 1	1	1
19	Введите название файла импульсной реакции	****.ir	****.ir
20	Нормированный отклик? Да – 1; Нет – 0	1	1
21	Записывать на диск ФЧХ? Нет – 0	0	0
22	Выводить на печать параметры канала? Нет – 0	0	0
23	Выводить на печать параметры отклика? Нет – 0	0	0
24	Рисовать ли отклик? Нет – 0	0	0
25	Рисовать ли АЧХ и ГВП? Нет – 0	0	0
26	Продолжать вычисления? Да – 1 (переход к 4 столбцу); Нет – 0 (закончить ввод данных)	1	0

В строчці 19 замість **** ввести чотирьохбуквені назви файлів.

3. Після заповнення четвертого стовпчика закрити «вікно» та відкрити MathCAD.

4. Знайти файли, набрані в пункті 19 табл. 5.1

5. Зарисувати отримані характеристики.

6. Зміст протоколу

1. Характеристики, які отримані при виконанні лабораторного завдання.
2. Висновки по результатам досліджень.

7. Список літератури

1. Фомина Г.Т. Моделирование характеристик канала тональной частоты. Метод.руководство. – Одесса: УГАС. – 2002. – 16 с.
2. Захарченко М.В. Функціональні пристрої телекомунікаційних систем. – Модуль 2: Спряження апаратури з лініями зв'язку, реєстрація, синхронізація, скремблювання та мультиплексування/ М.В. Захарченко, Є.М. Рудий. – Одеса: ВЦ ОНАЗ ім. О.С. Попова. – 2007. – 76 с.

Лабораторная работа 1.1 ИССЛЕДОВАНИЕ КАНАЛА ТОНАЛЬНОЙ ЧАСТОТЫ

1. Цель работы

Моделирование и исследование характеристик канала тональной частоты (КТЧ).

2. Ключевые положения

КТЧ называют комплекс аппаратуры связи, предназначенный для передачи сигналов в полосе частот (0,3 ...3,4) кГц. КТЧ используют для передачи сигналов телефонной связи и передачи данных.

При проектировании, настройке и эксплуатации КТЧ оцениваются электрическими характеристиками. Нормы на характеристики задаются как на всю магистраль, так и на определенные ее части. Электрические характеристики кабельных и радиорелейных магистралей нормируются для гипотетической цепи длиной 2500 км.

В данной лабораторной работе моделируются только две электрические характеристики КТЧ: амплитудно-частотная характеристика и фазочастотная характеристика.

Амплитудно-частотной характеристикой (АЧХ) называют зависимость модуля коэффициента передачи от частоты передаваемого сигнала. Оценивают неравномерность АЧХ относительным коэффициентом передачи Y , который представляет собой отношение коэффициента передачи $K(f)$ на данной частоте к коэффициенту передачи K_0 на средней частоте. Коэффициент передачи часто представляют в децибелах

$$Y = 20 \lg \frac{K(f)}{K_0}. \quad (1)$$

Величина Y не должна выходить за определенные границы поля допусков амплитудно-частотной характеристики.

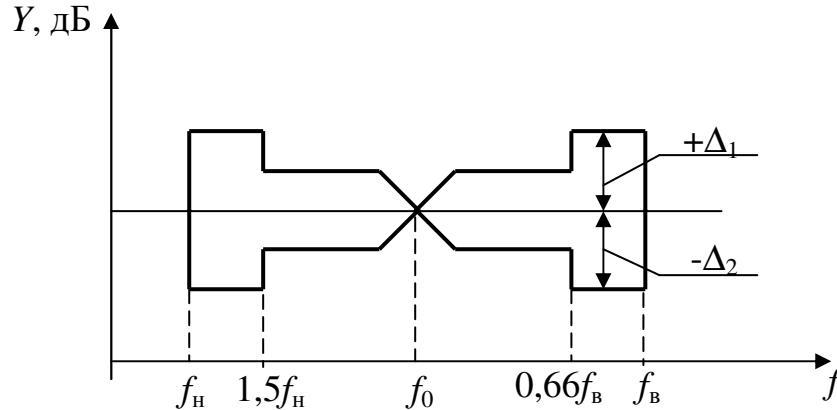


Рисунок 2.1 – Поле допусков амплитудно-частотной характеристики:
 f_n, f_v – нижняя и соответственно верхняя границы номинального диапазона частот;
 f_0 – средняя частота диапазона

Как правило

$$|+\Delta_1| < |-\Delta_2|. \quad (2)$$

Это обусловлено тем, что человек лучше замечает позитивные отклонения АЧХ, чем негативные.

Если характеристика устройства помещается в границах поля допусков АЧХ, то такое устройство соответствует нормам.

Коррекция характеристик сводится к формированию амплитудно-частотных характеристик, которые имеют обратный вид. Суммарная идеальная характеристика (устройство плюс звено коррекции) должна иметь вид прямой линии, параллельной оси частот.

Дополнительные искажения сигналов в системах связи могут возникать за счет нелинейности фазочастотной характеристики (ФЧХ). Практически тяжело измерять ФЧХ систем связи. Поэтому ФЧХ оценивают групповым временем прохождения, которое представляет собой производную от фазочастотной характеристики по частоте

$$t_r = \frac{d\varphi}{d\omega}. \quad (3)$$

Величина t_r имеет размерность времени. Наличие фазочастотных искажений свидетельствует о том, что отдельные частотные составляющие сигнала проходят по каналу связи через разные промежутки времени. На рис. 2.2 представлена нелинейная фазочастотная характеристика.

Причем

$$t_3 > t_2 > t_1. \quad (4)$$

Неравенство (4) свидетельствует о том, что нижняя частота сигнала проходит каналом связи через наименьший промежуток времени, а верхняя частота – через наибольший промежуток времени. Форма сигнала на выходе канала связи при этом изменяется.

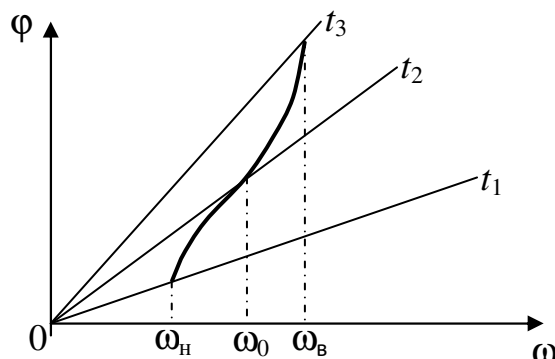


Рисунок 2.2 – Фазочастотная характеристика:
 ω_n , ω_0 , ω_b – нижняя, средняя и верхняя частоты сигнала

Фазовые искажения нормируются для нижних и верхних частот номинального диапазона:

$$\left. \begin{aligned} \Delta t_n &= \left| \frac{d\varphi}{d\omega} \right|_{\omega \rightarrow \omega_n} - \left| \frac{d\varphi}{d\omega} \right|_{\omega \rightarrow \omega_0} \\ \Delta t_b &= \left| \frac{d\varphi}{d\omega} \right|_{\omega \rightarrow \omega_b} - \left| \frac{d\varphi}{d\omega} \right|_{\omega \rightarrow \omega_0} \end{aligned} \right\} \quad (5)$$

Малые фазовые искажения не замечаются слухом человека. Если $\Delta t_b > 8$ мс, то фазовые искажения приводят к размытости речи и снижению разборчивости. Заметность фазовых искажений в области нижних частот меньше и может достигать величины 80 мс.

В многоканальных системах передачи фазовые искажения проявляются в виде увеличения группового времени прохождения на краях диапазона передаваемых частот. Если каналы тональной частоты используются для передачи телефонных передач, то фазовые корректоры при этом не используют. Передача дискретных сигналов каналами тональной частоты требует включения фазовых корректоров. Поскольку фазочастотные характеристики разных каналов тональной частоты могут существенно отличаться одна от другой, то это требует использования сменных фазовых корректоров, характеристики которых могут быть подстроены под характеристику какого угодно канала.

Математическое моделирование при помощи ЭВМ позволяет проанализировать параметры отдельных узлов аппаратуры и их характеристики без изготовления макетов аппаратуры, что значительно упрощает процесс разработки и позволяет выбрать оптимальный вариант.

3. Ключевые вопросы

1. Для чего предназначен КТЧ?
2. Какая длина гипотетической цепи?
3. Что называют амплитудно-частотной характеристикой?
4. Дайте определение поля припусков.
5. Как производится коррекция амплитудно-частотных характеристик?
6. Как оценивают фазочастотные искажения?
7. Для чего используют фазовые корректоры?

4. Домашнее задание

Подготовить ответы на ключевые вопросы.

5. Лабораторное задание

1. Запустить файл CANAL.exe, который есть на рабочем столе.
2. Ввести данные для работы программы согласно табл. 5.1.

Таблица 5.1 – Входные данные для одного и пяти переприемных участков

№ п/п	Текст запроса программы	Данные, вводимые по запросу программы	
1	Введите тактовую частоту, кГц	8	8
2	Введите номер АЧХ: 1 – хорошая, 2 – плохая	1	1
3	Введите число переприемов	1	5
4	Есть ли колебания АЧХ в полосе пропускания? Да – 1; Нет – 0	0	0
5	Используется ли сплайн-функция? Да – 1; Нет – 0	0	0
6	Используется ли согласованный фильтр? Да – 1; Нет – 0	0	0
7	Введите номер ФЧХ: 1 – хорошая, 2 – плохая, 3 – нулевая	1	1
8	Введите число переприемов по ФЧХ в диапазоне частот ниже 1,9 кГц	1	5
9	Введите число переприемов по ФЧХ в диапазоне частот ниже 1,9 кГц	1	5
10	Записывать на диск частотную характеристику? Нет – 0	0	0
11	Нужно ли усечение отклика? – Да – 1	1	1
12	Введите порог, в дБ	40	40
13	Нужно ли вычисление свертки с косинусным фильтром? Нет – 0	0	0
14	Нужен ли вывод АЧХ и ГВП? Нет – 0	0	0
15	Нужен ли вывод отклика? Нет – 0	0	0

16	Записывать на диск АЧХ? Нет – 0	0	0
17	Записывать на диск ГВП? Нет – 0	0	0
18	Записывать на диск отклик? Да – 1	1	1
19	Введите название файла импульсной реакции	****.ir	****.ir
20	Нормированный отклик? Да – 1; Нет – 0	1	1
21	Записывать на диск ФЧХ? Нет – 0	0	0
22	Выводить на печать параметры канала? Нет – 0	0	0
23	Выводить на печать параметры отклика? Нет – 0	0	0
24	Рисовать ли отклик? Нет – 0	0	0
25	Рисовать ли АЧХ и ГВП? Нет – 0	0	0
26	Продолжать вычисления? Да – 1 (переход к 4 столбцу); Нет – 0 (закончить ввод данных)	1	0

В строчке 19 вместо **** ввести четырехбуквенные названия файлов.

3. После заполнения четвертого столбика закрыть «окно» и открыть MathCAD.

4. Найти файлы, набранные в пункте 19 табл. 5.1.

5. Зарисовать полученные характеристики.

6. Содержание протокола

1. Характеристики, которые получены при выполнении лабораторной работы.

2. Выводы по результатам исследований.

7. Список литературы

1. Фомина Г.Т. Моделирование характеристик канала тональной частоты. Метод. руководство. – Одесса: УГАС. – 2002. – 16 с.

2. Захарченко М.В. Функціональні пристрої телекомунікаційних систем. – Модуль 2: Спряження апаратури з лініями зв'язку, реєстрація, синхронізація, скремблювання та мультиплексування/ М.В. Захарченко, Є.М. Рудий. – Одеса: ВЦ ОНАЗ ім. О.С. Попова. – 2007. – 76 с.

Лабораторна робота 1.2 МОДЕЛЮВАННЯ ГАРМОНІЧНОГО ВОКОДЕРА

1. Мета роботи

Одержання знань та вмінь в галузі стиснення мовних сигналів, а також моделювання гармонічного вокодера на ЕОМ, та оцінка якості одержаних після моделювання результатів.

2. Ключові положення

В теперішній час, коли розвиток інформаційних технологій набув шалених темпів, а результати наукової діяльності в сфері зв'язку набувають широкого практичного значення і застосування, дослідження в галузі стиснення мовних сигналів і тепер залишається актуальними. Стиснення мовних сигналів дозволяє підвищити ефективність використання каналу зв'язку і можливе завдяки дискретності слуху людини.

В системах зв'язку використовують вокодери, які дозволяють значно зменшити надмірність сигналу. Вокодери враховують природу мовоутворення. Так, наприклад в російській мові усього 42 фонем. Фонем це найменші звукові одиниці мови. Вони бувають голосні та приголосні. Інколи фонем і букви співпадають, а іноді відрізняються. Це яскраво видно у англійській мові. Людина вимовляє в секунду 10 – 15 фонем. Тобто від центральної нервової системи до мовного апарату сигнали керування передаються зі швидкістю

$$(10 - 15)[\log_2 42] = (60 - 90) \text{ біт/с.}$$

В той же час в цифровій телефонії використовується швидкість 64 кбіт/с. Лінійна модель мовоутворення уявляє мову як систему, яка складається з генератора сигналу збудження та лінійної системи з змінними параметрами, яка ними збуджується. В цій моделі не враховується взаємний вплив голосової щілини та голосового тракту.

Вокодери розрізняють на параметричні і мовоелементні. Робота параметричних вокодерів основана на виділенні параметрів при передачі, які визначають зрозумілість мови і упізнавання голосу абонента, а на прийомі синтезують мовні сигнали, використовуючи виділені параметри сигналу. В мовоелементних вокодерах при передачі розпізнають окремі елементи мови, а при прийомі синтезують ці елементи мови за певними правилами або вибирають їх з пам'яті прийомного пристрою.

Основна задача аналізатора параметричного вокодера полягає в виділенні параметрів, які характеризують спектрально-часову обвідну мовного сигналу і параметрів джерел мовних коливань, які обумовлюють частоту основного тону і зміну її з часом. Також визначають момент появи та закінчення основного тону, шумового сигналу і одинокого імпульсу.

Синтезатор вокодера містить генератори основного тону і шуму, якими керують виділені при передачі параметри. Сигнали цих генераторів подаються на систему, що імітує роботу мовного тракту при вимові звуків. Сигнали генератора основного тону застосовуються при синтезі дзвінких звуків мови. При синтезі глухих звуків мови використовують генератор шуму.

По типу використовуваних параметрів розрізняють смугові, гармонічні, формантні та фонемні вокодери. В смугових вокодерах при передачі визначають складові в вузьких смугах частот, в гармонічних – коефіцієнти Фур'є при зображенні сигналу у вигляді суми гармонік, в формантних – амплітуди і частоти формант, в фонемних – вимовлені звуки.

Вокодери дозволяють зменшити швидкість цифрового потоку від 64000 біт/с до 60 біт/с. Якщо використовувати статистичне кодування то коефіцієнт стиснення збільшується в 2,55 разів. Це було встановлено вченими ОНАЗ ім. О.С. Попова у 1997 році.

Використовуються вони зараз для кодування телефонних сигналів в військових та комерційних цифрових системах зв'язку. Формантні та смугові вокодери знаходять своє застосування також при цифровому зв'язку телефонних сигналів по короткохвильовим каналам радіозв'язку. При цьому якість синтетичної телефонії значно вище якості звичайного однополосного короткохвильового радіотелефонного каналу, так як використання вокодерів практично виключає всі спотворення телефонних сигналів, властивому радіотелефонному каналу.

Використання вокодерів на лініях телефонного зв'язку дозволяє підвищити пропускну спроможність та завадостійкість телефонних ліній на радіохвильових спектрах, а також забезпечує нерозголошеність переговорів. Ці переваги визначають також відповідну економічну користь.

Підвищення пропускну спроможності телефонних ліній зв'язку є звичайним наслідком зменшення смуги частот. Сучасні методи компресії дозволяють підвищити пропускну спроможність в 15 – 50 разів. Правда, при великих значеннях коефіцієнту компресії втрачається природність звучання мови.

Перші вокодери були запропоновані ще на початку минулого століття і будувалися як аналогові пристрої. В 1936 році Дадлі (Dudley) розробив пристрій, який перетворював мову в аналоговий сигнал зі смугою частот шириною 300 Гц. Цей пристрій він назвав вокодером (voice - голос, coder - кодівник), який давав синтезовану мову низької якості.

Як подальший розвиток вокодерної телефонії після смугових вокодерів з'явилися ортогональні (рис. 2.1), в основі яких є розкладання спектру в ортогональний ряд спектральної обвідної сигналу та передачі в канал зв'язку коефіцієнтів розкладання. В аналізаторі миттєвого спектру визначаються коефіцієнти $a_i(t)$ розкладання обвідної миттєвого спектра в заданий ортогональний ряд

$$|S(\omega)| = \sum_{i=0}^n a_i(t) \xi_i(\omega), \quad (1)$$

де $\xi_i(\omega)$ – система використовуємих ортогональних функцій. Сигнали $a_i(t)$ ідуть на систему передачі. Одночасно на систему передачі ідуть і сигнали основного тону та тон/шум. На прийомній стороні після декодування і роз'єднання групового сигналу вони поступають на синтезатор де сигнали ОТ та Т/Ш керують параметрами збудження, а коефіцієнти $a_i(t)$ – формою обвідної спектру.

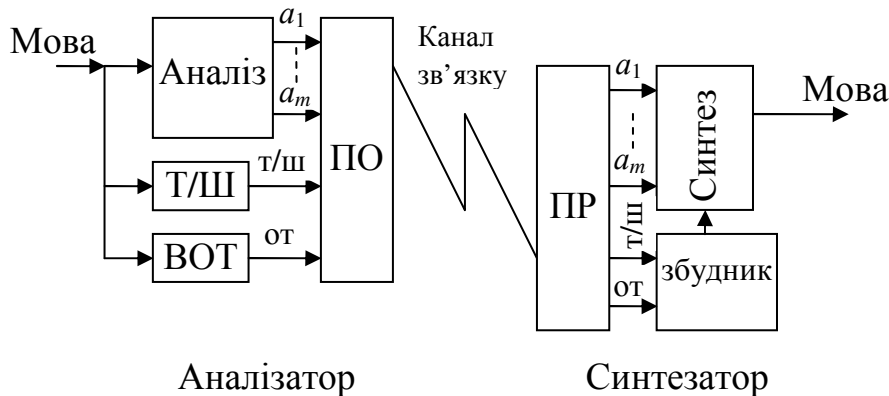


Рисунок 2.1 – Структурна схема ортогонального вокодера:
Т/Ш – виділювач сигналу тон/шум; ВОТ – виділювач сигналу основного тону;
ПО – пристрій об'єднання; ПР – пристрій роз'єднання.

Сучасні вокодери розробляються як цифрові прилади, основані на формальних математичних перетвореннях, наприклад як розкладання у ряд Фур'є, яке здійснює спектральний аналіз мовного сигналу.

Завдяки математичному апарату легко моделювати вокодери на ЕОМ. Існують багато різних алгоритмів стиснення мови, але в даній лабораторній роботі використовується для учбових цілей один, звичайно далеко не кращий, адаптивний алгоритм обробки сигналів, розроблений в нашій академії. Це дасть можливість краще зрозуміти усі особливості моделювання систем стиснення мови. В моделі вокодера стиснення виконується за допомогою дискретного косинусного перетворення Фур'є. Блок відліків X_n в часовій області перетворюється в блок Y_k в частотній області

$$Y_k = \sum_{n=0}^{N-1} X_n \cos \frac{(2n+1)k\pi}{2N}, \quad (2)$$

$$X_n = \frac{1}{N} \sum_{k=0}^{N-1} Y_k \cos \frac{(2n+1)k\pi}{2N}, \quad (3)$$

де $(n, k) = 0, 1, 2, \dots, N-1$, N – число відліків сигналу в одному блоці.

Якщо врахувати кореляційні зв'язки між відліками при передачі відліків сигналів в часовій області то можна зменшити швидкість цифрового потоку по каналу зв'язку. Але величина стиснення в цих системах мала. При передачі відліків сигналу в частотній області можна врахувати дискретність слуху та маскування сигналів і досягти більшого стиснення сигналів.

В учбовому алгоритмі використовується майже миттєве компадування сигналів в частотній області. Звуковий сигнал передається блочно, де в кожному блоці N знаходять найбільший частотний відлік та нормують всі інші відліки, які звичайно менші за найбільший, відносно найбільшого.

Детально алгоритм майже миттєвого компадування описаний в роботі [1].

Для передачі коефіцієнту масштабу квантування використовується m біт.

$$m = \log_2(a - b), \quad (4)$$

де a – число біт, що виділяється для передачі найбільшого відліку в часовій області, b – число біт, яке виділяється для найбільшого частотного відліку.

Коефіцієнт стиснення G можна найти як

$$G = \frac{aN}{m + bN}. \quad (5)$$

3. Ключові питання

1. Для чого потрібно стискати мову?
2. Яким приладом стискають мову людини, і завдяки чому можна стискати мову?
3. Поясніть, що розуміється під поняттям фонем.
4. Які особливості параметричних вокодерів, а які мовоелементних?
5. На які типи розрізняють параметричні вокодери?
6. Завдяки чому можливе моделювання вокодерів на ЕОМ?
7. На скільки сучасні вокодери можуть зменшити швидкість цифрового потоку?
8. Скажіть, яка на Вашу думку, може бути досягнута мінімальна швидкість цифрового потоку при використанні найдосконалішого вокодера?
9. Поясніть принцип дії ортогональних вокодерів.

4. Домашнє завдання

1. Підготувати відповіді на ключові запитання.
2. Підготувати таблицю для дослідження мови.

5. Опис лабораторного макета

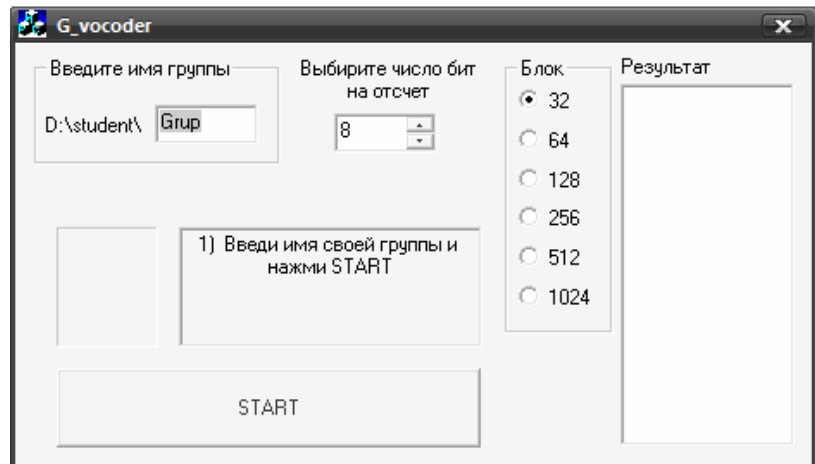
В якості лабораторного макету використовується ЕОМ. Завдяки універсальності сучасних персональних комп'ютерів є можливість моделювати різні прилади програмним шляхом, а не макетуванням. В даній лабораторній роботі використовується модель гармонічного вокодера основана на майже миттєвому компадуванні сигналів в частотній області. Реалізована модель у вигляді консольної програми, написаної на мові високого рівня C++.

Мовний сигнал, використовуємий в даній роботі для дослідження максимального стиснення, при якому не помітні спотворення на слух, записаний у форматі wav. Запис проводився з використанням звукової плати Creative Audigy через програму Microsoft® Звукозапис з частотою дискретизації 8,000 кГц; 16 біт; моно.

6. Лабораторне завдання

Підготовчий етап для виконання роботи:

Запустіть програму G_vocoder через ярлик на робочому столі. При цьому Ви побачите вікно, в якому необхідно ввести ім'я Вашої групи. Тим самим у папці Student на диску D з'явиться папка з введеним ім'ям, в яку будуть записуватися фрагменти мови після обробки в моделі. Далі необхідно при одному з вибраних значень блоку (32 ... 1024) виставити від 16 до 2 кількість біт на відлік. Запускаємо обробку кнопкою START. Для кожного значення блоку повинно вийти близько 14 фрагментів мови з різною кількістю біт на відлік. Формат отриманих фрагментів виглядає наступним чином:



rech	32	_	8	.wav
Ім'я	N – кількість відліків у блоці	розділювач	b – кількість біт на відлік	розширення

Виконання роботи:

Для виконання роботи необхідно прослухати підготовлені фрагменти мови і порівнюючи з оригіналом визначити, є даний фрагмент розбірливим і зрозумілим чи ні. Прослуховувати фрагменти необхідно в наступному порядку: Вибирати фрагмент з мінімальним значенням N і максимальним значенням b (rech32_16.wav). Якщо фрагмент розбірливий, то необхідно перейти до фрагмента зі значеннями N і $b-1$ (rech32_15.wav). Як тільки почуємо мову, яку буде неможливо розібрати, наприклад фрагмент rech32_2.wav, тоді в таблицю заносимо значення b попереднього фрагмента rech32_3.wav, в даному випадку рівне 3, і переходимо до прослуховування фрагментів з максимальним значенням b і збільшуємо кількість відліків в блоці N на 1 (rech64_16.wav). Таким чином, суб'єктивно оцінюючи якість записаних фрагментів, в процесі виконання лабораторної роботи необхідно заповнити в таблиці для кожного блоку значення b .

Далі Вам необхідно буде при різній кількості відліків в блоці N , та різному числу біт для найбільшого частотного відліку b знайти коефіцієнт квантування масштабу m по формулі (4), де a – число біт, що виділяється для передачі найбільшого відліку в часовій області взяти рівному 16. Коефіцієнт стиснення G по формулі (5).

Число біт найбільшого частотного відліку b необхідно зменшувати до появи помітних на слух спотворень в фрагментах мови, що пройшли через

модель. Так як запис фрагменту був зроблений з 16 біт на відлік, то при максимальному значенні $b = 16$ стиснення не буде проводитись.

Результати проведених робіт необхідно занести у табл. 6.1.

Таблиця 6.1 – Дослідження стиснення сигналів мови

N	32	64	128	256	512	1024
b						
m						
G						

7. Зміст протоколу

1. Таблиця з результатами досліджень та обчислень.
2. Висновки по результатам досліджень.

8. Список літератури

1. Рудий Є. М. Стиснення звукових сигналів: Навчальний посіб. – Одеса, УДАЗ, 1996. – 79 с.
2. Система электросвязи. Том второй. Системы радио-, телевизионного вещания и документальной электросвязи. Под ред. Н. В. Захарченко. – К.: Техніка. – 1998. – 239 с.

Лабораторная работа 1.2 МОДЕЛИРОВАНИЕ ГАРМОНИЧЕСКОГО ВОКОДЕРА

1. Цель работы

Получение знаний и умений в области сжатия речевых сигналов, а также моделирование гармонического вокодера на ЭВМ, и оценка качества полученных после моделирования результатов.

2. Ключевые положения

В настоящее время, когда развитие информационных технологий приобрело безумные темпы, а результаты научной деятельности в сфере связи приобретают широкое практическое значение, исследования в отрасли сжатия речевых сигналов, и теперь остаются актуальными. Сжатие речевых сигналов возможно благодаря дискретности слуха человека, позволяет повысить эффективность использования канала связи.

В системах связи используют вокодеры, которые позволяют значительно уменьшить избыточность сигнала. Вокодеры учитывают природу речестроения. Например в русском языке всего 41 фонемы. Фонемы это наименьшие звуковые единицы языка. Они бывают гласные и согласные. Иногда фонемы и буквы

совпадают, а иногда отличаются. Это ярко видно в английском языке. Человек произносит в секунду 10 – 15 фонем. То есть от центральной нервной системы к языковому аппарату сигналы управления передаются со скоростью

$$(10 - 15)[\log_2 42] = (60 - 90) \text{ бит/с.}$$

В то же время в цифровой телефонии используется скорость 64 кбит/с. Линейная модель речеобразования представляет язык как систему, которая состоит из генератора сигнала возбуждения и линейной системы с переменными параметрами, которая ими возбуждается. В этой модели не учитывается взаимное влияние голосовой щели и голосового тракта.

Вокодеры делятся на параметрические и речевые. Работа параметрических вокодеров основана на выделении параметров при передаче, которые определяют разборчивость речи и узнаваемость голоса абонента, а на приеме синтезируют речевые сигналы, используя выделенные параметры сигнала. В речевых вокодерах при передаче распознают отдельные элементы речи, а при приеме синтезируют эти элементы речи по определенным правилам или выбирают их из памяти приемного устройства.

Основная задача анализатора параметрического вокодера заключается в выделении параметров, которые характеризуют спектрально-временную характеристику речевого сигнала и параметров источников речевых колебаний, которые обуславливают частоту основного тона и изменение ее во времени. Также определяют момент появления и окончания основного тона, шумового сигнала и одиночного импульса.

Синтезатор вокодера содержит генераторы основного тона и шума, которыми руководят выделенные при передаче параметры. Сигналы этих генераторов подаются на систему, которая имитирует работу языкового тракта при произношении звуков. Сигналы генератора основного тона применяются при синтезе звонких звуков речи. При синтезе глухих звуков речи используют генератор шума.

По типу используемых параметров различают полосовые, гармонические, формантные и фонемные вокодеры. В полосовых вокодерах при передаче определяют параметры в узких полосах частот. В гармонических – коэффициенты Фурье при изображении сигнала в виде суммы гармоник. В формантных – амплитуды и частоты формант. В фонемных – произнесенные звуки.

Вокодеры позволяют уменьшить скорость цифрового потока от 64000 бит/с до 60 бит/с. Если использовать статистическую кодировку то коэффициент сжатия увеличивается в 2,55 раза. Это было установлено учеными ОНАС им. О.П. Попова 1997 году.

Используются вокодеры в настоящий момент для кодировки телефонных сигналов в военных и коммерческих цифровых системах связи. Наибольшее распространение они получили в мобильных телефонах. Формантные и полосовые вокодеры находят свое применение также в цифровой связи телефонных сигналов по коротковолновым каналам радиосвязи. При этом качество синтетической телефонии значительно выше качества обычного однополосного коротковолнового радиотелефонного канала, так как

использование вокодеров практически исключает все искажения телефонных сигналов, свойственному радиотелефонному каналу.

Использование вокодеров на линиях телефонной связи позволяет повысить пропускную способность и помехоустойчивость телефонных линий в радиочастотных спектрах, а также обеспечивают безопасность переговоров. Эти преимущества определяют также соответствующую экономическую пользу.

Повышение пропускной способности телефонных линий связи является обычным следствием уменьшения полосы частот. Современные методы компрессии позволяют повысить пропускную способность в 15 – 50 раз. Правда, при больших значениях коэффициента сжатия теряется естественность звучания речи.

Первые вокодеры были предложены еще в начале прошлого века и строились как аналоговые устройства. В 1936 году Дадли (Dudley) разработал устройство, которое превращало речь в аналоговый сигнал с шириной полосы частот 300 Гц. Это устройство он назвал вокодером (voice – голос, coder – кодер), который давал синтезированный язык низкого качества.

Как дальнейшее развитие вокодерной телефонии после полосовых вокодеров появились ортогональные (рис. 2.1), в основе которых стоит разложение спектра в ортогональный ряд спектральной характеристики сигнала и передаче в канал связи коэффициентов разложения. В анализаторе мгновенного спектра определяются коэффициенты разложения характеристики мгновенного спектра в заданный ортогональный ряд

$$|S(\omega)| = \sum_{i=0}^n a_i(t) \xi_i(\omega), \quad (1)$$

где $\xi_i(\omega)$ – система используемых ортогональных функций. Сигналы $a_i(t)$ поступают в систему передачи.

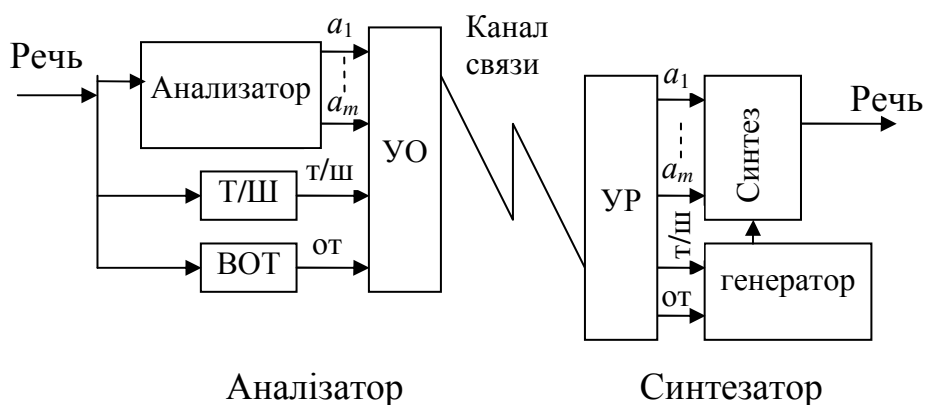


Рисунок 2.1 – Структурная схема ортогонального вокодера:
Т/Ш – выделитель сигнала тон/шум; ВОТ – выделитель сигнала основного тона;
УО – устройство объединения; УР – устройство разделения.

Одновременно в систему передачи поступают и сигналы основного тона и тон/шума. На приемной стороне после декодирования и разделения группового сигнала они поступают на синтезатор где сигналы ОТ та Т/Ш управляют параметрами возбуждения, а коэффициенты $a_i(t)$ – формой характеристики спектра.

Современные вокодеры разрабатываются как цифровые приборы, основанные на формальных математических преобразованиях, например как разложение в ряд Фурье, которое осуществляет спектральный анализ речевого сигнала.

Благодаря математическому аппарату легко моделировать вокодеры на ЭВМ. Существуют много разных алгоритмов сжатия речи, но в данной лабораторной работе используется для учебных целей один, обычно далеко не лучший, адаптивный алгоритм обработки сигналов, разработанный в нашей академии. Это дает возможность лучше понять все особенности моделирования систем сжатия речи. В модели вокодера сжатие выполняется с помощью дискретного косинусного преобразования Фурье. Блок отсчетов X_n во временной области преобразуется в блок Y_k в частотной области

$$Y_k = \sum_{n=0}^{N-1} X_n \cos \frac{(2n+1)k\pi}{2N}, \quad (2)$$

$$X_n = \frac{1}{N} \sum_{k=0}^{N-1} Y_k \cos \frac{(2n+1)k\pi}{2N}, \quad (3)$$

где $(n, k) = 0, 1, 2, \dots, N-1$, N – число отсчетов сигнала в одном блоке.

Если учесть корреляционные связи между отсчетами при передаче отсчетов сигналов во временной области то можно уменьшить скорость цифрового потока по каналу связи. Но величина сжатия в этих системах незначительна. При передаче отсчетов сигнала в частотной области можно учесть дискретность слуха и маскировку сигналов и достичь большего сжатия сигналов.

В учебном алгоритме используется почти мгновенное компадирование сигналов в частотной области. Речевой сигнал передается блоками, где в каждом блоке N находят наибольший частотный отсчет и нормируют все другие отсчеты, которые обычно меньше наибольшего, относительно наибольшего.

Детально алгоритм почти мгновенного компандирования описан в работе [1].

Для передачи коэффициента масштаба квантования используется m бит.

$$m = \log_2(a - b), \quad (4)$$

где a – число бит, которое выделяется для передачи наибольшего отсчета во временной области, b – число бит, которое выделяется для наибольшего частотного отсчета.

Коэффициент сжатия G можно определить как

$$G = \frac{aN}{m + bN}. \quad (5)$$

3. Ключевые вопросы

1. Для чего необходимо сжимать речь?
2. Каким устройством сжимают речь и благодаря чему можно сжимать её?
3. Поясните, что подразумевается под понятием фонема.
4. Какие особенности параметрических вокодеров, а какие речевых элементов?
5. На какие типы разделяются параметрические вокодеры?
6. Благодаря чему возможно моделирование вокодеров на ЭВМ?
7. Во сколько раз современные вокодеры могут уменьшить скорость цифрового потока?
8. Скажите, какая по Вашему мнению, может быть достигнута минимальная скорость цифрового потока при использовании совершенного вокодера?
9. Объясните принцип работы ортогональных вокодеров.

4. Домашнее задание

1. Подготовить ответы на ключевые вопросы.
2. Подготовить таблицу для исследования сжатия речи.

5. Описание лабораторного макета

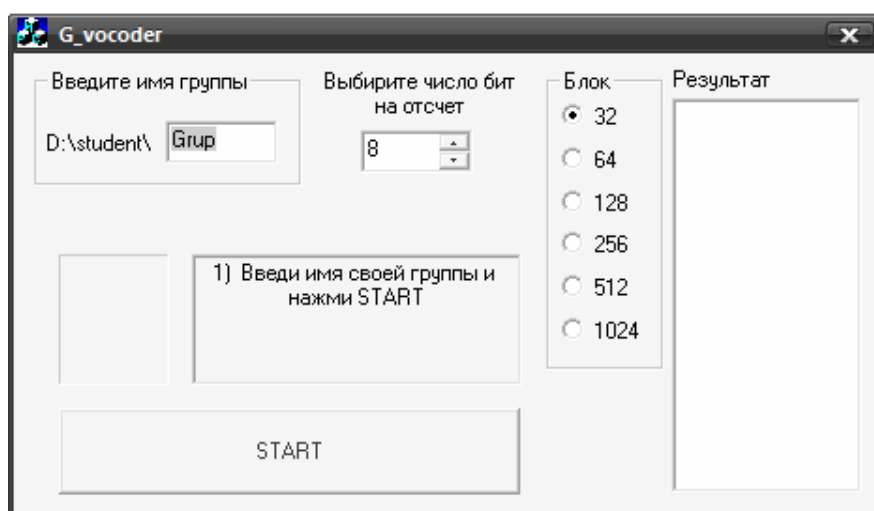
В качестве лабораторного макета используется ЭВМ. Благодаря универсальности современных персональных компьютеров есть возможность моделировать различные устройства программным путем, а не макетированием. В данной лабораторной работе используется модель гармонического вокодера, основанная на почти мгновенном компадировании сигналов в частотной области. Реализованная модель в виде консольной программы, написанной на языке высокого уровня C++.

Речевой сигнал, используемый в данной работе для исследования максимального сжатия, при котором не заметны искажения на слух, записанные в формате wav. Запись проводилась с использованием звуковой платы Creative Audigy через программу Microsoft® Звукозапись с частотой дискретизации 8000 Гц; 16 бит; моно.

6. Лабораторное задание

Подготовительный этап для выполнения работы: Запустите программу G_vocoder через ярлык на рабочем столе. При этом Вы увидите окно, в котором необходимо ввести имя Вашей группы. Тем самым в папке Student на диске D появится папка с введенным именем, в которую будут записываться фрагменты речи после обработки в модели. Далее необходимо при одном из выбранных значений блока (32 ... 1024) выставить от 16 до 2 количество бит на отсчет.

Запускаем обработку кнопкой START. Для каждого значения блока получится около 14 фрагментов речи с разным количеством бит на отсчет. Формат полученных фрагментов выглядит следующим образом:



rech	32	_	8	.wav
имя	N – количество отсчетов в блоке	разделитель	b – количество бит на отсчет	расширение

Выполнение работы:

Для выполнения работы необходимо прослушать подготовленные фрагменты речи и сравнивая с оригиналом определить, является данный фрагмент разборчивым и понятным или нет. Прослушивать фрагменты необходимо в следующем порядке: Выбирать фрагмент с минимальным значением N и максимальным значением b (rech32_16.wav). Если прослушиваемый фрагмент разборчив, то необходимо перейти к фрагменту со значениями N и $b-1$ (rech32_15.wav). Как только услышим речь, которую будет невозможно разобрать, например фрагмент rech32_2.wav, тогда в таблицу заносим значение b предыдущего фрагмента rech32_3.wav, в данном случае равное 3, и переходим к прослушиванию фрагментов с максимальным значением b и увеличиваем количество отсчетов в блоке N на 1 (rech64_16.wav). Таким образом, субъективно оценивая качество записанных фрагментов, в процессе выполнения лабораторной работы необходимо заполнить в таблице для каждого блока значение b .

Дальше Вам необходимо будет при разном количестве отсчетов в блоке N , и разном числе бит для наибольшего частотного отсчета b найти коэффициент квантования масштаба m по формуле (4), где a – число бит, которые выделяются для передачи наибольшего отсчета во временной области, взять равным 16. Коэффициент сжатия G считать по формуле (5). Так как запись фрагмента была сделана из 16 бит на отсчет, то при максимальном значении $b = 16$ сжатие не будет проводиться.

Результаты проведенных работ необходимо занести в табл. 6.1.

Таблица 6.1 – Исследование сжатия сигналов речи

N	32	64	128	256	512	1024
b						
m						
G						

7. Содержание протокола

1. Таблица с результатами исследований и вычислений.
2. Выводы по результатам исследований.

8. Список литературы

1. Рудый Е. М. Стиснення звукових сигналів: Навчальний посіб. – Одеса, УДАЗ, 1996. – 79 с.
2. Система электросвязи. Том второй. Системы радио-, телевизионного вещания и документальной электросвязи. Под ред. Н. В. Захарченко. – К.: Техніка. – 1998. – 239 с.

Лабораторна робота 1.3 ДОСЛІДЖЕННЯ МОДЕМА З ЕХОКОМПЕНСАТОРОМ

1. Мета роботи

Дослідження модему з ехокомпенсатором по рекомендації V.32 bis.

2. Ключові положення

Для сполучення абонентської двохпровідникової системи передачі з чотирьохпровідниковою застосовують диференційні системи, які виконуються за допомогою або диференційного трансформатора, або резисторів у вигляді мостової схеми. Найчастіше в апаратурі використовують диференційні системи на трансформаторах (рис. 2.1).

До точок 1-1 підключають абонентську лінію, в якій передача сигналів організовується в обох напрямках. Z_1 – опір абонентської лінії; Z_2 – опір балансного контура, який підмикається до точок 2-2.

Тракт передачі чотирьохпровідникової системи підмикають до точок 3-3, а тракт прийому – до точок 4-4. Диференційна система вважається врівноваженою, якщо згасання сигналів у напрямку від точок 4-4 до точок 3-3 є нескінченно велике.

Найчастіше використовують диференційні трансформатори з однаковими півобмотками, тобто кількість витків ω в них є однаковими.

$$\omega_1 = \omega_2 . \quad (1)$$

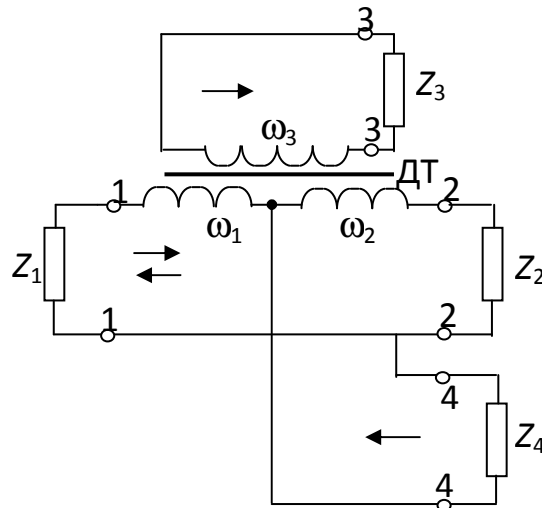


Рисунок 2.1 – Диференційна система на трансформаторі

Задля простоти аналізу вважається, що диференційний трансформатор (ДТ) є ідеальний, тобто активний опір обмоток дорівнює нулю, індуктивності обмоток є нескінченно великі, а розсіяння немає. Якщо диференційна система врівноважена, то

$$Z_1 = Z_2, \quad (2)$$

$$Z_3 = \frac{Z_1 + Z_2}{n^2}, \quad (3)$$

де n – коефіцієнт трансформації трансформатора,

$$n = \frac{\omega_1 + \omega_2}{\omega_3}, \quad (4)$$

а ω_3 – кількість витків третьої обмотки.

При цьому величина опору

$$Z_4 = 0,5 Z_1, \quad (5)$$

а струми, які протікають по півобмоткам є однаковими.

При цьому магнітні потоки, що виникли в півобмотках, є однакові за величиною, але спрямовані назустріч один до одного. Тому результуючий потік в осерді диференційного трансформатора дорівнює нулю, а електрорушійна сила у третій обмотці не наводиться і сигнал від точок 4-4 до точок 3-3 не проходить, тобто згасання сигналу дорівнює нескінченності.

Згасання між точками 4-4 і точками 1-1 буде рівно 3 дБ. Таке ж згасання буде між точками 1-1 і точками 4-4, оскільки схема зворотна.

Якщо виконуються вирази (2) і (3), то згасання між точками 3-3 і 1-1 теж дорівнює 3 дБ.

Згасання між точками 4-4 і 3-3

$$a_{43} = 6 + 20 \lg \left| \frac{Z_1 + Z_2}{Z_1 - Z_2} \right|. \quad (6)$$

Неузгодженість в точках підмикання абонентної лінії до чотирьохпровідникового телефонного каналу призводить до з'яви електричного еха. Суть з'яви еха (рис. 2.2) полягає в тому, що частина мовного сигналу абонента А надходить до нього ж, завдяки диференційній системі.

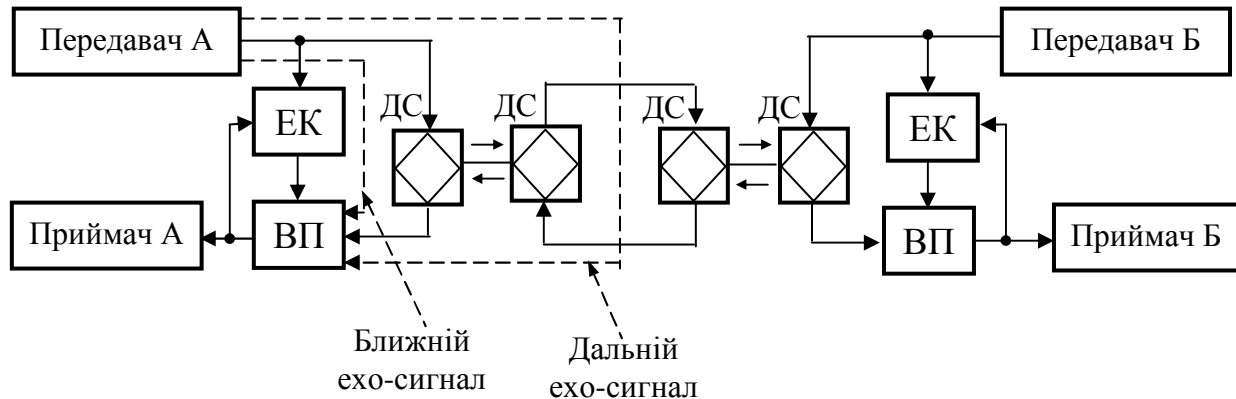


Рисунок 2.2 – Структурна схема зв'язку з ехокомпенсатором:
ЕК – ехокомпенсатор; ВП – вхідний пристрій; ДС – диференційна система

Ближній ехо-сигнал надходить до абонента А через найближчу до нього диференційну систему. Дальній ехо-сигнал проходить через більш віддалену диференційну систему. Той абонент, котрий говорить, почує свою мову через певний час, величина якого залежить від групового часу проходження сигналу каналом зв'язку. Цей сигнал називають перше ехо мовця. Далі сигнал ехо може знову потрапити через диференційну систему, розташовану на початку чотирьохпровідникового тракту, до напрямку передачі А-Б і далі до напрямку Б-А і мовець почує друге ехо мовця тощо.

Коли перше ехо мовця через диференційну систему і напрямок передачі А-Б досягне вуха слухача, останній почує перше ехо слухача. Потім слухач може почути друге ехо слухача і тощо до повного згасання сигналів мови.

Ехо мовця сприймається ним як перепитування. Ехо слухача знижує розбірливість мови.

Завадна дія ехо тим більша, чим менше є згасання ехо і більший час проходження сигналів каналом зв'язку. Час поширювання сигналів ехо дорівнює подвоєному часу проходження сигналів в одному напрямку каналом зв'язку. Окрім лінії зв'язку, певний груповий час проходження вносить кінцева та проміжна апаратура, тому що в ній є реактивні елементи (конденсатори, котушки індуктивності). Цей час може становити 10...20% групового часу проходження сигналів лінією зв'язку.

Ехо може бути помітним за слухового сприйняття, а може бути й непомітним. На рис. 2.3 наведено графік помітності ехо.

Величина a_e схарактеризує згасання ехо на шляху поширювання сигналів. Чим більший інтервал Δt , тим більша буде помітність еха:

$$\Delta t = t_e - t_c, \quad (7)$$

де t_e , t_c – груповий час проходження сигналів ехо і основного сигналу апаратурою та трактами зв'язку.

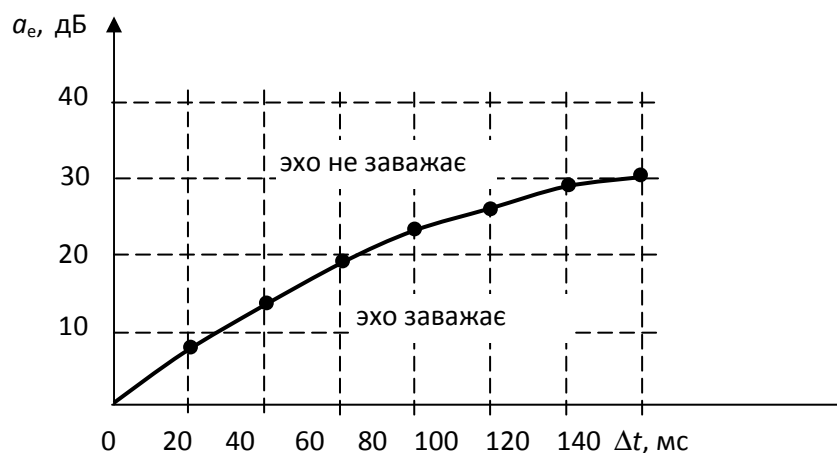


Рисунок 2.3 – Помітність ехо

В середньому величина a_e становить близько 19 дБ, якщо канал зв'язку нормально відрегульовано. З рис. 2.3 видно, що сигнали ехо можна не зважати, якщо час поширювання сигналів в одному напрямку є менше за 30 мс. Якщо цей час є більший, то треба збільшувати згасання сигналів ехо. Однак це приводить до зниження якості зв'язку. Тому більш доцільно використовувати ехокомпенсатори.

Модель ехокомпенсації подано на рис. 2.4, де $b(t)$ – сигнал далекого передавача, $a(t)$ – сигнал ближнього передавача.

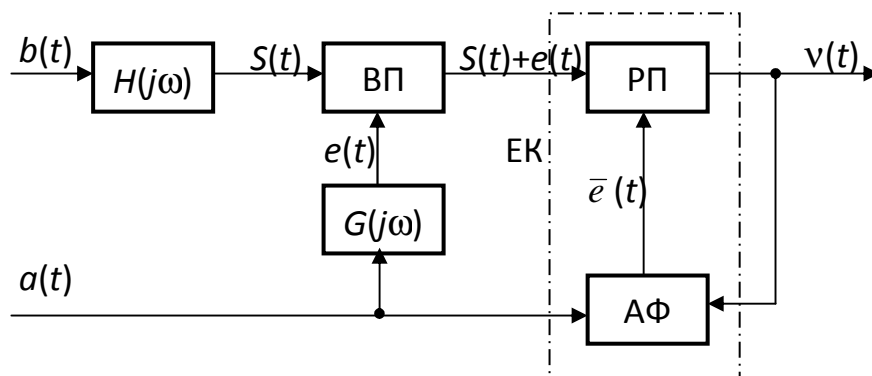


Рисунок 2.4 – Модель ехокомпенсації:

ВП – вхідний пристрій, РП – розв'язувальний пристрій, АФ – адаптивний фільтр, ЕК – ехокомпенсатор

На вхідний пристрій приймача надходять сигнали $S(t)$ далекого передавача після проходження через канал зв'язку з передаточною функцією $H(j\omega)$. Також на вхідний пристрій проходить сигнал ехо $e(t)$ від ближнього передавача. Сигнал ехо проходить через еквівалентний тракт з періодичною функцією $G(j\omega)$.

Завдання ехокомпенсатора полягає в тому, щоби при прийманні сигналу далекого передавача скомпенсувати сигнал ехо. Якщо передавальна функція

$G(j\omega)$ відома, то завдання спрощується, оскільки необхідно тільки розробити пристрій, який перетворює сигнал $a(t)$ на сигнал $\bar{e}(t)$, тобто сигнал, який компенсує ехо $e(t)$.

У більшості випадків передавальна функція $G(j\omega)$ відома лише частково і є необхідність використання адаптивного фільтра, який продукує сигнал $\bar{e}(t)$. На один вхід розв'язувального пристрою ехокомпенсатора подається сигнал $S(t) + e(t)$, а на другий вхід – сигнал $\bar{e}(t)$. На виході розв'язувального пристрою буде

$$v(t) = S(t) + e(t) - \bar{e}(t). \quad (8)$$

Настроювання ехокомпенсатора відбувається за мінімумом середньоквадратичної помилки. Адаптивний фільтр настраюється в такий спосіб, щоби сигнал $v(t)$ найбільшою мірою відповідав сигналові $S(t)$. Потужність сигналів на виході ехокомпенсатора при цьому буде мінімальною.

Піднесемо сигнал $v(t)$ до квадрата

$$v^2(t) = S^2(t) + [e(t) - \bar{e}(t)]^2 + 2S(t)[e(t) - \bar{e}(t)]. \quad (9)$$

Сигнал $S(t)$ та $[e(t) - \bar{e}(t)]$ є некорельовані. Тому математичне сподівання $m[v^2(t)]$ буде

$$m[v^2(t)] = m[S^2(t)] + m[e(t) - \bar{e}(t)]^2, \quad (10)$$

оскільки

$$m[2S(t) e(t) - \bar{e}(t)] \rightarrow 0. \quad (11)$$

При настроюванні величина $m[S^2(t)]$ не змінюється і

$$\min m[v^2(t)] = m[S^2(t)] + \min m[e(t) - \bar{e}(t)]^2. \quad (12)$$

Величина $\min m[v^2(t)]$ відповідає $\min m[e(t) - \bar{e}(t)]^2$, а потужність сигналів на виході ехокомпенсатора в цьому разі буде мінімальною. За ідеального настроювання

$$\min m[e(t) - \bar{e}(t)]^2 \rightarrow 0. \quad (13)$$

На рис. 2.5 наведено схему ехокомпенсатора на базі трансверсального фільтра.

Ланки затримки Δt обирають за формулою

$$\Delta t = \frac{1}{2f_B}, \quad (14)$$

де f_B – верхня частота передаваного сигналу.

Адаптивні коефіцієнти $C_i(K)$ обирають в такий спосіб, щоби потужність сигналу $v(t)$ на виході розв'язувального пристрою на K -тому кроці роботи була мінімальною. Це станеться тоді, коли компенсуючий сигнал $\bar{e}(K)$ за величиною дорівнюватиме $e(K)$.

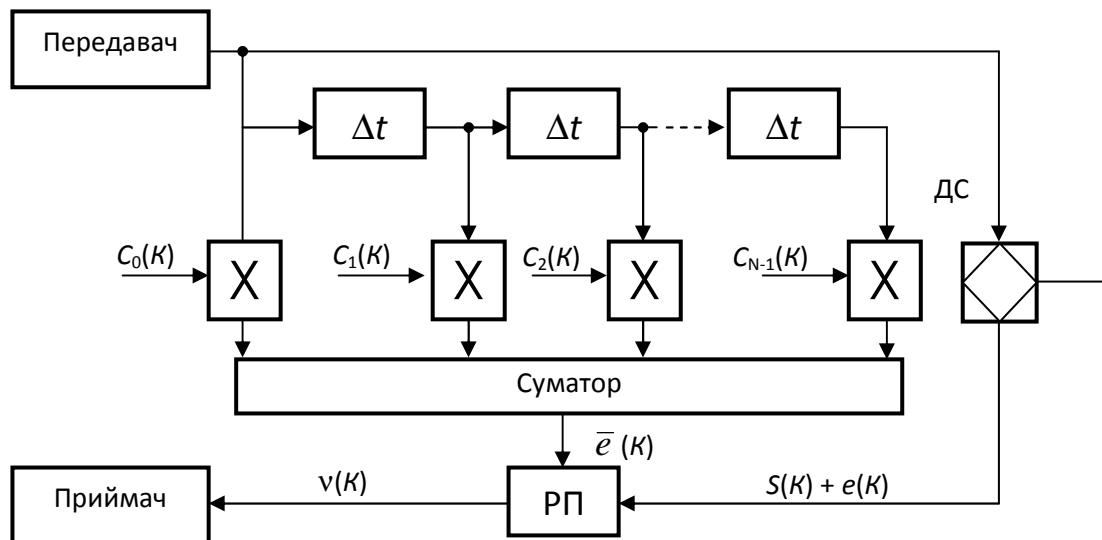


Рисунок 2.5 – Ехокомпенсатор на базі трансверсального фільтра:
РП – розв’язувальний пристрій, ДС – диференційна система, X – перемножувачі

3. Ключові питання

1. Для чого використовують диференційні системи?
2. Яка диференційна система вважається врівноваженою?
3. Як з'являється ближній ехо-сигнал?
4. Як з'являється дальній ехо-сигнал?
5. Як виникає перше та друге ехо мовця?
6. Як виникає перше та друге ехо слухача?
7. Яку величину в середньому складає згасання ехо на шляху поширення сигналів?
8. Для чого призначений ехокомпенсатор?
9. Як обирають коефіцієнти трансверсального фільтра?

4. Домашнє завдання

1. Підготувати відповіді на ключові питання.
2. Підготувати таблицю для дослідження демонстраційної моделі модему.

5. Опис лабораторного макету і лабораторне завдання

В лабораторному макеті використана реальна модель модема з ехокомпенсатором по Рекомендації V.32 bis. Передавач генерує псевдовипадкову послідовність, а приймач порівнює її з еталонною послідовністю.

На робочому столі знайти файл V32bis.exe і запустити його.

При натисненні клавіші F4 видно сигнальне сузір'я. Білі крапки – прийняті сигнали, червоні крапки – еталонні сигнали.

Клавіша F2 – динамічні параметри даних.

Клавіша F3 – параметри каналу.

Клавіша «S» – вмикає і вимикає покроковий режим.

Клавіша «Z» – обнуляє лічильник.

Для зміни параметрів каналу використовуйте клавіші «TAB», «↑», «↓».

Таблиця 5.1 – Результати моделювання

Вікно	Параметр	Величини параметрів		
		1 варіант	2 варіант	3 варіант
F3	Рівень шуму (Noise Level), дБ	–60/–60	– ** /60	–60/–60
F3	Рівень ближнього ехо (Local Echo Level), дБ	–10	–10	**
F2	Число трактів (Clock Counter)	<50000	<50000	<50000
F2	Коефіцієнт помилок (Bit Error Ratio)	0.00/0.00		
F2	Відношення сигнал/шум (R × SNR), дБ	39,4/39,6		
F2	Рівень подавленого ехо (Echo Reject, Lev.), дБ	–64,9/–64,8		

Спочатку виконують перший варіант. Потім змінюють рівень шуму кроками через 10 дБ до тих пір, поки не з'являться помилки.

Обнуляють лічильник і, коли число тактів стане більше 50000, заповнюють таблицю 5.1 (другий варіант).

Відновлюють рівень шуму і кроками через 1 дБ змінюють рівень ближнього ехо, поки не з'являться помилки. Як і раніше, обнуляють лічильник і заповнюють таблицю. Кількість кроків повинна бути не менше 50000 (третій варіант).

6. Зміст протоколу

1. Таблиці з результатами досліджень.
2. Висновки по роботі.

7. Список літератури

1. Фомина Г.Т. Исследование демонстрационной модели модема по Рекомендации V.32bis с эхокомпенсационным разделением передачи и приема. Метод. руководство. – Одесса: УГАС. – 2002. – 15 с.
2. Захарченко М.В. Функціональні пристрої телекомунікаційних систем. – Модуль 2: Спряження апаратури з лініями зв'язку, реєстрація, синхронізація, скремблювання та мультимплексування / М.В. Захарченко, Є.М. Рудий. – Одеса: ВЦ ОНАЗ ім. О.С. Попова. – 2007. – 76 с.

Лабораторная работа 1.3 ИССЛЕДОВАНИЕ МОДЕМА С ЭХОКОМПЕНСАТОРОМ

1. Цель работы

Исследование модема с эхокомпенсатором по рекомендации V.32 bis.

2. Ключевые положения

Для соединения абонентской двухпроводной системы передачи с четырехпроводной применяют дифференциальные системы, которые выполняются при помощи или дифференциального трансформатора, или резисторов в виде мостовой схемы. Чаще в аппаратуре используются дифференциальные системы на трансформаторах (рис. 2.1).

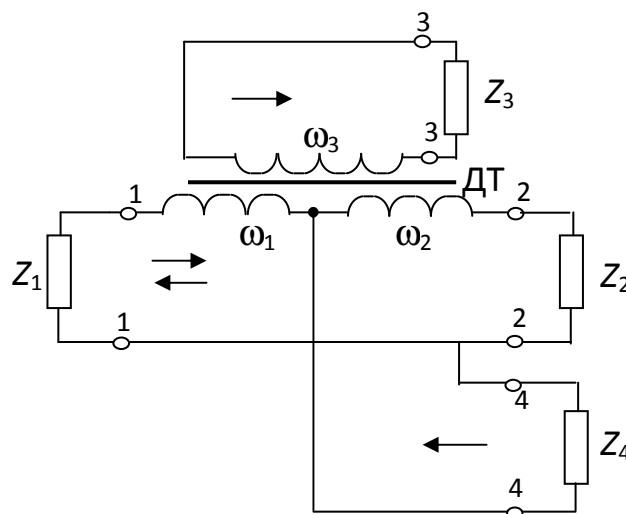


Рисунок 2.1 – Дифференциальная система на трансформаторе

К точкам 1-1 подключают абонентскую линию, в которой передача сигналов организуется в обоих направлениях. Z_1 – сопротивление абонентской линии; Z_2 – сопротивление балансного контура, который подключается к точкам 2-2.

Тракт передачи четырехпроводной системы подключают к точкам 3-3, а тракт приема – к точкам 4-4. Дифференциальная система считается уравновешенной, если затухание сигналов в направлении от точек 4-4 к точкам 3-3 есть бесконечно большое.

Чаще используют дифференциальные трансформаторы с одинаковыми полуобмотками, то есть количество витков ω в них одинаково.

$$\omega_1 = \omega_2. \quad (1)$$

Для простоты анализа считается, что дифференциальный трансформатор (ДТ) идеальный, то есть активное сопротивление обмоток равняется нулю, индуктивности обмоток бесконечно велики, а рассеивания нет. Если дифференциальная система уравновешена, то

$$Z_1 = Z_2, \quad (2)$$

$$Z_3 = \frac{Z_1 + Z_2}{n^2}, \quad (3)$$

где n – коэффициент трансформации трансформатора,

$$n = \frac{\omega_1 + \omega_2}{\omega_3}, \quad (4)$$

а ω_3 – количество витков третьей обмотки.

При этом величина сопротивления

$$Z_4 = 0,5 Z_1, \quad (5)$$

а токи, протекающие по полуобмоткам одинаковы.

При этом магнитные потоки, возникшие в полуобмотках, одинаковы по величине, но направлены навстречу друг к другу. Поэтому результирующий поток в сердечнике дифференциального трансформатора равен нулю, а электродвижущая сила в третьей обмотке не наводится и сигнал от точек 4-4 к точкам 3-3 не проходит, то есть затухание сигнала равно бесконечности.

Затухание между точками 4-4 и точками 1-1 будет равно 3 дБ. Такое же затухание будет между точками 1-1 и точками 4-4, поскольку схема обратима.

Если выполняются выражения (2) и (3), то затухание между точками 3-3 и 1-1 тоже равно 3 дБ.

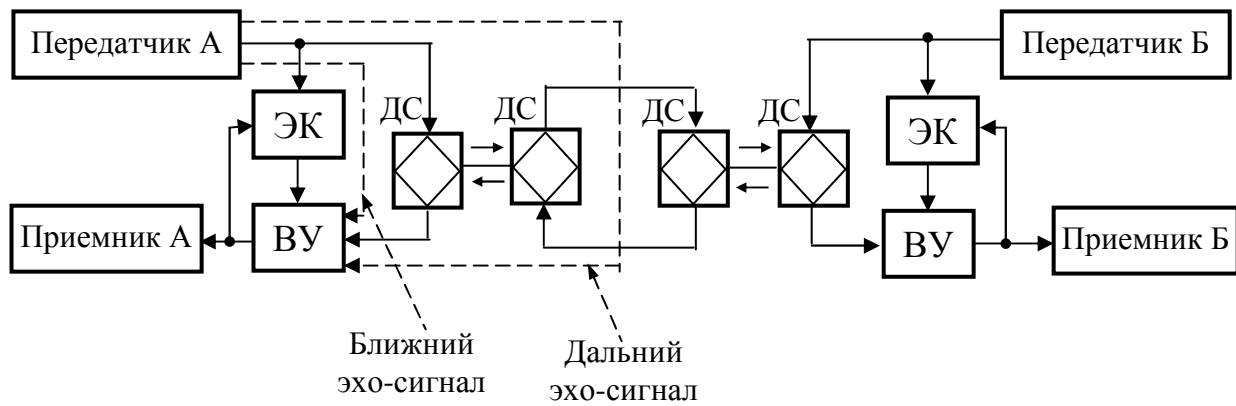
Затухание между точками 4-4 и 3-3

$$a_{43} = 6 + 20 \lg \left| \frac{Z_1 + Z_2}{Z_1 - Z_2} \right|. \quad (6)$$

Несогласованность в точках подключения абонентской линии к четырехпроводному телефонному каналу приводит к появлению электрического эха. Суть появления эха (рис. 2.2) состоит в том, что часть речевого сигнала абонента А приходит к нему же, благодаря дифференциальной системе

Ближайший эхо-сигнал поступает к абоненту А через наиболее близкую к нему дифференциальную систему. Дальний эхо-сигнал проходит через более отдаленную дифференциальную систему. Тот из абонентов, который говорит, услышит свою речь через определенное время, величина которого зависит от группового времени прохождения сигнала каналом связи. Этот сигнал называют первое эхо говорящего. Далее сигнал эха может снова попасть через дифференциальную систему, размещенную в начале четырехпроводного тракта, по направлению передачи А-Б и дальше по направлению Б-А и говорящий услышит второе эхо говорящего.

Когда первое эхо говорящего по дифференциальной системе и напряжение передачи А-Б достигнет уха слушателя, последний услышит первое



эхо слушателя. Затем слушатель может услышать второе эхо слушателя и так до полного затухания сигналов речи.

Рисунок 2.2 – Структурная схема связи с эхокомпенсатором:

ЭК – эхокомпенсатор; ВУ – входное устройство; ДС – дифференциальная система

Эхо говорящего воспринимается им как переспрашивание. Эхо слушателя снижает разборчивость речи.

Мешающее действие эха тем больше, чем меньше затухание эха и больше время прохождения сигналов каналом связи. Время распространения сигналов эха равняется удвоенному времени прохождения сигналов в одном направлении каналом связи. Кроме линии связи, определенное групповое время прохождения вносит оконечная и промежуточная аппаратура, потому что в ней присутствуют реактивные элементы (конденсаторы, катушки индуктивности). Это время может составлять 10...20% группового времени прохождения сигналов линий связи.

Эхо может быть заметным по слуховому восприятию, а может быть и незаметным. На рис. 2.3 приведен график заметности эха.

Величина a_3 характеризует затухание эхо на пути распространения сигналов. Чем больше интервал Δt , тем будет больше заметность эхо:

$$\Delta t = t_3 - t_c, \quad (7)$$

где t_3 , t_c – групповое время прохождения сигналов эхо и основного сигнала аппаратурой и трактами связи.

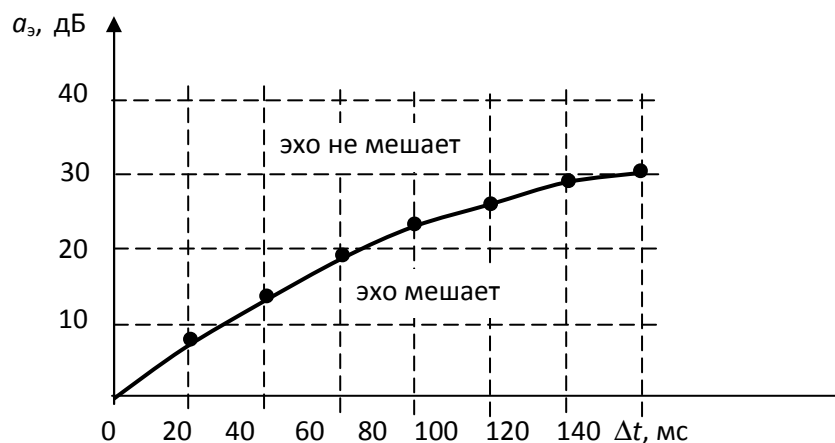


Рисунок 2.3 – Заметность эхо

В среднем величина a_3 составляет около 19 дБ, если канал связи нормально отрегулирован. Из рис. 2.3 видно, что сигналы эхо можно не учитывать, если время распространения сигналов в одном направлении менее 30 мс.

Если это время больше, то необходимо увеличивать затухание сигналов эхо. Однако это приводит к снижению качества звука. Поэтому наиболее целесообразно использовать эхокомпенсаторы.

Модель эхокомпенсации представлена на рис. 2.4, где $b(t)$ – сигнал дальнего передатчика, $a(t)$ – сигнал ближнего передатчика.

На входное устройство передатчика приходит сигнал $S(t)$ дальнего передатчика с передаточной функцией $H(j\omega)$. Также на входное устройство приходит сигнал эхо $e(t)$ от ближнего передатчика. Сигнал эхо проходит через эквивалентный тракт с периодичной функцией $G(j\omega)$.

Задача эхокомпенсатора состоит в том, чтобы при принятии сигнала дальнего передатчика скомпенсировать сигнал эхо. Если передаточная функция $G(j\omega)$ известна, то задача упрощается, поскольку необходимо только разработать устройство, которое преобразует сигнал $a(t)$ на сигнал $\bar{e}(t)$.

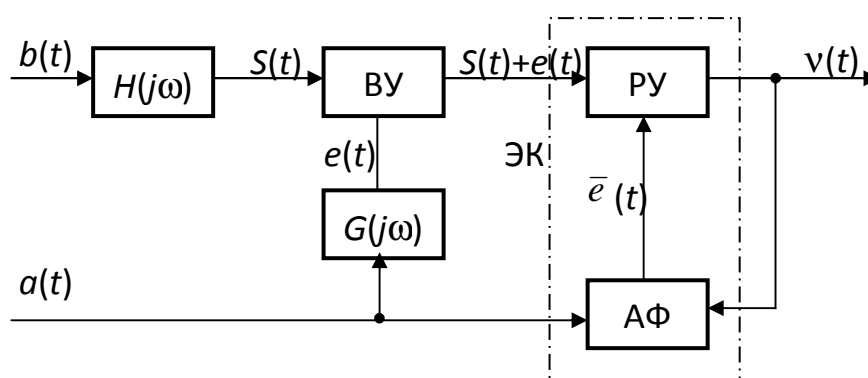


Рисунок 2.4 – Модель эхокомпенсации:

ВУ – входное устройство, РУ – решающее устройство, АФ – адаптивный фильтр, ЭК – эхокомпенсатор

В большинстве случаев передаточная функция $G(j\omega)$ известна лишь частично и существует необходимость использования адаптивного фильтра, который вырабатывает сигнал $\bar{e}(t)$. На один вход решающего устройства эхокомпенсатора подается сигнал $S(t) + e(t)$, а на другой вход сигнал $\bar{e}(t)$. На выходе решающего устройства будет

$$v(t) = S(t) + e(t) - \bar{e}(t). \quad (8)$$

Настройка эхокомпенсатора проводится по минимуму среднеквадратичной ошибки. Адаптивный фильтр настраивается таким образом, что бы сигнал $v(t)$ в наибольшей степени соответствовал сигналу $S(t)$. Мощность сигналов на выходе эхокомпенсатора при этом будет минимальной.

Возведем сигнал $v(t)$ в квадрат

$$v^2(t) = S^2(t) + [e(t) - \bar{e}(t)]^2 + 2S(t)[e(t) - \bar{e}(t)]. \quad (9)$$

Сигнал $S(t)$ и $[e(t) - \bar{e}(t)]$ – некоррелированы.

Поэтому математическое совпадение $m[v^2(t)]$ будет

$$m[v^2(t)] = m[S^2(t)] + m[e(t) - \bar{e}(t)]^2, \quad (10)$$

ПОСКОЛЬКУ

$$m[2S(t) e(t) - \bar{e}(t)] \rightarrow 0. \quad (11)$$

При настройке величина $m[S^2(t)]$ не изменяется и

$$\min m[v^2(t)] = m[S^2(t)] + \min m[e(t) - \bar{e}(t)]^2. \quad (12)$$

Величина $\min m[v^2(t)]$ соответствует $\min m[e(t) - \bar{e}(t)]^2$, а мощность сигналов на выходе эхокомпенсатора в этом случае будет минимальной. При идеальной настройке

$$\min m[e(t) - \bar{e}(t)]^2 \rightarrow 0. \quad (13)$$

На рис. 2.5 приведена схема эхокомпенсатора на трансверсальном фильтре.

Цепь задержки Δt вычисляют по формуле

$$\Delta t = \frac{1}{2f_{\text{R}}}, \quad (14)$$

где f_B – верхняя частота передаваемого сигнала.

Адаптивные коэффициенты $C_i(K)$ выбирают таким образом, чтобы мощность сигнала $v(t)$ на выходе решающего устройства на K -том шаге работы была минимальной. Это произойдет тогда, когда компенсирующий сигнал $\bar{e}(K)$ по величине равен $e(K)$.

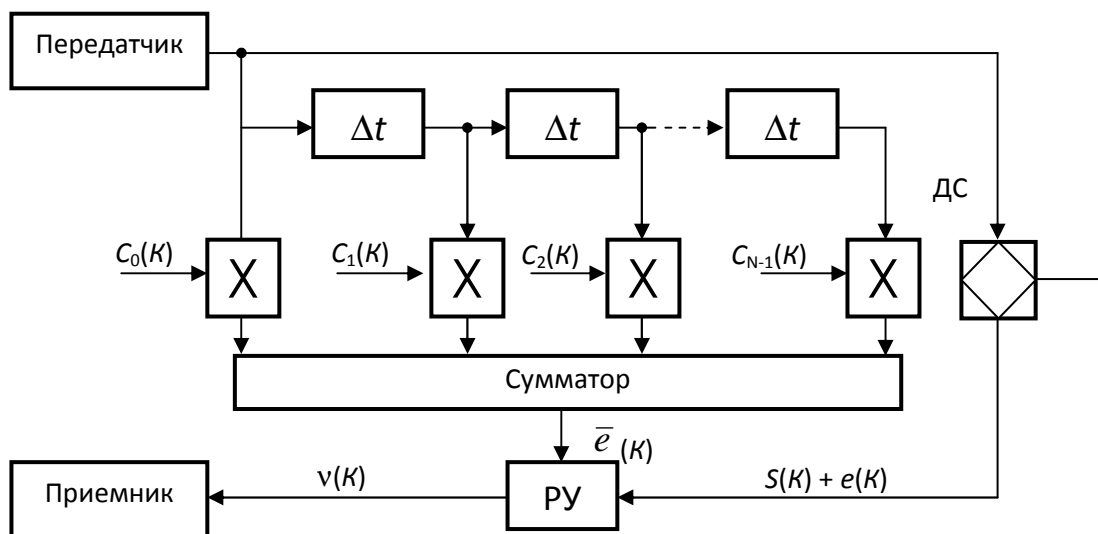


Рисунок 2.5 – Эхокомпенсатор на базе трансверсального фильтра
 РУ – решающее устройство, ДС – дифференциальная система, Х – умножитель

3. Ключевые вопросы

1. Для чего используют дифференциальные системы?
2. Какая дифференциальная система считается уравновешенной?
3. Как появляется ближайший эхо-сигнал?
4. Как появляется дальний эхо-сигнал?
5. Как возникает первое и второе эхо говорящего?
6. Как возникает первое и второе эхо слушателя?
7. Какую величину в среднем составляет затухание эхо на пути распространения сигналов?
8. Для чего предназначен эхокомпенсатор?
9. Как выбирают коэффициенты трансверсального фильтра?

4. Домашнее задание

1. Подготовить ответы на ключевые вопросы.
2. Подготовить таблицу для исследования демонстрационной модели модема.

5. Описание лабораторного макета и лабораторное задание

В лабораторном макете использована реальная модель модема с эхокомпенсатором по Рекомендации V.32bis. Передатчик генерирует псевдослучайную последовательность, а приемник сравнивает ее с эталонной последовательностью.

На рабочем столе найти файл V32bis.exe и запустить его.

При нажатии клавиши F4 видно сигнальное созвездие. Белые точки – принятые сигналы, красные точки – эталонные сигналы. Клавиша F2 – динамические параметры данных. Клавиша F3 – параметры канала. Клавиша «S» – включает и выключает пошаговый режим. Клавиша «Z» – обнуляет счетчик.

Для смены параметров канала используйте клавиши «TAB», «↑», «↓».

Таблица 5.1 – Результаты моделирования

Окно	Параметр	Величины параметров		
		1 вариант	2 вариант	3 вариант
F3	Уровень шума (Noise Level), дБ	–60/–60	– ** /60	–60/–60
F3	Уровень ближнего эхо (Local Echo Level), дБ	–10	–10	**
F2	Число трактов (Clock Counter)	<50000	<50000	<50000
F2	Коэффициент ошибок (Bit Error Ratio)	0.00/0.00		
F2	Отношение сигнал/шум ($R \times SNR$), дБ	39,4/39,6		
F2	Уровень подавленного эхо (Echo Reject, Lev.), дБ	–64,9/–64,8		

Сначала выполняют первый вариант. Потом изменяют уровень шума шагами через 10 дБ до тех пор, пока не появятся ошибки.

Обнуляют счетчик и, когда число тактов станет больше 50000, заполняют таблицу 5.1 (второй вариант).

Восстанавливают уровень шума и шагами через 1 дБ изменяют уровень ближнего эхо, пока не появятся ошибки. Как и ранее, обнуляют счетчик и заполняют таблицу. Количество шагов должно быть не менее 50000 (третий вариант).

6. Содержание протокола

1. Таблицы с результатами исследований
2. Выводы по работе.

7. Список литературы

1. Фомина Г.Т. Исследование демонстрационной модели модема по Рекомендации V. 32bis с эхокомпенсационным разделением передачи и приема. Метод. руководство. – Одесса: УГАС. – 2002. – 15 с.

2. Захарченко М.В. Функціональні пристрої телекомунікаційних систем. – Модуль 2: Спряження апаратури з лініями зв'язку, реєстрація, синхронізація, скремблювання та мультиплексування / М.В. Захарченко, Є.М. Рудий. – Одеса: ВЦ ОНАЗ ім. О.С. Попова. – 2007. – 76 с.

Лабораторна робота 1.4 МОДЕЛЮВАННЯ СИСТЕМИ СТИСНЕННЯ СИГНАЛІВ ЗВУКОВОГО МОВЛЕННЯ

1. Мета роботи

Формування знань та вмінь в галузі стиснення сигналів звукового мовлення, моделювання на ЕОМ структурних схем стиснення сигналів.

2. Ключові положення

Для підвищення ефективності використання каналів зв'язку застосовують різні методи і засоби. Важливу роль відіграють методи скорочення надмірності даних, які дозволяють зменшити обсяг переданої інформації, що еквівалентно підвищенню пропускної спроможності каналу зв'язку. Зменшення надмірності даних називають *стисненням*.

Стиснення завжди супроводжується спотворенням сигналів. Стиснути сигнал без спотворень неможливо. Спотворення можуть бути помітні при

слуховому сприйнятті, а можуть бути й не помітні. Завдяки тому, що існують спотворення не помітні для слуху людини, можливе стиснення сигналів звукового мовлення.

Не можна дозволити, щоб під час передачі програми звукового мовлення були помітні спотворення сигналів, бо слухачі не захочуть користуватись таким мовленням. При розробці систем стиснення враховують властивості вуха людини та характеристики акустичних сигналів.

В системах стиснення враховують дискретність слуху по амплітуді і частоті, та взаємне маскування сигналів. Якщо поступово збільшувати амплітуду сигналів даної частоти, то людині здається, що сигнал збільшується стрибками, а не поступово. В області середніх частот кількість таких стрибків біля 250. На низьких і високих частотах кількість стрибків зменшується до 150.

Дискретність слуху спостерігається не тільки по амплітуді сигналів, але й по частоті. В усьому частотному діапазоні людина сприймає біля 250 градацій частоти, причому кількість градацій залежить від інтенсивності звукових сигналів. При великій інтенсивності сигналів сприймається до 650 градацій частоти.

Границі частотного діапазону слуху людини 20...20000 Гц. У людей, які знаходяться під впливом гучних звуків (трактористи, любителі поп-музики та ін.) слух невинно погіршується і частотний діапазон зменшується. Це є захисною реакцією організму людини на дуже потужний подразник, якими є гучні звукові сигнали. Від впливом гучних звуків поріг гучності підвищується (людина втрачає слух). Тому музику та передачі по радіо і телебаченню слід прослуховувати при мінімально допустимій гучності сигналів.

При одночасній дії на слух двох сигналів один з них може не прослуховуватись в присутності іншого. Це явище називають *маскуванням* сигналів. Маскування залежить від частот і амплітуд сигналів. Експериментально встановлено, що низькочастотні сигнали сильніше маскують високочастотні. Коли частота завади набагато більша ніж корисний сигнал, то маскування менше.

Якщо сигнал маскується вузькосмуговим шумом, то при збільшенні смуги шуму маскування збільшується, тому що енергія маскуючого сигналу зростає зі збільшенням смуги шуму. Але маскування збільшується до тих пір, поки ширина смуги не досягне деякого значення, яке називають *критичною смугою слуху*.

При реалізації систем стиснення сигналів звукового мовлення враховують дискретність слуху людини та маскування в критичних смугах слуху. Досліджують системи стиснення за допомогою ЕОМ. При цьому ЕОМ використовують в якості макета обладнання системи стиснення сигналів.

В ОНАЗ ім. О.С. Попова дослідження по стисненню сигналів звукового мовлення проводяться більше 30 років. Ряд алгоритмів стиснення використані в супутникових системах передачі. Фірми-розробники охоче рекламують свої досягнення, пишуть в загальних рисах про особливості кодування, не розкриваючи технічні деталі обробки сигналів, що не дозволяє повторити

системи стиснення сигналів без купівлі ліцензій або обладнання, яке випускається фірмами.

В даній лабораторній роботі використовується для учбових цілей один, звичайно далеко не кращий, адаптивний алгоритм обробки сигналів, розроблений в нашій академії. В моделі системи стиснення за допомогою дискретного косинусного перетворення Фур'є блок відліків $X(N)$ в часовій області перетворюється в блок $Y(K)$ в частотній області та навпаки.

Розрахунки виконуються за такими формулами:

$$Y(1) = \frac{1}{\sqrt{N1}} \sum_{N=1}^{N1} X(N), \quad (1)$$

$$Y(K) = \sqrt{\frac{2}{N1}} \sum_{N=1}^{N1} X(N) \cos \frac{(2N-1)(K-1)\pi}{2N1}, \quad (2)$$

$$X(1) = \frac{1}{\sqrt{N1}} Y(1), \quad (3)$$

$$X(N) = \sqrt{\frac{2}{N1}} \sum_{K=1}^{N1} Y(K) \cos \frac{(2N-1)(K-1)\pi}{2N1}, \quad (4)$$

де $(N, K) = 1, 2, 3, \dots, N1$. $N1$ – число відліків сигналів в одному блоці в часовій та частотній областях.

Вхідний масив чисел являє собою часові відліки реальних мовних сигналів, які отримані при рівномірному квантуванні. Величина масивів була такою, щоб час звучання фрагментів сигналів, які використовувались при дослідженнях, був 7...10 с. Така невелика протяжність звучання фрагментів вибиралась тому, щоб уникнути заважаючого впливу слухової пам'яті людини при виконанні суб'єктивно-статистичних експертиз.

При передачі відліків сигналів в часовій області можна врахувати кореляційні зв'язки між відліками і зменшити швидкість цифрового потоку по каналу зв'язку. Але величина стиснення в цих системах мала. Якщо передаються відліки сигналів в частотній області, то можна врахувати дискретність слуху та маскування сигналів і досягти більшого стиснення сигналів.

Системи стиснення сигналів звукового мовлення набагато складніші ніж системи стиснення телефонних сигналів, тому що при прослуховуванні передач не повинні бути помітні на слух завади, викликані системами стиснення. Кращі зарубіжні зразки систем стиснення звукового мовлення дозволяють зменшити швидкість цифрового потоку з 768 кбіт/с до 128 кбіт/с, коли мовлення ведеться в смузі частот 20...15000 Гц. Системи стиснення, розроблені в ОНАЗ ім. О.С. Попова, при такій самій смузі частот зменшують швидкість цифрового потоку до 42 кбіт/с.

В учбовому алгоритмі стиснення сигналів звукового мовлення використовується майже миттєве компадування сигналів в частотній області.

При цьому в блоці відліків сигналів в частотній області знаходять найбільший частотний відлік. Всі інші відліки сигналу, які, звичайно, менші за найбільший, нормують відносно найбільшого.

$$Y_{ni} = \frac{Y_i}{Y_1}, \quad (5)$$

де Y_1 – найбільший частотний відлік в блоці; Y_i – величина i -того частотного відліку; Y_{ni} – нормована величина i -того частотного відліку.

При такому підході завжди виконується умова

$$Y_{ni} \leq 1. \quad (6)$$

Після проходження сигналів по каналу зв'язку відновлюють відліки сигналів в блоці в частотній області

$$Y_i = Y_1 \cdot Y_{ni}. \quad (7)$$

На кожний відлік сигналів в часовій області виділяють 16 біт. Кількість біт, яка виділяється на найбільший частотний відлік у блоці, знаходять по результатам суб'єктивно-статистичних експертиз.

Найчастіше використовують п'ятибальну шкалу оцінок. Експерти порівнюють сигнал А на вході системи стиснення з сигналом Б, який оброблений системою стиснення. Як правило, експертів попереджують про те, коли вони слухають сигнал А, а коли – Б.

Для перевірки об'єктивності показань експертів, інколи, за сигналом А повторюють замість сигналу Б сигнал А. Шкала оцінок така:

- 5 – спотворення Б в порівнянні з А непомітні;
- 4 – спотворення Б в порівнянні з А помітні, але прийнятні;
- 3 – спотворення трохи роздратовують, але допустимі;
- 2 – спотворення роздратовують та недопустимі;
- 1 – спотворення надто роздратовуючі та недопустимі.

Бажано при експертизах використовувати так звані критичні фрагменти програм, при звучанні яких спотворення, викликані стисненням сигналів, найбільш помітні.

Встановлено, що в якості критичних фрагментів може використовуватись звучання фортеп'яно без супроводу оркестру, звучання дзвіночків та ксилофона.

Для створюваних цифрових систем передачі рекомендується гранична експертна оцінка 4,5 бали. Тобто, одна половина експертів не помічає спотворення, а друга половина експертів, хоча і помічає спотворення, але вважає їх прийнятними.

Величина стиснення G сигналів звукового мовлення оцінюється як

$$G = \frac{v_1}{v_2}, \quad (8)$$

де v_1, v_2 – швидкість передачі інформації по каналу зв'язку без стиснення і з стисненням інформації.

Числову величину коефіцієнта стиснення можна знайти, уявивши, що всі відліки в блоці дорівнюють максимальному відліку, на який розраховано кодер каналу зв'язку.

В цьому випадку

$$G = \frac{a N_1}{m + b N_1}, \quad (9)$$

де a – число біт, яке необхідне для передачі максимального відліку сигналу в блоці в часовій області; b – число біт, яке виділяється для передачі найбільшого відліку сигналу в блоці в частотній області; m – коефіцієнт масштабу квантування

$$m \leq \log_2 b. \quad (10)$$

Величина m заокруглюється до найближчого більшого цілого числа. $a = 16$ біт. Величина b в бітах задається при прослуховуванні кожного фрагменту програми в процесі виконання суб'єктивно-статистичних експертиз.

Якщо в системі стиснення сигналів використовують статистичне кодування за методами Шеннона-Фано або Хафмена, то величина G збільшується у 2,5 рази.

3. Ключові питання

1. Який процес називають стисненням сигналів?
2. Чому спотворення, викликані системами стиснення сигналів звукового мовлення, не повинні бути помітні вухом людини?
3. Що називають дискретністю слуху людини по амплітуді і частоті?
4. Чому погіршується слух людини при тривалому впливі гучних сигналів?
5. Яке явище називають маскуванням сигналів?
6. Що називають критичною смугою слуху?
7. Які властивості слуху людини використовують при реалізації систем стиснення сигналів?
8. Чому при передачі сигналів в частотній області досягають більшого стиснення?
9. Яка швидкість цифрового потоку необхідна при передачі сигналів звукового мовлення в смузі 20...15000 Гц, якщо використовуються системи стиснення?
10. Як здійснюється майже миттєве компадування сигналів в частотній області?
11. Як оцінюється величина стиснення сигналів?

4. Домашнє завдання

1. Підготувати відповіді на ключові питання.
2. Підготувати таблиці для дослідження двох фрагментів сигналів звукового мовлення.

5. Опис лабораторного макета

В якості лабораторного макета використовується ЕОМ. Всі, без винятку, системи стиснення сигналів звукового мовлення використовуються не в аналоговому виді, а в цифровому. Інколи аналоговими бувають тільки фільтри на вході кодера і фільтри на виході декодера. Обробка сигналу в передавачі та приймачі в цифровому виді дозволяє зменшити розміри обладнання та уніфікувати окремі блоки.

В сучасних ЕОМ також використовують цифрову обробку сигналів. Тому при дослідженні систем стиснення нема необхідності виготовляти окремі макети обладнання, а всі алгоритми обробки сигналів можна моделювати на ЕОМ, в яких є плата, що дозволяє записувати звукові сигнали в реальному масштабі часу в пам'ять ЕОМ.

В процесі моделювання блоки відліків сигналів записуються також в пам'ять ЕОМ. При прослуховуванні фрагментів сигналів звукового мовлення можна порівнювати еталонний сигнал і сигнал, який прийшов по моделі каналу зв'язку, в якому є система стиснення сигналів.

6. Лабораторне завдання

На диску F знайти папку Group. В цій папці записані фрагменти сигналів естрадної та симфонічної музики, які прийшли по моделі стиснення сигналів. Запис `estr32_3.wav` означає, що число відліків в блоці дорівнює 32, а 3 виділено на передачу найбільшого частотного відліку у блоці, музика – естрадна. Оригінал фрагменту естрадної музики – `estr_original.wav`.

Фрагменти симфонічної музики записані так:

`simf_original.wav`, `simf256_7.wav` (256 відліків у блоці, 7 біт – для передачі найбільшого частотного відліку у блоці).

Для кожного з двох фрагментів сигналів при різній кількості відліків в блоці N знайти, при слуховому сприйнятті, величини b та m , при яких помітні спотворення, викликані системою стиснення сигналів звукового мовлення.

Величину b в бітах знаходять при прослуховуванні фрагментів сигналів, що пройшли по моделі системи стиснення. Коефіцієнт масштабу квантування m обчислюють по формулі (10), округляючи m до більшого цілого числа. Величину G знаходять по формулі (9).

На основі одержаних результатів та обчислень заповніть табл. 6.1 та 6.2.

Таблиця 6.1 – Дослідження стиснення сигналів естрадної музики

<i>N</i>	32	64	128	256	512	1024
<i>b</i>						
<i>m</i>						
<i>G</i>						

Таблиця 6.2 – Дослідження стиснення сигналів симфонічної музики

<i>N</i>	32	64	128	256	512	1024
<i>b</i>						
<i>m</i>						
<i>G</i>						

При проведенні досліджень зверніть увагу на те, що спотворення, викликані системою стиснення, більш помітні при прослуховуванні сигналів симфонічної музики.

7. Зміст протоколу

1. Таблиці з результатами досліджень та обчислень.
2. Висновки по результатам досліджень.

8. Список літератури

1. Рудий Є.М. Стиснення звукових сигналів: навчальний посіб. – Одеса: УДАЗ. – 1996. – 79 с.
2. Захарченко Н.В. Функциональные устройства телекоммуникационных систем. – Модуль 1: Проблемы сжатия информационных сигналов: учеб. пособие / Н.В. Захарченко, Е.М. Рудый. – Одесса: ИЦ ОНАС им. А.С. Попова. – 2005. – 96 с.
3. Система электросвязи. Том второй. Системы радио-, телевизионного вещания и документальной электросвязи. Под ред. Н.В. Захарченко. – К.: Техніка. – 1998. – 239 с.

Лабораторная работа 1.4 МОДЕЛИРОВАНИЕ СИСТЕМЫ СЖАТИЯ СИГНАЛОВ ЗВУКОВОГО ВЕЩАНИЯ

1. Цель работы

Формирование знаний и умений в отрасли сжатия сигналов звукового вещания, моделирование на ЭВМ структурных схем сжатия сигналов.

2. Ключевые положения

Для повышения эффективности использования каналов связи используют различные методы и средства. Важную роль играют методы сокращения избыточности данных, которые позволяют уменьшить объем переданной информации, что эквивалентно повышению пропускной способности канала связи. Уменьшение избыточности данных называют *сжатием*.

Сжатие всегда сопровождается искажением сигналов. Сжать сигнал без искажений невозможно. Искажения могут быть заметны при слуховом восприятии, а могут быть и незаметны. Именно потому, что существуют искажения, которые не замечаются ухом человека, возможно сжатие сигналов звукового вещания.

Нельзя позволить, чтобы во время передачи программы звукового вещания были заметны искажения сигналов, так как слушатели не захотят пользоваться таким вещанием.

При разработке системы сжатия учитывают свойства слуха человека и характеристики акустических сигналов.

В системах сжатия учитывают дискретность слуха по амплитуде и частоте, и взаимное маскирование сигналов. Если постепенно увеличивать амплитуду сигналов данной частоты, то человеку кажется, что сигнал увеличивается скачками, а не постепенно. В области средних частот количество таких скачков около 250. На низких и высоких частотах количество скачков уменьшается до 150.

Дискретность слуха наблюдается не только по амплитуде сигналов, но и по частоте. Во всем частотном диапазоне человек воспринимает около 250 градаций частоты, причем количество градаций зависит от интенсивности звуковых сигналов. При большой интенсивности сигналов воспринимается до 650 градаций частоты.

Граница частотного диапазона слуха человека 20...20000 Гц. У людей, которые находятся под воздействием громких звуков (трактористы, любители поп-музыки и др.) слух неуклонно ухудшается и частотный диапазон уменьшается. Это является защитной реакцией организма человека на очень мощный раздражитель, которыми являются громкие звуковые сигналы. Под воздействием громких звуков порог слышимости повышается (человек теряет слух). Поэтому музыку и передачи по радио и телевидению необходимо прослушивать при минимально допустимой громкости сигналов.

При одновременном воздействии на слух двух сигналов один из них может не прослушиваться в присутствии другого. Это явление называют *маскировкой сигнала*. Маскировка зависит от частоты и амплитуды сигналов. Экспериментально установлено, что низкочастотные сигналы сильнее маскируют высокочастотные. Если частота помехи намного больше полезного сигнала, то маскировка меньше.

Если сигнал маскируется узкополосным шумом, то при увеличении полосы шума, маскировка увеличивается, потому что энергия маскирующего сигнала растет с увеличением полосы шума. Но маскировка увеличивается до тех пор, пока ширина полосы пропускания не достигнет некоторого значения, которое называют *критической полосой слуха*.

При реализации систем сжатия сигналов звукового вещания учитывают дискретность слуха человека и маскировку в критических полосах слуха. Исследуют системы сжатия при помощи ЭВМ. При этом ЭВМ используют в качестве макета оборудования системы сжатия сигналов.

В ОНАС им. А.С. Попова исследования в области сжатия ведутся более 30 лет. Алгоритмы сжатия использованы в спутниковых системах передачи. Фирмы-разработчики охотно рекламируют свои достижения, пишут в общих чертах об особенностях кодирования, не раскрывая технические детали обработки сигналов, что не позволяет повторить системы сжатия сигналов без покупки лицензий или оборудования, которое выпускается фирмами.

В данной лабораторной работе используется для учебных целей один, конечно далеко не лучший, адаптивный алгоритм обработки сигналов, разработанный в нашей академии. В модели системы сжатия при помощи дискретного косинусного преобразования Фурье блок отсчетов $X(N)$ во временной области преобразуется в блок отсчетов $Y(K)$ в частотной области и наоборот.

Расчеты выполняются по таким формулам:

$$Y(1) = \frac{1}{\sqrt{N1}} \sum_{N=1}^{N1} X(N), \quad (1)$$

$$Y(K) = \sqrt{\frac{2}{N1}} \sum_{N=1}^{N1} X(N) \cos \frac{(2N-1)(K-1)\pi}{2N1}, \quad (2)$$

$$X(1) = \frac{1}{\sqrt{N1}} Y(1), \quad (3)$$

$$X(N) = \sqrt{\frac{2}{N1}} \sum_{K=1}^{N1} Y(K) \cos \frac{(2N-1)(K-1)\pi}{2N1}, \quad (4)$$

где $(N, K) = 1, 2, 3, \dots, N1$. $N1$ – число отсчетов сигналов в одном блоке во временной и частотной областях.

Входной массив чисел представляет собой временные отсчеты реальных речевых сигналов, полученных при равномерном квантовании. Величина массивов была такой, чтобы время звучания фрагментов сигналов, которые использовались при испытаниях было 7...10 с. Такая небольшая протяженность звучания фрагментов выбиралась для того, чтобы избежать мешающего воздействия слуховой памяти человека при проведении субъективно-статистических экспертиз.

При передаче отсчетов сигналов во временной области можно учесть корреляционные связи между отсчетами и уменьшить скорость цифрового потока по каналу связи. Но величина сжатия в этих системах мала. Если передаются отсчеты сигналов в частотной области, то можно учесть дискретность слуха и маскировку сигналов, что позволяет достичь большего сжатия сигналов.

Системы сжатия сигналов звукового вещания намного сложнее, чем системы сжатия телефонных сигналов потому, что при прослушивании передач не должны быть заметны на слух помехи, вызванные системами сжатия. Лучшие зарубежные образцы систем сжатия звукового вещания позволяют уменьшить скорость цифрового потока с 768 кбит/с до 128 кбит/с, если вещание ведется в полосе частот 20...15000 Гц. Системы сжатия, разработанные в ОНАС им. А.С. Попова, при такой же полосе частот уменьшают скорость цифрового потока до 42 кбит/с.

В учебном алгоритме сжатия сигналов звукового вещания используется почти мгновенное компандирование сигналов в частотной области. При этом в блоке отсчетов сигналов в частотной области находят наибольший частотный отсчет. Все другие отсчеты сигналов, которые, естественно меньше наибольшего, нормируют относительно наибольшего.

$$Y_{ni} = \frac{Y_i}{Y_1}, \quad (5)$$

где Y_1 – наибольший частотный отсчет в блоке; Y_i – величина i -того частотного отсчета; Y_{ni} – нормированная величина i -того частотного отсчета.

При таком подходе всегда выполняется условие

$$Y_{ni} \leq 1. \quad (6)$$

После прохождения сигналов по каналу связи возобновляют отсчет сигналов в блоке в частотной области

$$Y_i = Y_1 \cdot Y_{ni}. \quad (7)$$

На каждый отсчет сигналов во временной области выделяют 16 бит. Количество бит, которое выделяется на наибольший частотный отсчет в блоке, определяют по результатам субъективно-статистических экспертиз.

Чаще используют пятибалльную шкалу оценок. Эксперт сравнивает сигнал А на входе системы сжатия с сигналом Б, который обработан системой сжатия. Как правило, экспертов предупреждают о том, когда они слушают сигнал А, когда Б. Для проверки объективности показаний экспертов, иногда, за сигналом А повторяют вместо сигнала Б сигнал А.

Шкала оценок следующая:

- 5 – искажения Б в сравнении с А незаметны;
- 4 – искажения Б в сравнении с А заметны, но приемлемы;
- 3 – искажения слегка раздражающие, но допустимые;
- 2 – искажения раздражающие и недопустимые;
- 1 – искажения сильно раздражающие и недопустимые.

Желательно при экспертизах использовать так называемые критические фрагменты программ, при звучании которых искажения, вызванные сжатием сигналов, наиболее заметны.

Установлено, что в качестве критических фрагментов может использоваться звучание фортепиано без сопровождения оркестром, звучание звоночков и ксилофона.

Для создаваемых цифровых систем передачи рекомендуется граничная экспертная оценка 4,5 балла, то есть одна половина экспертов не замечает искажения, а вторая половина экспертов, хотя и замечает искажения, но считает их приемлемыми.

Величина сжатия G сигналов звукового вещания оценивается как

$$G = \frac{v_1}{v_2}, \quad (8)$$

где v_1 , v_2 – скорости передачи информации по каналу связи без сжатия и со сжатием информации.

Числовую величину коэффициента сжатия можно определить, предположив, что все отсчеты в блоке равны максимальному отсчету, на который рассчитан кодер канала связи.

В этом случае

$$G = \frac{aN_1}{m + bN_1}, \quad (9)$$

где a – число бит, которое необходимо для передачи максимального отсчета сигнала в блоке во временной области; b – число бит, которое выделяется для передачи наибольшего отсчета сигнала в блоке в частотной области; m – коэффициент масштаба квантования

$$m \leq \log_2 b. \quad (10)$$

Величина m округляется до ближайшего большего целого числа. $a = 16$ бит. Величина b в битах задается при прослушивании каждого фрагмента программ во время проведения субъективно-статистических экспертиз.

Если в системе сжатия сигналов используют статистическое кодирование по методу Шеннона-Фано или Хаффмена, то величина G увеличивается приблизительно в 2,5 раза.

3. Ключевые вопросы

1. Какой процесс называют сжатием сигналов?
2. Почему искажения, вызванные системами сжатия сигналов звукового вещания, не должны быть заметны ухом человека?
3. Что называется дискретностью слуха человека по амплитуде и частоте?
4. Почему ухудшается слух человека при длительном воздействии громких сигналов?
5. Какое явление называют маскировкой сигналов?

6. Что называют критической полосой слуха?

7. Какие особенности слуха человека используют при реализации систем сжатия сигналов?

8. Почему при передаче сигналов в частотной области достигают большего сжатия?

9. Какая скорость цифрового потока необходима при передаче сигналов звукового вещания в полосе 20...15000 Гц, если используют системы сжатия?

10. Как осуществляется почти мгновенное компандирование сигналов в частотной области?

11. Как оценивается величина сжатия сигналов?

4. Домашнее задание

1. Подготовить ответы на ключевые вопросы.

2. Подготовить таблицы для исследования двух фрагментов сигналов звукового вещания.

5. Описание лабораторного макета

В качестве лабораторного макета используется ЭВМ. Все, без исключения, системы сжатия сигналов звукового вещания выполняются не в аналоговом виде, а в цифровом. Иногда аналоговыми бывают только фильтры на входе кодера и фильтры на выходе декодера. Обработка сигналов в передатчике и приемнике в цифровом виде позволяет уменьшить размеры оборудования и унифицировать отдельные блоки.

В современных ЭВМ также используют цифровую обработку сигналов. Поэтому при исследовании систем сжатия нет необходимости изготавливать отдельные макеты оборудования, а все алгоритмы обработки сигналов можно моделировать на ЭВМ, в которых есть плата, позволяющая записывать звуковые сигналы в реальном масштабе времени в память ЭВМ.

В процессе моделирования блоки отсчетов сигналов записываются также в память ЭВМ. При прослушивании фрагментов сигналов звукового вещания можно сравнивать эталонный сигнал и сигнал прошедший по модели канала связи, в котором есть система сжатия сигналов.

6. Лабораторное задание

На диске F найти папку Group. В этой папке записаны фрагменты сигналов эстрадной и симфонической музыки, которые пришли по модели сжатия сигналов. Запись estr32_3.wav обозначает, что число отсчетов в блоке равно 32, а 3 бита выделено на передачу наибольшего частотного отсчета в блоке, музыка – эстрадная. Оригинал фрагмента эстрадной музыки – estr_original.wav.

Фрагменты симфонической музыки записаны так:

simf_original.wav, simf256_7.wav (256 отсчетов в блоке, 7 бит – для передачи наибольшего частотного отсчета в блоке).

Для каждого из двух фрагментов сигналов при разном количестве отсчетов в блоке N , найти, при слуховом восприятии величины b и m , при которых заметны искажения, вызванные сжатием сигналов звукового вещания.

Величину b в битах находят при прослушивании фрагментов сигналов, которые прошли по модели сжатия. Коэффициент масштаба квантования m вычисляют по формуле (10), округляя m к ближайшему большему целому числу. Величину G находят по формуле (9).

На основании полученных результатов и вычислений заполнить табл. 6.1 и 6.2.

Таблица 6.1 – Исследование сжатия сигналов эстрадной музыки

N	32	64	128	256	512	1024
b						
m						
G						

Таблица 6.2 – Исследование сжатия сигналов симфонической музыки

N	32	64	128	256	512	1024
b						
m						
G						

При проведении исследований обратите внимание на то, что искажения, вызванные системой сжатия, более заметны при прослушивании сигналов симфонической музыки.

7. Содержание протокола

1. Таблицы с результатами исследований и вычислений.
2. Выводы по результатам исследований.

8. Список литературы

1. Рудий Є.М. Стиснення звукових сигналів: навчальний посіб. – Одеса: УДАЗ. – 1996. – 79 с.
2. Захарченко Н.В. Функциональные устройства телекоммуникационных систем. – Модуль 1: Проблемы сжатия информационных сигналов: учеб. пособие/ Н.В. Захарченко, Е.М. Рудый. – Одесса: ИЦ ОНАС им. А.С. Попова. – 2005. – 96 с.
3. Система электросвязи. Том второй. Системы радио-, телевизионного вещания и документальной электросвязи. Под ред. Н.В. Захарченко. – К.: Техніка. – 1998. – 239 с.

Лабораторна робота 2.1

ПОРІВНЯННЯ АЛГОРИТМІВ СТИСНЕННЯ БЕЗ ВТРАТ

1. Мета роботи

Порівняння алгоритмів стиснення даних без втрат, застосованих до різних класів зображень.

2. Ключові положення

Переважає більшість сучасних форматів запису даних представлена у виді, зручному для швидкого маніпулювання і зручного прочитання користувачами. При цьому дані займають об'єм більший, ніж це дійсно потрібно для їх зберігання. Алгоритми, які усувають надмірність запису даних, називаються *алгоритмами стиснення даних*, або *алгоритмами архівації*.

Растрова і векторна графіка

Растрове зображення - зображення, що є сіткою пікселів або кольорових точок (зазвичай прямокутною) на комп'ютерному моніторі, папері та інших пристроях, що відображають, а також матеріалах.

Важливими характеристиками зображення є:

- кількість пікселів – дозвіл. Може вказуватися окремо кількість пікселів по ширині і висоті (1024*768, 640*480, ...) або ж, рідко, загальна кількість пікселів;

- кількість використовуваних кольорів або глибина кольору (ці характеристики мають наступну залежність: $N = 2^k$, де N – кількість кольорів, а k – глибина кольору);

- колірний простір RGB, CMYK, XYZ, YCbCr та ін.

RGB (абревіатура англійських слів Red, Green, Blue – червоний, зелений, синій) – аддитивна колірна модель, що як правило, описує спосіб синтезу кольору для кольоровідтворення.

Вибір основних кольорів обумовлений особливостями фізіології сприйняття кольору сітківкою людського ока. Колірна модель RGB знайшла широке застосування в техніці.

Чотириколірна автотипія (CMYK: Cyan, Magenta, Yellow, Key color) – субтрактивна схема формування кольору, використовувана передусім в поліграфії для стандартного тріадного друку. Схема CMYK, як правило, має (порівняно з RGB) невелике колірне охоплення.

Українською ці кольори часто називають блакитним, пурпурним і жовтим, хоча перший точніше називати синьо-зеленим, а маджента – лише частина пурпурного спектру, в якості четвертого кольору використовується чорний. Друк чотирма фарбами, відповідними CMYK, також називають друком тріадними фарбами.

CIE XYZ – лінійна 3-компонентна колірна модель, що ґрунтується на результатах виміру характеристик людського ока. Побудована на основі зорових можливостей "стандартного спостерігача", тобто гіпотетичного глядача, можливості якого були ретельно вивчені і зафіксовані в ході тривалих досліджень людського зору, проведених комітетом CIE.

В колірному просторі **YCbCr** є один компонент, який представляє яскравість (сигнал яскравості Y), і два інші компоненти, які представляють колір (сигнали колірності Cb – хроматичний синій, Cr – хроматичний червоний).

Растрову графіку редагують за допомогою растрових графічних редакторів. Створюється растрова графіка фотоапаратами, сканерами, безпосередньо в растровому редакторі, також шляхом експорту з векторного редактора або у вигляді скриншотів.

Достоїнства:

- Растрова графіка дозволяє створити практично будь-який малюнок, незалежно від складності. На відміну, наприклад від векторної, де неможливо точно передати ефект переходу від одного кольору до іншого без втрат у розмірі файлу.

- Поширеність – растрова графіка використовується зараз практично скрізь: від маленьких значків до плакатів.

- Висока швидкість обробки складних зображень, якщо не потрібне масштабування.

- Растрове представлення зображення природне для більшості пристроїв введення-виведення графічної інформації, таких як монітори (за винятком векторних), матричні і струминні принтери, цифрові фотоапарати, сканери.

Недоліки:

- Великий розмір файлів з простими зображеннями.

- Нemoжливiсть iдеального масштабування.

- Нemoжливiсть виводу на друк на плоттер.

Із за цих недоліків для зберігання простих малюнків рекомендують замість навіть стислої растрової графіки використати векторну графіку.

Растрові зображення зазвичай зберігаються в стислому виді. Також в графічному файлі може зберігатися додаткова інформація: про автора файлу, фотокамеру і її налаштування, кількість точок на дюйм при друці та ін.

Формати растрових зображень:

- BMP або Windows Bitmap – зазвичай використовується без стиснення, хоча можливе використання алгоритму RLE.

- GIF (Graphics Interchange Format) – застаріваючий формат, що підтримує не більше 256 кольорів одночасно.

- PCX застарілий формат, що дозволяв добре стискувати прості рисовані зображення.

- JPEG дуже широко використовуваний формат зображень. Стиснення засноване на усереднюванні кольору сусідніх пікселів (інформація про

яскравість при цьому не усереднюється) і відкиданні високочастотних складових в просторовому спектрі фрагмента зображення.

- TIFF підтримує великий діапазон зміни глибини кольору, різні колірні простори, різні налаштування стиснення (як з втратами, так і без) та ін.

- RAW зберігає інформацію, що безпосередньо отримується з матриці цифрового фотоапарата або аналогічного пристрою без застосування до неї яких-небудь перетворень, а також зберігає налаштування фотокамери. Практично усі напівпрофесійні і професійні цифрові фотоапарати дозволяють зберігати RAW зображення. Формат файлу залежить від моделі фотоапарата, єдиного стандарту не існує.

Векторна графіка – спосіб представлення об'єктів і зображень в комп'ютерній графіці, заснований на використанні геометричних примітивів, таких як точки, лінії, сплайни і багатокутники.

Розглянемо, приміром, такий графічний примітив, як коло радіусу r . Для його побудови необхідно і достатньо наступних початкових даних: координати центру кола; значення радіусу r ; колір заповнення (якщо коло не прозоре); колір і товщина контура (у разі наявності контура).

Переваги векторної графіки над растровою графікою

- Розмір, займаний описовою частиною, не залежить від реальної величини об'єкту, що дозволяє, використовуючи мінімальну кількість інформації, описати скільки завгодно великий об'єкт файлом мінімального розміру.

- У зв'язку з тим, що інформація про об'єкт зберігається в описовій формі, можна нескінченно збільшити графічний примітив, наприклад, дугу кола, і вона залишиться гладкою. З іншого боку, якщо крива представлена у вигляді ламаної лінії, збільшення покаже, що вона насправді не крива.

- Параметри об'єктів зберігаються і можуть бути легко змінені. Також це означає що переміщення, масштабування, обертання, заповнення і т.д. не погіршує якості малюнка.

- При збільшенні або зменшенні об'єктів товщина ліній може бути задана постійною величиною, незалежно від реального контура.

Фундаментальні недоліки векторної графіки:

- Не кожен об'єкт може бути легко зображений у векторному виді.

- Перетворення векторної графіки в растр досить просте. Але зворотного шляху, як правило, немає.

Технічні характеристики процесів стиснення

Метою процесу стиснення є отримання компактнішого вихідного потоку з некомпактного вхідного потоку. Основними технічними характеристиками процесів стиснення і результатів їх роботи є:

- коефіцієнт стиснення $K_{ст}$, або відношення об'ємів початкового і результуючого потоків;

– швидкість стиснення – час, що витрачається на стиснення вхідного потоку, до отримання з нього еквівалентного вихідного потоку;

– якість стиснення – величина, що показує, на скільки сильно упакований вихідний потік, за допомогою застосування до нього повторного стиснення за тим самим або іншим алгоритмом.

Усі способи стиснення можна розділити на дві категорії: *оборотне і безповоротне*.

Під *безповоротним стисненням* (архівація з втратами) мають на увазі таке перетворення вхідного потоку даних, при якому вихідний потік представляє досить схожий на вхідний потік об'єкт, проте відрізняється від нього об'ємом.

Міра схожості вхідного і вихідного потоків визначається мірою відповідності деяких властивостей об'єкту (тобто стислої і нестислої інформації відповідно до деякого певного формату даних), що представляється цим потоком інформації.

Оборотне стиснення (архівація без втрат) завжди призводить до зниження об'єму вихідного потоку інформації без зміни його інформативності, тобто - без втрати інформаційної структури. З вихідного потоку, за допомогою поновлюючого алгоритму, можна отримати вхідний. Процес відновлення називається декомпресією або розпаковуванням.

Для того, щоб коректніше оцінювати міру стиснення, треба ввести поняття класу зображень. Під класом розумітиметься деяка сукупність зображень, застосування до яких алгоритму архівації дає якісно однакові результати. Наприклад, для одного класу алгоритм дає дуже високу міру стиснення, для іншого – майже не стискає, для третього – збільшує файл в розмірі. (Відомо, що багато алгоритмів у гіршому разі збільшують файл). Розглянемо наступні приклади неформального визначення класів зображень:

Клас 1. Зображення з невеликою кількістю кольорів (4-16) і великими областями, заповненими одним кольором. Плавні переходи кольорів відсутні. Приклади: ділова графіка – гістограми, діаграми, графіки і тому подібне.

Клас 2. Зображення з плавними переходами кольорів, побудовані на комп'ютері. Приклади: графіка презентацій, ескізні моделі в системах автоматизованого проектування (САПР), зображення, побудовані за методом Гуро¹.

Клас 3. Фотореалістичні зображення. Приклад: відскановані фотографії.

Клас 4. Фотореалістичні зображення з накладенням ділової графіки. Приклад: реклама.

Алгоритм RLE

Цей алгоритм надзвичайно простий в реалізації. Групове кодування – від англійського Run Length Encoding (RLE) – один з найстаріших і найпростіших алгоритмів архівації. Стискування в RLE відбувається за рахунок того, що в

¹ Метод Гуро - метод зафарбовування в тривимірній комп'ютерній графіці (затінювання), призначений для створення ілюзії гладкої криволінійної поверхні

початковому потоці зустрічаються ланцюжки однакових байт. Заміна їх на пари <лічильник повторень, значення> зменшує надмірність даних. Ситуація, коли файл збільшується, для цього простого алгоритму не так вже рідкісна.

Розглянемо приклад архівації. Нехай 44 44 44 11 11 11 11 11 01 33 97 22 22 (13 байт) – початкова послідовність. У стислій послідовності перший байт вказує, скільки разів треба повторити наступний байт, тому для передачі перших трьох байт використовуватиметься 2 байти – 03 і 44, замість п'яти подальших – 05 і 11. Якщо далі знаходяться дані, що не повторюються, то перший байт дорівнює 00, а потім йде лічильник, що показує скільки за ним йде даних, що не повторюються, і самі дані, – 00 03 01 33 97. Стисла послідовність матиме вигляд: 03 44 05 11 00 03 01 03 97 02 22 (11 байт). Коефіцієнт стиснення при цьому складе $K_{\text{ст}} = 13/11 = 1,18$.

Ці методи, як правило, досить ефективні для стиснення растрових графічних зображень (BMP, PCX, TIFF, GIF), оскільки останні містять досить багато довгих серій послідовностей байтів, що повторюються. Недоліком методу RLE є досить низька міра стиснення.

Алгоритм LZW

Назву алгоритм дістав по перших буквах прізвищ його розробників – Lempel, Ziv і Welch. Стиснення в ньому, на відміну від RLE, здійснюється вже за рахунок однакових ланцюжків байт. Існує досить велике сімейство LZ-подібних алгоритмів, що розрізняються методом пошуку ланцюжків, що повторюються.

Перша річ, яка робиться при LZW-стисненні – ініціалізація ланцюжка символів. Щоб зробити це, необхідно вибрати код розміру (кількість біт) і знати, скільки можливих значень можуть приймати символи. Покладемо вихідний код розміром 12 біт, що означає можливість запам'ятовування 4096 елементів в таблиці рядків. Також припустимо, що маємо 256 можливих різних символів (наприклад, ASCII-таблиця² або зображення з глибиною кольору 8 біт). Щоб ініціалізувати таблицю, встановимо відповідність коду <0> символу 0, кода <1> символу 1, і так далі, до коду <255> і символу 255. Для коду очищення і коду кінця інформації зарезервовані значення <256> і <257>. Під коди для рядків залишаються значення від <258> до <4095>.

Процес стиснення виглядає досить просто. Прочитуємо послідовно символи вхідного потоку і перевіряємо, чи є в створеній таблиці рядків такий рядок. Якщо рядок є, то прочитуємо наступний символ, а якщо рядка немає, то заносимо в потік код для попереднього знайденого рядка, заносимо рядок в таблицю і починаємо пошук знову. Рядки, що додаються, записуються в таблицю послідовно, при цьому індекс рядка в таблиці стає її кодом.

Нехай стискаємо послідовність 45 55 55 151 55 55 55. Спочатку помістимо у вихідний потік код очищення <256>, потім додамо до порожнього рядка 45 і перевіримо, чи є рядок <45> у таблиці. Оскільки при ініціалізації

² ASCII -таблиця - англ. American Standard Code for Information Interchange - американська стандартна кодировочна таблиця для друкарських символів і деяких спеціальних кодів

занесли в таблицю усі рядки з одного символу, то рядок <45> є в таблиці. Далі читаємо наступний символ 55 з вхідного потоку і перевіряємо, чи є рядок 45 55 в таблиці. Такого рядка в таблиці доки немає. Заносимо в таблицю рядок 45 55 (з першим вільним кодом <258>) і записуємо в потік код <45> і так далі (див. табл. 1.1).

Таблиця 1.1 – Стиснення послідовності за алгоритмом LZW

Рядок	Наявність в таблиці рядків	Додавання в таблицю рядків	Вихідний потік
45	є (номер <45>)		
45 55	немає	<258> 45 55	<45>
55	є (номер <55>)		
55 55	немає	<259> 55 55	<55>
55	є (номер <55>)		
55 151	немає	<260> 55 151	<55>
151	є (номер <151>)		
151 55	немає	<261> 151 55	<151>
55	є (номер <55>)		
55 55	є (номер <259>)		
55 55 55	немає	<262> 55 55 55	<259>
55	є (номер <55>)		<55>

Послідовність кодів для цього прикладу, що потрапляють у вихідний потік: <256>, <45>, <55>, <55>, <151>, <259>, <55> і код завершення інформації <257>.

Особливість LZW полягає в тому, що для декомпресії не потрібно зберігати таблицю рядків у файл для розпаковування. Алгоритм побудований таким чином, що можна відновити таблицю рядків, користуючись тільки потоком кодів.

Зауваження. Можна відмітити, що записувані в потік коди поступово зростають. До тих пір, поки в таблиці не з'явиться, наприклад, вперше код 512, усі коди будуть менше 512. Крім того, при компресії і при декомпресії коди в таблиці додаються при обробці одного і того ж символу, тобто це відбувається "синхронно". Можна скористатися цією властивістю алгоритму для того, щоб підвищити міру компресії. Поки в таблицю не доданий 512 символ, писатимемо у вихідний бітовий потік коди з 9 біт, а відразу при додаванні 512 – коди з 10 біт. Відповідно декомпресор також повинен буде сприймати усі коди вхідного потоку 9-бітовими до моменту додавання в таблицю коду 512, після чого сприйматиме усі вхідні коди як 10-бітові. Аналогічно ми поступатимемо при додаванні в таблицю кодів 1024 і 2048. Цей прийом дозволяє приблизно на 15% підняти міру компресії.

3. Ключові питання

1. Назвіть достоїнства і недоліки растрової графіки.
2. Назвіть переваги векторної графіки над растровою.
3. Назвіть основні технічні характеристики процесів стиснення.

4. У чому різниця між алгоритмами з втратою інформації і без втрати інформації?
5. На який клас зображень орієнтований алгоритм RLE?
6. Наведіть два приклади "поганих" зображень для алгоритму RLE, для яких файл максимально збільшиться в розмірі.
7. За рахунок чого відбувається стиснення в алгоритмі LZW?
8. Наведіть приклад "поганого" зображення для алгоритму LZW, для якого файл максимально збільшиться в розмірі.
9. Порівняйте алгоритми стиснення зображень без втрат.
10. У яких випадках алгоритми стиснення без втрат стискають зображення з втратами?

4. Домашнє завдання

За розділом 2 даного посібника вивчіть принципи побудови стислих послідовностей за допомогою кожного з наведених алгоритмів та підготуйте відповіді на ключові запитання.

5. Лабораторне завдання

1. За двома останніми цифрами номера залікової книжки з табл. 5.1 вибрати інформаційні послідовності для стиснення за допомогою алгоритмів RLE і LZW. Для алгоритму LZW використати послідовність, повторену двічі.

Таблиця 5.1 – Інформаційні послідовності для стиснення

0	11 44 11 44 44 44 44
1	151 56 78 78 78 78
2	56 56 23 23 23 02 02
3	02 02 02 44 44 11 11
4	23 23 23 11 11 78 89
5	02 56 11 11 44 44 44
6	44 44 11 11 11 44 56
7	20 20 11 11 11 89 02
8	44 44 44 11 78 89 89
9	151 151 151 02 02

2. Розрахувати коефіцієнти стиснення для отриманих послідовностей (при розрахунку коефіцієнта для алгоритму LZW слід пам'ятати, що вхідний потік був 8-бітовим, а вихідний – 9-бітовим). Зробити висновки про отриманих результатах.

3. Визначити клас зображень, представлених в папці "Стиснення без втрат" на Робочому столі комп'ютера.

4. Для кожного класу окремо провести ту, що перекодувала зображень у формат tiff з використанням стиснення за алгоритмом LZW, змішаним алгоритмом Хаффмана і RLE, а також модифікованим алгоритмом Хаффмана.

5. Розрахувати коефіцієнт стискування для кожного алгоритму і класу. Результати обчислень занести в табл. 5.2.

Таблиця 5.2 – Порівняння алгоритмів

Им'я файлу	Клас зображень	Алгоритм стиснення	Розмір вихідного файлу, кб	Розмір стислого файлу, кб	Коефіцієнт сжатия, $K_{ст}$	Втрати (є/немає)
	1	LZW				
		Хафман і RLE				
		CCITT Group 3				
	2	LZW				
		Хафман і RLE				
		CCITT Group 3				
	3	LZW				
		Хафман і RLE				
		CCITT Group 3				

6. За результатами обчислень зробити висновки про придатність алгоритмів стиснення для запропонованих класів зображень.

6. Зміст протоколу

1. Тема й мета лабораторної роботи.
2. Відповіді на ключові питання.
3. Послідовності, стиснуті за допомогою алгоритмів RLE и LZW, а також розраховані коефіцієнти стиснення.
4. Підсумкова таблиця 5.2.
5. Висновки по виконаній роботі.

7. Список літератури

1. Рудий Є. М. Технології передачі дискретних повідомлень. Т.1: Стиснення сигналів: підручник. – Одеса: ОНАЗ ім. О.С. Попова, 2011. – 170 с.

Лабораторная работа 2.1 СРАВНЕНИЕ АЛГОРИТМОВ СЖАТИЯ БЕЗ ПОТЕРЬ

1. Цель работы

Сравнение алгоритмов сжатия данных без потерь, примененных к различным классам изображений.

2. Ключевые положения

Подавляющее большинство современных форматов записи данных представлены в виде, удобном для быстрого манипулирования и удобного прочтения пользователями. При этом данные занимают объем больший, чем это действительно требуется для их хранения. Алгоритмы, которые устраняют избыточность записи данных, называются *алгоритмами сжатия данных*, или *алгоритмами архивации*.

Растровая и векторная графика

Растровое изображение – изображение, представляющее собой сетку пикселей или цветных точек (обычно прямоугольную) на компьютерном мониторе, бумаге и других отображающих устройствах и материалах.

Важными характеристиками изображения являются:

- количество пикселей – разрешение. Может указываться отдельно количество пикселей по ширине и высоте (1024*768, 640*480, ...) или же, редко, общее количество пикселей;

- количество используемых цветов или глубина цвета (эти характеристики имеют следующую зависимость: $N = 2^k$, где N – количество цветов, а k – глубина цвета);

- цветовое пространство RGB, CMYK, XYZ, YCbCr и др.

RGB (аббревиатура английских слов **R**ed, **G**reen, **B**lue – красный, зелёный, синий) – аддитивная цветовая модель, как правило, описывающая способ синтеза цвета для цветовоспроизведения.

Выбор основных цветов обусловлен особенностями физиологии восприятия цвета сетчаткой человеческого глаза. Цветовая модель RGB нашла широкое применение в технике.

Четырёхцветная автотипия (CMYK: Cyan, Magenta, Yellow, Key color) – субтрактивная схема формирования цвета, используемая прежде всего в полиграфии для стандартной триадной печати. Схема CMYK, как правило, обладает (сравнительно с RGB) небольшим цветовым охватом.

По-русски эти цвета часто называют голубым, пурпурным и жёлтым, хотя первый точнее называть сине-зелёным, а маджента – лишь часть пурпурного спектра, а в качестве четвёртого используется чёрный. Печать четырьмя красками, соответствующими CMYK, также называют печатью *триадными красками*.

CIE XYZ – линейная 3-компонентная цветовая модель, основанная на результатах измерения характеристик человеческого глаза. Построена на основе зрительных возможностей «стандартного наблюдателя», то есть гипотетического зрителя, возможности которого были тщательно изучены и зафиксированы в ходе длительных исследований человеческого зрения, проведённых комитетом CIE.

В цветовом пространстве **YCbCr** есть один компонент, который представляет яркость (сигнал яркости Y), и два других компонента, которые представляют цвет (сигналы цветности Cb – хроматический синий, Cr – хроматический красный).

Растровую графику редактируют с помощью растровых графических редакторов. Создается растровая графика фотоаппаратами, сканерами, непосредственно в растровом редакторе, также путем экспорта из векторного редактора или в виде скриншотов.

Достоинства

- Растровая графика позволяет создать практически любой рисунок, вне зависимости от сложности. В отличие, например, от векторной, где невозможно

точно передать эффект перехода от одного цвета к другому без потерь в размере файла.

- Распространённость — растровая графика используется сейчас практически везде: от маленьких значков до плакатов.

- Высокая скорость обработки сложных изображений, если не нужно масштабирование.

- Растровое представление изображения естественно для большинства устройств ввода-вывода графической информации, таких как мониторы (за исключением векторных), матричные и струйные принтеры, цифровые фотоаппараты, сканеры.

Недостатки

- Большой размер файлов с простыми изображениями.

- Невозможность идеального масштабирования.

- Невозможность вывода на печать на плоттер.

- Из-за этих недостатков для хранения простых рисунков рекомендуют вместо даже сжатой растровой графики использовать векторную графику.

Растровые изображения обычно хранятся в сжатом виде. Также в графическом файле может храниться дополнительная информация: об авторе файла, фотокамере и её настройках, количестве точек на дюйм при печати и др.

Форматы растровых изображений

- BMP или Windows Bitmap – обычно используется без сжатия, хотя возможно использование алгоритма RLE.

- GIF (Graphics Interchange Format) – устаревающий формат, поддерживающий не более 256 цветов одновременно.

- PCX устаревший формат, позволявший хорошо сжимать простые рисованные изображения.

- JPEG очень широко используемый формат изображений. Сжатие основано на усреднении цвета соседних пикселей (информация о яркости при этом не усредняется) и отбрасывании высокочастотных составляющих в пространственном спектре фрагмента изображения.

- TIFF поддерживает большой диапазон изменения глубины цвета, разные цветовые пространства, разные настройки сжатия (как с потерями, так и без) и др.

- RAW хранит информацию, непосредственно получаемую с матрицы цифрового фотоаппарата или аналогичного устройства без применения к ней каких-либо преобразований, а также хранит настройки фотокамеры. Практически все полупрофессиональные и профессиональные цифровые фотоаппараты позволяют сохранять RAW изображения. Формат файла зависит от модели фотоаппарата, единого стандарта не существует.

Векторная графика – способ представления объектов и изображений в компьютерной графике, основанный на использовании геометрических примитивов, таких как точки, линии, сплайны и многоугольники.

Рассмотрим, к примеру, такой графический примитив, как окружность радиуса r . Для её построения необходимо и достаточно следующих исходных

данных: координаты центра окружности; значение радиуса r ; цвет заполнения (если окружность не прозрачная); цвет и толщина контура (в случае наличия контура).

Преимущества векторной графики над растровой графикой

- Размер, занимаемый описательной частью, не зависит от реальной величины объекта, что позволяет, используя минимальное количество информации, описать сколько угодно большой объект файлом минимального размера.

- В связи с тем, что информация об объекте хранится в описательной форме, можно бесконечно увеличить графический примитив, например, дугу окружности, и она останется гладкой. С другой стороны, если кривая представлена в виде ломаной линии, увеличение покажет, что она на самом деле не кривая.

- Параметры объектов хранятся и могут быть легко изменены. Также это означает что перемещение, масштабирование, вращение, заполнение и т. д. не ухудшает качества рисунка.

- При увеличении или уменьшении объектов толщина линий может быть задана постоянной величиной, независимо от реального контура.

Фундаментальные недостатки векторной графики.

- Не каждый объект может быть легко изображен в векторном виде.
- Перевод векторной графики в растр достаточно прост. Но обратного пути, как правило, нет.

Технические характеристики процессов сжатия

Целью процесса сжатия является получение более компактного выходного потока из изначально некомпактного входного потока. Основными техническими характеристиками процессов сжатия и результатов их работы являются:

- коэффициент сжатия $K_{сж}$, или отношение объемов исходного и результирующего потоков;

- скорость сжатия - время, затрачиваемое на сжатие входного потока, до получения из него эквивалентного выходного потока;

- качество сжатия - величина, показывающая, на сколько сильно упакован выходной поток, при помощи применения к нему повторного сжатия по этому же или иному алгоритму.

Все способы сжатия можно разделить на две категории: *обратимое и необратимое*.

Под *необратимым сжатием* (архивация с потерями) подразумевают такое преобразование входного потока данных, при котором выходной поток представляет достаточно похожий на входной поток объект, однако отличается от него объемом.

Степень сходства входного и выходного потоков определяется степенью соответствия некоторых свойств объекта (т.е. сжатой и несжатой информации в соответствии с некоторым определенным форматом данных), представляемого данным потоком информации.

Обратимое сжатие (архивация без потерь) всегда приводит к снижению объема выходного потока информации без изменения его информативности, т.е. – без потери информационной структуры. Из выходного потока, при помощи восстанавливающего алгоритма, можно получить входной. Процесс восстановления называется декомпрессией или распаковкой.

Для того чтобы корректнее оценивать степень сжатия, нужно ввести понятие класса изображений. Под классом будет пониматься некая совокупность изображений, применение к которым алгоритма архивации дает качественно одинаковые результаты. Например, для одного класса алгоритм дает очень высокую степень сжатия, для другого – почти не сжимает, для третьего – увеличивает файл в размере. (Известно, что многие алгоритмы в худшем случае увеличивают файл). Рассмотрим следующие примеры неформального определения классов изображений:

Класс 1. Изображения с небольшим количеством цветов (4-16) и большими областями, заполненными одним цветом. Плавные переходы цветов отсутствуют. Примеры: деловая графика – гистограммы, диаграммы, графики и т.п.

Класс 2. Изображения, с плавными переходами цветов, построенные на компьютере. Примеры: графика презентаций, эскизные модели в системах автоматизированного проектирования (САПР), изображения, построенные по методу Гуро³.

Класс 3. Фотореалистичные изображения. Пример: отсканированные фотографии.

Класс 4. Фотореалистичные изображения с наложением деловой графики. Пример: реклама.

Алгоритм RLE

Данный алгоритм необычайно прост в реализации. Групповое кодирование – от английского Run Length Encoding (RLE) – один из самых старых и самых простых алгоритмов архивации. Сжатие в RLE происходит за счет того, что в исходном потоке встречаются цепочки одинаковых байт. Замена их на пары <счетчик повторений, значение> уменьшает избыточность данных. Ситуация, когда файл увеличивается, для этого простого алгоритма не так уж редка.

Рассмотрим пример архивации. Пусть 44 44 44 11 11 11 11 11 01 33 97 22 22 (13 байт) – исходная последовательность. В сжатой последовательности первый байт указывает, сколько раз нужно повторить следующий байт, поэтому для передачи первых трех байт будет использоваться 2 байта – 03 и 44, вместо пяти последующих – 05 и 11. Если далее находятся неповторяющиеся данные, то первый байт равен 00, а затем идет счетчик, показывающий сколько за ним следует неповторяющихся данных и сами данные – 00 03 01 33 97. Сжатая последовательность будет иметь вид: 03 44 05 11 00 03 01 03 97 02 22 (11 байт). Коэффициент сжатия при этом составит $K_{\text{сж}} = 13/11 = 1,18$.

³ Метод Гуро - метод закрашивания в трёхмерной компьютерной графике (затенения), предназначенный для создания иллюзии гладкой криволинейной поверхности

Данные методы, как правило, достаточно эффективны для сжатия растровых графических изображений (BMP, PCX, TIFF, GIF), т.к. последние содержат достаточно много длинных серий повторяющихся последовательностей байтов. Недостатком метода RLE является достаточно низкая степень сжатия.

Алгоритм LZW

Название алгоритм получил по первым буквам фамилий его разработчиков – Lempel, Ziv и Welch. Сжатие в нем, в отличие от RLE, осуществляется уже за счет одинаковых цепочек байт. Существует довольно большое семейство LZ-подобных алгоритмов, различающихся методом поиска повторяющихся цепочек.

Первая вещь, которая делается при LZW-сжатии – инициализация цепочки символов. Чтобы сделать это, необходимо выбрать код размера (количество бит) и знать, сколько возможных значений могут принимать символы. Положим выходной код размером 12 бит, что означает возможность запоминания 4096 элементов в таблице строк. Также предположим, что имеем 256 возможных различных символов (например, ASCII-таблица⁴ или изображение с глубиной цвета 8 бит). Чтобы инициализировать таблицу, установим соответствие кода <0> символу 0, кода <1> символу 1, и т.д., до кода <255> и символа 255. Для кода очистки и кода конца информации зарезервированы значения <256> и <257>. Под коды для строк остаются значения от <258> до <4095>.

Процесс сжатия выглядит достаточно просто. Считываем последовательно символы входного потока и проверяем, есть ли в созданной таблице строк такая строка. Если строка есть, то считываем следующий символ, а если строки нет, то заносим в поток код для предыдущей найденной строки, заносим строку в таблицу и начинаем поиск снова. Добавляемые строки записываются в таблицу последовательно, при этом индекс строки в таблице становится ее кодом.

Пусть сжимаем последовательность 45 55 55 151 55 55 55. Сначала поместим в выходной поток код очистки <256>, потом добавим к изначально пустой строке 45 и проверим, есть ли строка <45> в таблице. Поскольку при инициализации занесли в таблицу все строки из одного символа, то строка <45> есть в таблице. Далее читаем следующий символ 55 из входного потока и проверяем, есть ли строка 45 55 в таблице. Такой строки в таблице пока нет. Заносим в таблицу строку 45 55 (с первым свободным кодом <258>) и записываем в поток код <45> и т.д. (см. табл. 1.1)

Последовательность кодов для данного примера, попадающих в выходной поток: <256>, <45>, <55>, <55>, <151>, <259>, <55> и код завершения информации <257>.

⁴ ASCII-таблица - англ. *American Standard Code for Information Interchange* — американская стандартная кодировочная таблица для печатных символов и некоторых специальных кодов.

Таблица 1.1 – Сжатие последовательности по алгоритму LZW

Строка	Наличие в таблице строк	Добавление в таблицу строк	Выходной поток
45	есть (номер <45>)		
45 55	нет	<258> 45 55	<45>
55	есть (номер <55>)		
55 55	нет	<259> 55 55	<55>
55	есть (номер <55>)		
55 151	нет	<260> 55 151	<55>
151	есть (номер <151>)		
151 55	нет	<261> 151 55	<151>
55	есть (номер <55>)		
55 55	есть (номер <259>)		
55 55 55	нет	<262> 55 55 55	<259>
55	есть (номер <55>)		<55>

Особенность LZW заключается в том, что для декомпрессии не надо сохранять таблицу строк в файл для распаковки. Алгоритм построен таким образом, что можно восстановить таблицу строк, пользуясь только потоком кодов.

Замечание. Можно отметить, что записываемые в поток коды постепенно возрастают. До тех пор, пока в таблице не появится, например, в первый раз код 512, все коды будут меньше 512. Кроме того, при компрессии и при декомпрессии коды в таблице добавляются при обработке одного и того же символа, т.е. это происходит “синхронно”. Можно воспользоваться этим свойством алгоритма для того, чтобы повысить степень компрессии. Пока в таблицу не добавлен 512 символ, будем писать в выходной битовый поток коды из 9 бит, а сразу при добавлении 512 – коды из 10 бит. Соответственно декомпрессор также должен будет воспринимать все коды входного потока 9-битными до момента добавления в таблицу кода 512, после чего будет воспринимать все входные коды как 10-битные. Аналогично мы будем поступать при добавлении в таблицу кодов 1024 и 2048. Данный прием позволяет примерно на 15% поднять степень компрессии:

3. Ключевые вопросы

1. Назовите основные технические характеристики процессов сжатия.
2. В чем разница между алгоритмами с потерей информации и без потери информации?
3. На какой класс изображений ориентирован алгоритм RLE?
4. Приведите два примера “плохих” изображений для алгоритма RLE, для которых файл максимально увеличится в размере.
5. За счет чего происходит сжатие в алгоритме LZW?
6. Приведите пример “плохого” изображения для алгоритма LZW, для которого файл максимально увеличится в размере.
7. На какой класс изображений ориентирован классический алгоритм Хаффмана?

8. Приведите пример “плохого” изображения для классического алгоритма Хаффмана.
9. Сравните алгоритмы сжатия изображений без потерь.
10. В каких случаях алгоритмы сжатия без потерь сжимают изображение с потерями?
11. Назовите достоинства и недостатки растровой графики.
12. Назовите преимущества векторной графики над растровой.

4. Домашнее задание

По разделу 2 данного пособия изучить принципы построения сжатых последовательностей с помощью каждого из приведенных алгоритмов и подготовьте ответы на ключевые вопросы.

5. Лабораторное задание

1. По 2-м последним цифрам номера зачетной книжки из таблицы 5.1 выбрать информационные последовательности для сжатия с помощью алгоритмов RLE и LZW. Для алгоритма LZW использовать последовательность, повторенную дважды.

Таблица 5.1 – Информационные последовательности для сжатия

0	11 44 11 44 44 44 44
1	151 56 78 78 78 78
2	56 56 23 23 23 02 02
3	02 02 02 44 44 11 11
4	23 23 23 11 11 78 89
5	02 56 11 11 44 44 44
6	44 44 11 11 11 44 56
7	20 20 11 11 11 89 02
8	44 44 44 11 78 89 89
9	151 151 151 02 02

2. Рассчитать коэффициенты сжатия для полученных последовательностей (при расчете коэффициента для алгоритма LZW следует помнить, что входной поток был 8-битным, а выходной – 9-битным). Сделать выводы о полученных результатах.
3. Определить класс изображений, представленных в папке «Сжатие без потерь» на Рабочем столе компьютера.
4. Для каждого класса в отдельности провести перекодировку изображений в формат tiff с использованием сжатия по алгоритму LZW, смешанному алгоритму Хаффмана и RLE, а также модифицированному алгоритму Хаффмана.
5. Рассчитать коэффициент сжатия для каждого алгоритма и класса. Результаты вычислений занести в табл. 5.2.

Таблица 5.2 – Сравнение алгоритмов сжатия

Имя файла	Класс изображений	Алгоритм сжатия	Размер исходного файла, кб	Размер сжатого файла, кб	Коэффициент сжатия, $K_{сж}$	Потери (есть/ нет)
	1	LZW				
		Хаффман и RLE				
		CCITT Group 3				
	2	LZW				
		Хаффман и RLE				
		CCITT Group 3				
	3	LZW				
		Хаффман и RLE				
		CCITT Group 3				

6. По результатам вычислений сделать выводы о пригодности алгоритмов сжатия для предложенных классов изображений

6. Содержание протокола

1. Тема и цель лабораторной работы.
2. Ответы на ключевые вопросы.
3. Последовательности, сжатые с помощью алгоритмов RLE и LZW, а также рассчитанные коэффициенты сжатия.
4. Итоговая табл. 5.2.
5. Выводы по проделанной работе.

7. Список литературы

1. Рудий Є.М. Технології передачі дискретних повідомлень. Т.1: Стиснення сигналів: підручник. – Одеса: ОНАЗ ім. О.С. Попова, 2011. – 170 с.

Лабораторна робота 2.2 АЛГОРИТМ СТИСНЕННЯ ДАНИХ JPEG

1. Мета роботи

Вивчення алгоритму стиснення JPEG, оцінка якості стислого зображення й визначення найкращого фактора якості для заданих зображень.

2. Ключові положення

В 1986 році підгрупою CCITT були розпочаті дослідження методів стиснення кольорових і півтонових даних для факсимільного зв'язку. Застосовувані при цьому методи стиснення кольорових даних дуже нагадували

ті, які досліджувалися групою JPEG. Тому було ухвалене рішення об'єднати ресурси цих груп для спільної роботи над єдиним стандартом.

JPEG не був визначений у якості стандартного формату файлів зображень, однак на його основі були створені нові або модифіковані існуючі файлові формати.

Оперує алгоритм областями 8×8 біт, на яких яскравість і колір змінюються порівняно плавно. Внаслідок цього, при розкладанні матриці такої області в подвійний ряд по косинусах значимими виявляються тільки перші коефіцієнти. Таким чином, стиснення в JPEG здійснюється за рахунок плавності зміни кольорів у зображенні.

У цілому алгоритм заснований на дискретному косинусоїдальному перетворенні (ДКП), яке є різновидом дискретного перетворення Фур'є, застосовуваному до матриці зображення для одержання деякої нової матриці коефіцієнтів. Для одержання вихідного зображення застосовується зворотне перетворення.

ДКП розкладає зображення по амплітудах деяких частот. Таким чином, при перетворенні ми одержуємо матрицю, у якій багато коефіцієнтів або близькі, або дорівнюють нулю. Крім того, завдяки недосконалості людського зору, можна апроксимувати коефіцієнти більш грубо без помітної втрати якості зображення.

Для цього використовується квантування коефіцієнтів. У найпростішому випадку – це арифметичний побітовий зсув вправо. При цьому перетворенні втрачається частина інформації, але можуть досягатися більші коефіцієнти стиснення.

Процес стиснення за схемою JPEG включає ряд етапів (рис. 2.1):

- Перетворення зображення в оптимальний колірний простір.
- Застосування дискретних косинусних перетворень для зменшення надмірності даних зображення.
- Квантування кожного блоку коефіцієнтів ДКП із застосуванням вагових функцій, оптимізованих з урахуванням візуального сприйняття людиною.
- Кодування результуючих коефіцієнтів (даних зображення) із застосуванням алгоритму групового кодування й алгоритму Хафмана для видалення надмірності інформації.

Розглянемо коротенько особливості кожного з перерахованих етапів. При цьому хотілося б звернути увагу на те, що декодування JPEG здійснюється у зворотному порядку.

Перетворення зображення в оптимальний колірний простір

У принципі алгоритм JPEG здатний кодувати зображення, засновані на будь-якому типі колірного простору (наприклад, розбитті кольорів на три складові – [червоний, зелений і синій] або [яскравість, хроматичний червоний, хроматичний синій] і ін.). JPEG кодує кожний компонент колірної моделі окремо, що забезпечує його повну незалежність від будь-якої моделі колірного простору.

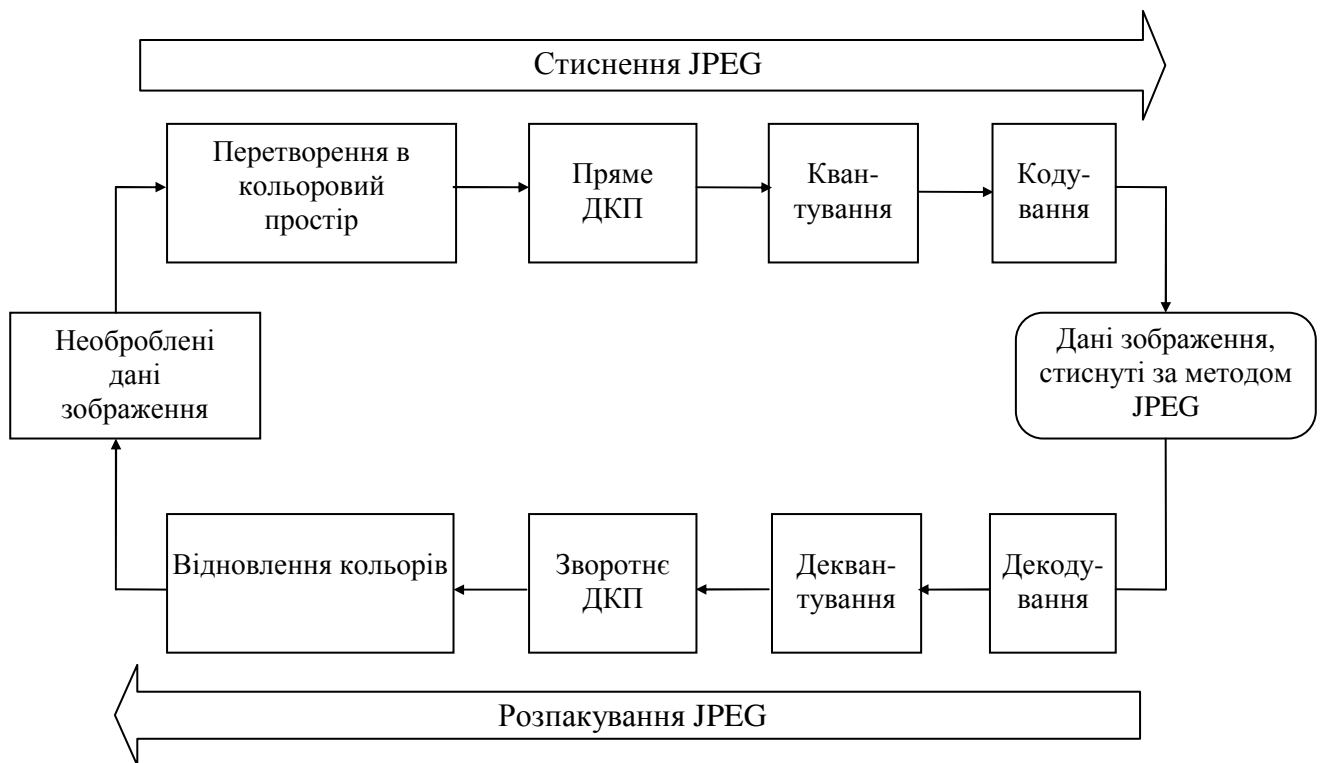


Рисунок 2.1 – Структура JPEG-перетворень

У випадку застосування колірного простору яскравість/кольоровість, наприклад такого, як $YCbCr$, досягається кращий ступінь стиснення. Компонента Y являє собою інтенсивність, а Cb і Cr – кольоровість (хроматичний червоний, хроматичний синій). Ця модель може бути переведена в RGB за допомогою перетворення без якої-небудь корекції насиченості. Для півтонових зображень (у градаціях сірого) використовується тільки одна складова Y .

Спрощено переведення з колірного простору RGB у колірний простір $YCrCb$ можна представити в такий спосіб:

$$\begin{aligned}
 Y &= 0.299R + 0.587G + 0.114B \\
 Cr &= 0.5R - 0.4184G - 0.0813B + 128 \\
 Cb &= 0.1687R - 0.3313G + 0.5B + 128
 \end{aligned}$$

Зворотнє перетворення здійснюється так:

$$\begin{aligned}
 R &= Y + 1.402Cb \\
 G &= Y - 0.34414Cr - 0.71414Cb - 128 \\
 B &= Y + 1.772Cr - 128
 \end{aligned}$$

Сегментація зображення

Сегментація зображення застосовується з метою розподілу його на дві й більше частин (підзображень). Це полегшує буферизацію даних зображення в пам'яті ПЕОМ, прискорює їхню довільну вибірку з диска, і дозволяє зберігати зображення розміром понад 64*64 кб. JPEG підтримує три типи сегментації зображень: просту, пірамідальну й комбіновану.

При *простій сегментації* зображення ділиться на два або більше сегментів фіксованого розміру. Усі прості сегменти кодуються ліворуч праворуч і зверху вниз, є суміжними й не перекриваються. Сегменти повинні мати однакову кількість вибірок і ідентифікаторів компонентів, і бути закодованими по одній схемі. Сегменти в нижній і правій частинах зображення можуть бути меншого розміру, чим "внутрішні" сегменти, оскільки величина зображення не обов'язково повинна бути кратною розмірам сегмента.

При *пірамідальній сегментації* зображення також ділиться на сегменти, а кожний з них, у свою чергу, – на ще більш дрібні сегменти. При цьому використовуються різні рівні дозволу. Моделлю такого процесу є сегментована піраміда зображення JPEG (JPEG Tiled Image Pyramid, JTIP), що відбиває процедуру створення пірамідального JPEG-зображення з декількома рівнями дозволу.

Комбінована сегментація дозволяє зберігати й відтворювати версії зображень із декількома рівнями дозволу у вигляді мозаїки. Комбінована сегментація допускає наявність сегментів, що перекриваються, різних розмірів, з різними коефіцієнтами масштабування й параметрами стиснення. Кожний сегмент кодується окремо й може комбінуватися з іншими сегментами без повторної дискретизації.

Наприклад, у випадку використання сегментів розміром 8*8 пікселів, для кожного блоку формується набір чисел. Перші кілька чисел представляють колір блоку в цілому, у той час як наступні числа відбивають більш тонкі деталі. Спектр деталей базується на зоровому сприйнятті людини, тому великі деталі більш помітні.

На наступному етапі, залежно від обраного рівня якості, відкидається певна частина чисел, що представляють тонкі деталі.

Таким чином, чим вище рівень компресії, тим більше даних відкидається й тем нижче якість зображення. Використовуючи JPEG можна одержати файл в 1-500 раз менше, ніж формат незжатих зображень BMP.

Дискретне косинусне перетворення

Ключовим компонентом роботи алгоритму є дискретне косинусне перетворення. Дискретне косинусне перетворення являє собою різновид перетворення Фур'є й, так само як і воно, має зворотне перетворення. Графічне зображення можна розглядати як сукупність просторових хвиль, причому осі X і Y збігаються із шириною й висотою картинки, а по осі Z відкладається значення кольору відповідного пікселя зображення. Дискретне косинусне перетворення дозволяє переходити від просторового подання картинки до її спектрального подання й зворотно. Впливаючи на спектральне подання

картинки, що складається з “гармонік”, тобто, відкидаючи найменш значимі з них, можна балансувати між якістю відтворення й ступенем стиснення. При цьому утворюється матриця, у якій коефіцієнти в лівому верхньому куті відповідають низькочастотній складовій зображення, а в правому нижньому – високочастотній.

Це перетворення можна представити так:

$$Y[u, v] = \frac{1}{\sqrt{2n}} \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} C(i, u) \times C(j, v) \times y[i, j], \quad (1)$$

де $C(i, u) = A(u) \times \cos\left(\frac{(2 \times i + 1) \times u \times \pi}{2 \cdot n}\right)$ – гармоніка сигналу, (2)

$$A(u) = \begin{cases} \frac{1}{\sqrt{2}}, & u \equiv 0 \\ 1, & u \neq 0 \end{cases} \quad \text{– постійна складова.} \quad (3)$$

Вираз для зворотного перетворення матриці “гармонік”, застосовуваний при розпакуванні зображення записується у вигляді

$$y[i, j] = \frac{1}{\sqrt{2n}} \sum_{u=0}^{n-1} \sum_{v=0}^{n-1} C(i, u) \times C(j, v) \times Y[u, v], \quad (4)$$

По визначенню дискретного косинусного перетворення для його реалізації потрібно два вкладені цикли, і тіло циклів буде виконуватися $n \times n$ разів для кожного елемента матриці дискретного косинусного перетворення. Значно більш ефективний варіант обчислення коефіцієнтів дискретного косинусного перетворення реалізований через перемножування матриць. У цьому випадку схему обчислення частотних коефіцієнтів матриці Y доцільно представити у вигляді множення матриць відповідно до відношення

$$Y = X \cdot y \cdot X^T, \quad (5)$$

де y – матриця вхідного зображення,

X – матриця постійних коефіцієнтів косинусного перетворення розміру $n \times n$, значення елементів якої обчислюються по формулі

$$X(i, j) = \begin{cases} \frac{1}{\sqrt{n}}, & i = 0 \\ \sqrt{\frac{2}{n}} \cdot \cos\left[\frac{(2j+1)i\pi}{2n}\right], & i > 0 \end{cases}, \quad (6)$$

X^T – транспонована матриця X .

Зворотнє перетворення виглядає наступним чином:

$$y = X^T \cdot Y \cdot X, \quad (7)$$

Цей варіант реалізації ДКП більш привабливий ще й тому, що сучасні архітектури багатопроцесорних обчислювачів виконують стандартні матричні

операції множення й транспонування. При перемножуванні двох матриць розміру $n \times n$ для обчислення одного елемента результуючої матриці необхідно виконати n множень і n додавань.

Квантування

Дискретне косинусне перетворення являє собою перетворення інформації без втрат і не здійснює ніякого стиснення. Напроти, дискретне косинусне перетворення підготовляє інформацію для етапу стиснення із втратами або округлення.

Округлення являє собою процес зменшення кількості битів, необхідних для зберігання коефіцієнтів матриці дискретного косинусного перетворення за рахунок втрати точності.

Стандарт JPEG реалізує цю процедуру через матрицю квантування. Для кожного елемента матриці дискретного косинусного перетворення існує відповідний елемент матриці квантування. Результуюча матриця утворюється діленням кожного елемента матриці дискретного косинусного перетворення на відповідний елемент матриці квантування й наступним округленням результату до найближчого цілого числа.

$$Yq[u, v] = E \left[\frac{Y[u, v]}{q[u, v]} \right], \quad (8)$$

де $E[]$ – ціла частина від ділення,
 $q[u, v]$ – матриця квантування.

Як правило, значення елементів матриці квантування зростають за напрямком ліворуч праворуч і зверху вниз. Від вибору матриці квантування залежить баланс між ступенем стиснення зображення і його якістю після відновлення. Стандарт JPEG дозволяє використовувати будь-яку матрицю квантування, однак ISO розробила набір матриць округлення.

На цьому етапі більшість JPEG-компресорів управляються за допомогою установки якості. Компресор використовує вбудовану таблицю, розраховану на середню якість, і нарощує або зменшує значення кожного елемента таблиці назад пропорційно необхідному якості. Застосовувані таблиці квантування записуються в стислий файл, щоб декомпресор знав, як відновити коефіцієнти ДКП.

Із квантуванням зв'язані й специфічні ефекти алгоритму. При малих значеннях фактора якості втрати в низьких частотах можуть бути настільки великі, що вихідне зображення (рис. 2.2,а) розпадеться на квадрати 8×8 (рис. 2.2,б). Втрати у високих частотах можуть виявитися в так званому “ефекті Гібса”, коли навколо контурів з різким переходом кольору утворюється своєрідний “німб” (рис. 2.2,в).

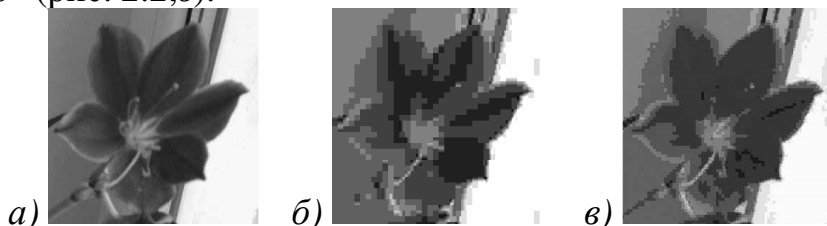


Рисунок 2.2 – Специфічні артефакти алгоритму JPEG

Кодування

Переводимо матрицю 8×8 в 64-елементний вектор за допомогою “зигзаг” - сканування, тобто беремо елементи з індексами $(0,0)$, $(0,1)$, $(1,0)$, $(2,0)$..., як показано на рис. 2.3.

0,0	0,1	0,2	0,3	0,4	0,5	0,6	0,7
1,0	1,1	1,2	1,3	1,4	1,5	1,6	1,7
2,0	2,1	2,2	2,3	2,4	2,5	2,6	2,7
3,0	3,1	3,2	3,3	3,4	3,5	3,6	3,7
4,0	4,1	4,2	4,3	4,4	4,5	4,6	4,7
5,0	5,1	5,2	5,3	5,4	5,5	5,6	5,7
6,0	6,1	6,2	6,3	6,4	6,5	6,6	6,7
7,0	7,1	7,2	7,3	7,4	7,5	7,6	7,7

Рисунок 2.3 – "Зигзаг" – сканування матриці

Таким чином, на початку вектора ми одержуємо коефіцієнти матриці, відповідні до низьких частот, а наприкінці – високим.

Заключна стадія роботи кодера JPEG – це власне кодування. Воно включає три дії над округленою матрицею дискретного косинусного перетворення, для того, щоб підвищити ступінь стиснення.

Перша дія – це заміна абсолютного значення коефіцієнта, розташованого в комірці $(0,0)$ матриці, на відносне. Тому що сусідні блоки зображення в значній мірі “схожі” один на одного, то кодування чергового $(0,0)$ елемента як різниці з попереднім дає менше значення.

Коефіцієнти матриці дискретного косинусного перетворення обходяться зигзагом. Після чого нульові значення кодуються з використанням алгоритму кодування повторів (RLE), а потім результат обробляється за допомогою “кодування ентропії”, тобто алгоритмів Хафмана або арифметичного кодування, залежно від реалізації.

Реалізація алгоритму в лабораторній роботі

У лабораторній роботі в якості вихідного використовується півтонове зображення, тобто зображення, що містить 3 однакові матриці інтенсивності Y . Тому всі процедури проводяться тільки для однієї матриці, і перетворення в колірний простір не використовується. Сегментація поділяє зображення на блоки 8×8 , після чого кожний блок зазнає ДКП.

У даній лабораторній роботі застосовується спрощений метод квантування, що використовує маску квантування, яка обнулює низькочастотні коефіцієнти залежно від фактора якості. Потім відбувається кодування кодами RLE і Хафмана.

Маска квантування – це матриця розміром 8×8 , що складається з одиниць і нулів, кількість одиниць у якій пропорційно фактору якості N . При факторі якості 100% маска містить усі одиниці. При факторі якості 1% – усього одну одиницю в елементі з індексом $(0,0)$.

Маска заповнюється за принципом «зигзаг»-сканування. Наприклад, $N = 83$. Кількість одиниць розраховується в такий спосіб: $[83 \cdot 64 / 100] = 53$. Маска квантування виглядає таким чином (рис. 2.4):

1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1

Рисунок 2.4 – Маска квантування для фактору якості 83

Оцінка втрат якості

Одна із серйозних проблем машинної графіки полягає в тому, що дотепер не знайдений адекватний критерій оцінки втрат якості зображення. А втрачається воно постійно – при оцифруванні, при переведенні в обмежену палітру кольорів, при переведенні в іншу систему кольоропредставлення для друку, і, що для нас особливо важливо, при архівації із втратами. Можна навести приклад простого критерію: середньоквадратичне відхилення (СКВ) значень пікселів:

$$d(x, y) = \sqrt{\frac{\sum_{i=1}^n \sum_{j=1}^n (x_{ij} - y_{ij})^2}{n^2}}. \quad (9)$$

По ньому зображення буде сильно зіпсоване при зниженні яскравості всього на 5% (око цього не помітить – у різних моніторів настроювання яскравості варіюється набагато сильніше). У той же час зображення зі “снігом” – різкою зміною кольору окремих точок, слабкими смугами або “муаром” будуть визнані такими, що “майже не змінилися”. Свої неприємні сторони є й в інших критеріях. Розглянемо, наприклад, максимальне відхилення (МВ):

$$d(x, y) = \max_{i,j} |x_{ij} - y_{ij}|. \quad (10)$$

Ця міра, як можна здогадатися, у край чутлива до биття окремих пікселів. Тобто у всьому зображенні може суттєво змінитися тільки значення одного пікселя (що практично непомітно для ока), однак згідно із цією мірою зображення буде сильно зіпсоване.

Міра, який зараз використовується на практиці, називається мірою відношення сигналу до шуму (ВСШ):

$$d(x, y) = 10 \log_{10} \frac{255 \cdot n^2}{\sum_{i=1}^n \sum_{j=1}^n (x_{ij} - y_{ij})^2}. \quad (11)$$

Дана міра, по суті, аналогічна середньоквадратичному відхиленню, однак користуватися нею трохи зручніше за рахунок логарифмічного масштабу шкали. Їй властиві ті ж недоліки, що й середньоквадратичному відхиленню.

У якості ще однієї оцінки якості стиснення використовують коефіцієнт стиснення, який дорівнює відношенню вихідного й стислого розмірів файлів. Однак даний показник не завжди точно відображає співвідношення якості/ступінь стиснення.

Найкраще втрати якості зображень оцінюють наші очі. Відмінною вважається архівація, при якій неможливо на око розрізнити початкове й розархівоване зображення. Доброю – коли сказати, яке із зображень зазнало архівації, можна тільки порівнюючи дві картинки, що перебувають поруч. При подальшому збільшенні ступеня стиснення, як правило, стають помітні побічні ефекти, характерні для даного алгоритму. На практиці, навіть при відмінному збереженні якості, у зображення можуть бути внесені регулярні специфічні зміни.

3. Ключові питання

1. За рахунок чого здійснюється стиснення в алгоритмі JPEG?
2. Для чого необхідне перетворення кольорового простору зображення?
3. Назвіть види сегментації зображень і їх відмінності. Яка сегментація використовується в алгоритмі JPEG?
4. Для чого потрібно дискретне косинусне перетворення?
5. На якому етапі стиснення відбуваються втрати якості зображення? Як вони пов'язані з фактором якості?
6. Яким способом проявляються втрати якості в алгоритмі JPEG?
7. Які алгоритми кодування використовує JPEG?
8. Приведіть приклади мір втрати інформації й опишіть їхні недоліки.

4. Домашнє завдання

1. За розділом 2 даного посібника вивчіть та зарисуйте алгоритм JPEG і підготуйте відповіді на ключові питання.
2. Підготуйте 3 маски квантування відповідно до фактора якості N , де N вибирається в такий спосіб:
 - 1) останні 2 цифри залікової книжки студента;
 - 2) $n*2$, де n – номер бригади;
 - 3) $100 - 2*n$.

5. Лабораторне завдання

1. Провести розрахунок дискретного косинусного перетворення для матриці вихідного зображення, що заданна в виді:

N	$N+1$	$N+2$	$N-1$
$N+3$	$N-1$	$N-2$	N
$N-3$	$N+2$	$N+1$	$N+2$
$N+1$	$N+2$	$N+3$	$N+1$

де N – дві останні цифри залікової книжки.

2. Проквантувати отриману матрицю.
3. Відновити матрицю зображення за допомогою зворотного ДКП.
4. Оцінити втрати якості зображення. Для парних варіантів визначити середньоквадратичне відхилення, для непарних – відношення сигнал/шум.
5. Запустити з робочого стола програму MATLAB. У командному рядку написати команду *guide*. У вікні, що з'явилося, перейти на вкладку Open Existing GUI і вибрати файл JPEG_Compression. Натиснути кнопку  на панелі інструментів.

6. Для початку роботи натиснути кнопку «Пуск» і ввести маску квантування, розраховану в домашньому завданні. Отримані на екрані дані звести в таблицю вигляду табл. 5.1:

Таблиця 5.1 – Підсумки виконання лабораторної роботи

Фактор якості	Розмір файлу, байт		Коефіцієнт стиснення	Оцінка якості			
	вхідного	стислого		СКВ	МВ	ВСП	Суб'єктивна

Стиснення необхідно провести для трьох факторів якості, заданих у домашньому завданні. Коефіцієнт стиснення розраховується самостійно. Суб'єктивна оцінка стиснення проводиться по зображенню «Відхилення».

7. Змінюючи маску квантування, знайти фактор якості, відповідний до суб'єктивної оцінки «5», при якому стиснення буде максимально. Дані додати в таблицю 1.

6. Зміст протоколу

1. Тема й мета лабораторної роботи.
2. Структурна схема JPEG-перетворень.
2. Відповіді на ключові питання.
3. Три маски квантування відповідно до домашнього завдання.
4. Підсумкова таблиця 1 із заповненими *чотирма* рядками відповідно до п.п. 2 і 3 лабораторного завдання.
5. Висновки по виконаній роботі.

7. Список літератури

1. Рудий Є.М. Технології передачі дискретних повідомлень. Т.1: Стиснення сигналів: підручник. – Одеса: ОНАЗ ім. О.С. Попова, 2011. – 170 с.

Лабораторная работа 2.2

АЛГОРИТМ СЖАТИЯ ДАННЫХ JPEG

1. Цель работы

Изучение алгоритма сжатия JPEG, оценка качества сжатого изображения и определение наилучшего фактора качества для заданных изображений.

2. Ключевые положения

В 1986 году подгруппой CCITT были начаты исследования методов сжатия цветных и полутоновых данных для факсимильной связи. Применяемые при этом методы сжатия цветных данных очень напоминали те, которые исследовались группой JPEG. Поэтому было принято решение объединить ресурсы этих групп для совместной работы над единым стандартом.

JPEG не был определен в качестве стандартного формата файлов изображений, однако на его основе были созданы новые или модифицированы существовавшие файловые форматы.

Оперирует алгоритм областями 8×8 бит, на которых яркость и цвет меняются сравнительно плавно. Вследствие этого, при разложении матрицы такой области в двойной ряд по косинусам значимыми оказываются только первые коэффициенты. Таким образом, сжатие в JPEG осуществляется за счет плавности изменения цветов в изображении.

В целом алгоритм основан на дискретном косинусоидальном преобразовании (ДКП), которое является разновидностью дискретного преобразования Фурье, применяемом к матрице изображения для получения некоторой новой матрицы коэффициентов. Для получения исходного изображения применяется обратное преобразование.

ДКП раскладывает изображение по амплитудам некоторых частот. Таким образом, при преобразовании мы получаем матрицу, в которой многие коэффициенты либо близки, либо равны нулю. Кроме того, благодаря несовершенству человеческого зрения, можно аппроксимировать коэффициенты более грубо без заметной потери качества изображения.

Для этого используется квантование коэффициентов. В самом простом случае – это арифметический побитовый сдвиг вправо. При этом преобразовании теряется часть информации, но могут достигаться большие коэффициенты сжатия.

Процесс сжатия по схеме JPEG включает ряд этапов (рис. 2.1):

- Преобразование изображения в оптимальное цветовое пространство.
- Применение дискретных косинусных преобразований для уменьшения избыточности данных изображения.
- Квантование каждого блока коэффициентов ДКП с применением весовых функций, оптимизированных с учетом визуального восприятия человеком.

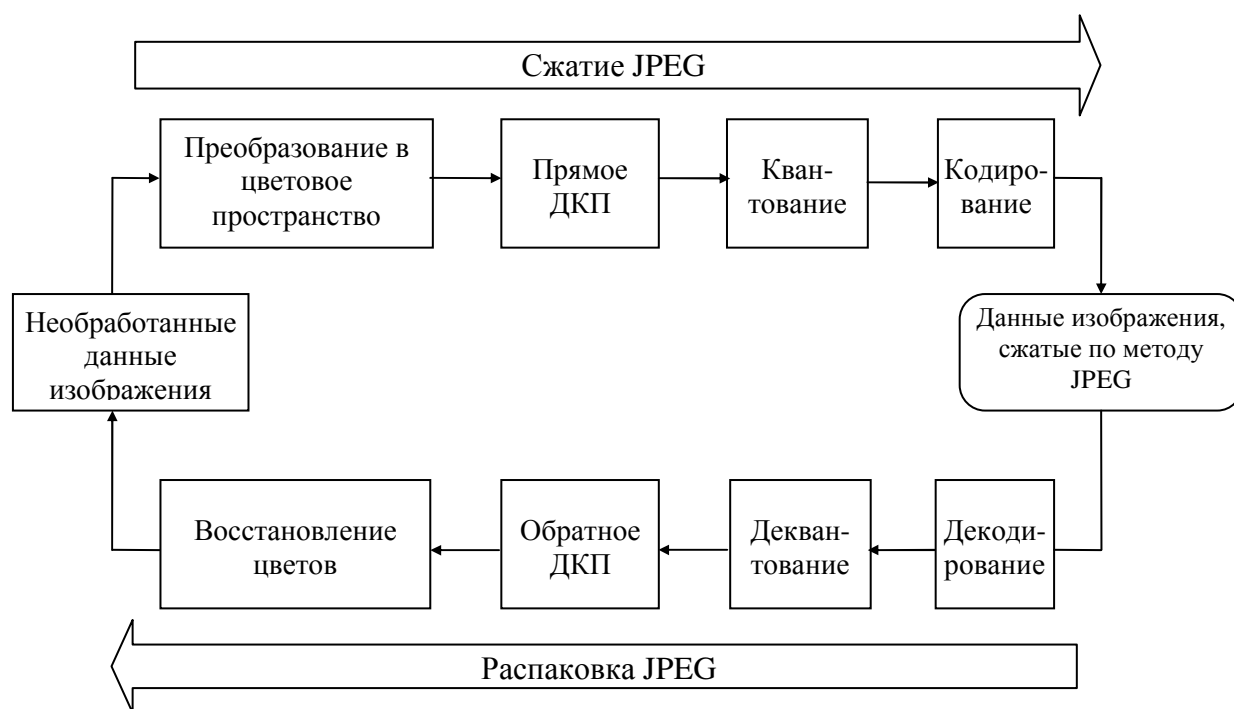


Рисунок 2.1 – Структура JPEG-преобразований

– Кодирование результирующих коэффициентов (данных изображения) с применением алгоритма группового кодирования и алгоритма Хаффмана для удаления избыточности информации.

Рассмотрим вкратце особенности каждого из перечисленных этапов. При этом хотелось бы обратить внимание на то, что декодирование JPEG осуществляется в обратном порядке.

Преобразование изображения в оптимальное цветное пространство

В принципе алгоритм JPEG способен кодировать изображения, основанные на любом типе цветного пространства (например, разбиении цветов на три составляющие – [красный, зеленый и синий] или [яркость, хроматический красный, хроматический синий] и др.). JPEG кодирует каждый компонент цветовой модели отдельно, что обеспечивает его полную независимость от любой модели цветного пространства.

В случае применения цветного пространства яркость/цветность, например такого, как $YCbCr$, достигается лучшая степень сжатия. Компонента Y представляет собой интенсивность, а Cb и Cr – цветность (хроматический красный, хроматический синий). Эта модель может быть переведена в RGB посредством преобразования без какой-либо коррекции насыщенности. Для полутоновых изображений (в градациях серого) используется только одна составляющая Y .

Упрощенно перевод из цветного пространства RGB в цветное пространство $YCrCb$ можно представить таким образом:

$$\begin{aligned}
 Y &= 0.299R + 0.587G + 0.114B \\
 Cr &= 0.5R - 0.4184G - 0.0813B + 128 \\
 Cb &= 0.1687R - 0.3313G + 0.5B + 128
 \end{aligned}$$

Обратное преобразование осуществляется так:

$$\begin{aligned} R &= Y + 1.402Cb \\ G &= Y - 0.34414Cr - 0.71414Cb - 128 \\ B &= Y + 1.772Cr - 128 \end{aligned}$$

Сегментация изображения

Сегментация изображения применяется с целью деления его на две и более частей (подизображений). Это облегчает буферизацию данных изображения в памяти ПЭВМ, ускоряет их произвольную выборку с диска, и позволяет хранить изображения размером свыше 64*64 кб. JPEG поддерживает три типа сегментации изображений: простую, пирамидальную и комбинированную.

При *простой сегментации* изображение делится на два или более сегментов фиксированного размера. Все простые сегменты кодируются слева направо и сверху вниз, являются смежными и не перекрывающимися. Сегменты должны иметь одинаковое количество выборок и идентификаторов компонентов, и быть закодированными по одной схеме. Сегменты в нижней и правой частях изображения могут быть меньшего размера, чем "внутренние" сегменты, поскольку величина изображения не обязательно должна быть кратной размерам сегмента.

При *пирамидальной сегментации* изображение также делится на сегменты, а каждый из них, в свою очередь, – на еще более мелкие сегменты. При этом используются различные уровни разрешения. Моделью такого процесса является сегментированная пирамида изображения JPEG (JPEG Tiled Image Pyramid, JTIIP), отражающая процедуру создания пирамидального JPEG-изображения с несколькими уровнями разрешения.

Комбинированная сегментация позволяет хранить и воспроизводить версии изображений с несколькими уровнями разрешения в виде мозаики. Комбинированная сегментация допускает наличие перекрывающихся сегментов разных размеров, с разными коэффициентами масштабирования и параметрами сжатия. Каждый сегмент кодируется отдельно и может комбинироваться с другими сегментами без повторной дискретизации.

Например, в случае использования сегментов размером 8*8 пикселей, для каждого блока формируется набор чисел. Первые несколько чисел представляют цвет блока в целом, в то время как последующие числа отражают более тонкие детали. Спектр деталей базируется на зрительном восприятии человека, поэтому крупные детали более заметны.

На следующем этапе, в зависимости от выбранного уровня качества, отбрасывается определенная часть чисел, представляющих тонкие детали.

Таким образом, чем выше уровень компрессии, тем больше данных отбрасывается и тем ниже качество изображения. Используя JPEG можно получить файл в 1-500 раз меньше, чем формат несжатых изображений BMP.

Дискретное косинусное преобразование

Ключевым компонентом работы алгоритма является дискретное косинусное преобразование. Дискретное косинусное преобразование представляет собой разновидность преобразования Фурье и, так же как и оно, имеет обратное преобразование. Графическое изображение можно рассматривать как совокупность пространственных волн, причем оси X и Y совпадают с шириной и высотой картинки, а по оси Z откладывается значение цвета соответствующего пикселя изображения. Дискретное косинусное преобразование позволяет переходить от пространственного представления картинки к ее спектральному представлению и обратно. Воздействуя на спектральное представление картинки, состоящее из “гармоник”, то есть, отбрасывая наименее значимые из них, можно балансировать между качеством воспроизведения и степенью сжатия. При этом образуется матрица, в которой коэффициенты в левом верхнем углу соответствуют низкочастотной составляющей изображения, а в правом нижнем – высокочастотной.

Это преобразование можно представить так:

$$Y[u, v] = \frac{1}{\sqrt{2n}} \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} C(i, u) \times C(j, v) \times y[i, j], \quad (1)$$

где $C(i, u) = A(u) \times \cos\left(\frac{(2 \times i + 1) \times u \times \pi}{2 \cdot n}\right)$ – гармоника сигнала, (2)

$$A(u) = \begin{cases} \frac{1}{\sqrt{2}}, & u \equiv 0 \\ 1, & u \neq 0 \end{cases} \text{ – постоянная составляющая.} \quad (3)$$

Выражение для обратного преобразования матрицы “гармоник”, применяемое при распаковке изображения записывается в виде

$$y[i, j] = \frac{1}{\sqrt{2n}} \sum_{u=0}^{n-1} \sum_{v=0}^{n-1} C(i, u) \times C(j, v) \times Y[u, v], \quad (4)$$

По определению дискретного косинусного преобразования для его реализации требуется два вложенных цикла, и тело циклов будет выполняться $n \times n$ раз для каждого элемента матрицы дискретного косинусного преобразования. Значительно более эффективный вариант вычисления коэффициентов дискретного косинусного преобразования реализован через перемножение матриц. В этом случае схему вычисления частотных коэффициентов матрицы Y целесообразно представить в виде умножения матриц в соответствии с отношением

$$Y = X \cdot y \cdot X^T, \quad (5)$$

где y – матрица исходного изображения,

X – матрица постоянных коэффициентов косинусного преобразования размера $n \times n$, значения элементов которой вычисляются по формуле

$$X(i, j) = \begin{cases} \frac{1}{\sqrt{n}}, i = 0 \\ \sqrt{\frac{2}{n}} \cdot \cos\left[\frac{(2j+1)i\pi}{2n}\right], i > 0 \end{cases}, \quad (6)$$

X^T – транспонированная матрица X .

Обратное преобразование выглядит следующим образом:

$$y = X^T \cdot Y \cdot X, \quad (7)$$

Этот вариант реализации ДКП более привлекателен еще и потому, что современные архитектуры многопроцессорных вычислителей выполняют стандартные матричные операции умножения и транспонирования. При перемножении двух матриц размера $n \times n$ для вычисления одного элемента результирующей матрицы необходимо выполнить n умножений и n сложений.

Квантование

Дискретное косинусное преобразование представляет собой преобразование информации без потерь и не осуществляет никакого сжатия. Напротив, дискретное косинусное преобразование подготавливает информацию для этапа сжатия с потерями или округления.

Округление представляет собой процесс уменьшения количества битов, необходимых для хранения коэффициентов матрицы дискретного косинусного преобразования за счет потери точности.

Стандарт JPEG реализует эту процедуру через матрицу квантования. Для каждого элемента матрицы дискретного косинусного преобразования существует соответствующий элемент матрицы квантования. Результирующая матрица получается делением каждого элемента матрицы дискретного косинусного преобразования на соответствующий элемент матрицы квантования и последующим округлением результата до ближайшего целого числа.

$$Yq[u, v] = E\left[\frac{Y[u, v]}{q[u, v]}\right], \quad (8)$$

где $E[]$ – целая часть от деления,

$q[u, v]$ – матрица квантования.

Как правило, значения элементов матрицы квантования растут по направлению слева направо и сверху вниз. От выбора матрицы квантования зависит баланс между степенью сжатия изображения и его качеством после восстановления. Стандарт JPEG позволяет использовать любую матрицу квантования, однако ISO разработала набор матриц округления.

На этом этапе большинство JPEG-компрессоров управляются с помощью установки качества. Компрессор использует встроенную таблицу, рассчитанную на среднее качество, и наращивает или уменьшает значение каждого элемента таблицы обратно пропорционально требуемому качеству. Применяемые таблицы квантования записываются в сжатый файл, чтобы декомпрессор знал, как восстановить коэффициенты ДКП.

С квантованием связаны и специфические эффекты алгоритма. При малых значениях фактора качества потери в низких частотах могут быть настолько велики, что исходное изображение (рис. 2.2,*а*) распадется на квадраты 8*8 (рис. 2.2,*б*). Потери в высоких частотах могут проявиться в так называемом “эффекте Гиббса”, когда вокруг контуров с резким переходом цвета образуется своеобразный “нимб” (рис. 2.2,*в*).

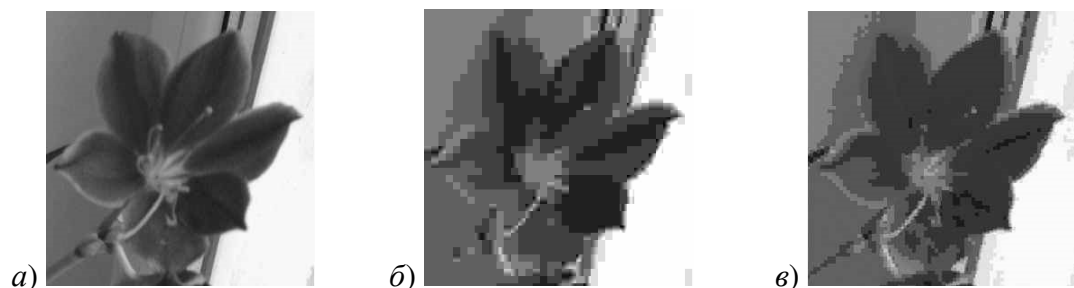


Рисунок 2.2 – Специфические артефакты алгоритма JPEG

Кодирование

Переводим матрицу 8*8 в 64-элементный вектор при помощи “зигзаг” - сканирования, т.е. берем элементы с индексами (0,0), (0,1), (1,0), (2,0)...., как показано на рис. 2.3.

0,0	0,1	0,2	0,3	0,4	0,5	0,6	0,7
1,0	1,1	1,2	1,3	1,4	1,5	1,6	1,7
2,0	2,1	2,2	2,3	2,4	2,5	2,6	2,7
3,0	3,1	3,2	3,3	3,4	3,5	3,6	3,7
4,0	4,1	4,2	4,3	4,4	4,5	4,6	4,7
5,0	5,1	5,2	5,3	5,4	5,5	5,6	5,7
6,0	6,1	6,2	6,3	6,4	6,5	6,6	6,7
7,0	7,1	7,2	7,3	7,4	7,5	7,6	7,7

Рисунок 2.3 – "Зигзаг" – сканирование матрицы

Таким образом, в начале вектора мы получаем коэффициенты матрицы, соответствующие низким частотам, а в конце – высоким.

Заключительная стадия работы кодера JPEG – это собственно кодирование. Оно включает три действия над округленной матрицей дискретного косинусного преобразования, для того, чтобы повысить степень сжатия.

Первое действие – это замена абсолютного значения коэффициента, расположенного в ячейке (0,0) матрицы, на относительное. Так как соседние блоки изображения в значительной степени “похожи” друг на друга, то кодирование очередного (0,0) элемента как разницы с предыдущим дает меньшее значение.

Коэффициенты матрицы дискретного косинусного преобразования обходятся зигзагом. После чего нулевые значения кодируются с использованием алгоритма кодирования повторов (RLE), а потом результат обрабатывается с помощью “кодирования энтропии”, то есть алгоритмов Хаффмана или арифметического кодирования, в зависимости от реализации.

Реализация алгоритма в лабораторной работе

В лабораторной работе в качестве исходного используется полутоновое изображение, т.е. изображение, содержащее 3 одинаковые матрицы интенсивности Y . Поэтому все процедуры производятся только для одной матрицы, и преобразование в цветное пространство не используется. Сегментация делит изображение на блоки 8×8 , после чего каждый блок подвергается ДКП.

В данной лабораторной работе применяется упрощенный метод квантования, использующий маску квантования, которая обнуляет низкочастотные коэффициенты в зависимости от фактора качества. Затем происходит кодирование кодами RLE и Хаффмана.

Маска квантования – это матрица размером 8×8 , состоящая из единиц и нулей, количество единиц в которой пропорционально фактору качества N . При факторе качества 100% маска содержит все единицы. При факторе качества 1% – всего одну единицу в элементе с индексом (0,0).

Маска заполняется по принципу «зигзаг»-сканирования. Например, $N = 83$. Количество единиц рассчитывается следующим образом: $[83 \cdot 64 / 100] = 53$. Тогда маска квантования выглядит так (рис. 2.4):

1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	0
1	1	1	1	1	1	1	0
1	1	1	1	1	1	0	0
1	1	1	1	1	0	0	0
1	1	1	1	0	0	0	0

Рисунок 2.4 – Маска квантования для фактора качества 83

Оценка потерь качества

Одна из серьезных проблем машинной графики заключается в том, что до сих пор не найден адекватный критерий оценки потерь качества изображения. А теряется оно постоянно – при оцифровке, при переводе в ограниченную палитру цветов, при переводе в другую систему цветопредставления для печати, и, что для нас особенно важно, при архивации с потерями. Можно привести пример простого критерия: среднеквадратичное отклонение (СКО) значений пикселей:

$$d(x, y) = \sqrt{\frac{\sum_{i=1}^n \sum_{j=1}^n (x_{ij} - y_{ij})^2}{n^2}}. \quad (9)$$

По нему изображение будет сильно испорчено при понижении яркости всего на 5% (глаз этого не заметит – у разных мониторов настройка яркости варьируется гораздо сильнее). В то же время изображения со “снегом” – резким изменением цвета отдельных точек, слабыми полосами или “муаром” будут

признаны “почти не изменившимися”. Свои неприятные стороны есть и у других критериев. Рассмотрим, например, максимальное отклонение (МО):

$$d(x, y) = \max_{i,j} |x_{ij} - y_{ij}|. \quad (10)$$

Эта мера, как можно догадаться, крайне чувствительна к биению отдельных пикселей. Т.е. во всем изображении может существенно измениться только значение одного пикселя (что практически незаметно для глаза), однако согласно этой мере изображение будет сильно испорчено.

Мера, которую сейчас используют на практике, называется мерой отношения сигнала к шуму (ОСШ):

$$d(x, y) = 10 \log_{10} \frac{255 \cdot n^2}{\sum_{i=1}^n \sum_{j=1}^n (x_{ij} - y_{ij})^2}. \quad (11)$$

Данная мера, по сути, аналогична среднеквадратичному отклонению, однако пользоваться ей несколько удобнее за счет логарифмического масштаба шкалы. Ей присущи те же недостатки, что и среднеквадратичному отклонению.

В качестве еще одной оценки качества сжатия используют коэффициент сжатия, равный отношению исходного и сжатого размеров файлов. Однако данный показатель не всегда точно отображает соотношение качество/степень сжатия.

Лучше всего потери качества изображений оценивают наши глаза. Отличной считается архивация, при которой невозможно на глаз различить первоначальное и разархивированное изображения. Хорошей – когда сказать, какое из изображений подвергалось архивации, можно только сравнивая две находящиеся рядом картинки. При дальнейшем увеличении степени сжатия, как правило, становятся заметны побочные эффекты, характерные для данного алгоритма. На практике, даже при отличном сохранении качества, в изображение могут быть внесены регулярные специфические изменения.

3. Ключевые вопросы

1. За счет чего осуществляется сжатие в алгоритме JPEG?
2. Для чего необходимо преобразование цветового пространства изображения?
3. Назовите виды сегментации изображений и их отличия. Какая сегментация используется в алгоритме JPEG?
4. Для чего нужно дискретное косинусное преобразование?
5. На каком этапе сжатия происходят потери качества изображения? Как они связаны с фактором качества?
6. Каким образом проявляются потери качества в алгоритме JPEG?
7. Какие алгоритмы кодирования использует JPEG?
8. Приведите примеры мер потери информации и опишите их недостатки.

4. Домашнее задание

1. По разделу 2 данного пособия изучите и зарисуйте алгоритм JPEG и подготовьте ответы на ключевые вопросы.

2. Подготовьте 3 маски квантования в соответствии с фактором качества N , где N выбирается следующим образом:

- 1) последние 2 цифры зачетной книжки студента;
- 2) $n*2$, где n – № бригады;
- 3) $100 - 2*n$.

5. Лабораторное задание

1. Провести расчет дискретного косинусного преобразования для матрицы исходного изображения, заданной в виде:

$$\begin{matrix} N & N+1 & N+2 & N-1 \\ N+3 & N-1 & N-2 & N \\ N-3 & N+2 & N+1 & N+2 \\ N+1 & N+2 & N+3 & N+1 \end{matrix}$$

где N – две последние цифры зачетной книжки.

2. Проквантовать полученную матрицу.

3. Восстановить матрицу изображения с помощью обратного ДКП.

4. Оценить потери качества изображения. Для четных вариантов определить среднеквадратическое отклонение, для нечетных – отношение сигнал/шум.

5. Запустить с рабочего стола программу MATLAB. В командной строке написать команду *guide*. В появившемся окне перейти на вкладку Open Existing GUI и выбрать файл JPEG_Compression. Нажать кнопку  на панели инструментов.

6. Для начала работы нажать кнопку «Пуск» и ввести маску квантования, рассчитанную в домашнем задании. Полученные на экране данные свести в таблицу вида табл. 5.1:

Таблица 5.1 – Итоги выполнения лабораторной работы

Фактор качества	Размер файла, байт		Коэффициент сжатия	Оценка качества			
	исходного	сжатого		СКО	МО	ОСШ	Субъективная

Сжатие необходимо провести для трех факторов качества, заданных в домашнем задании. Коэффициент сжатия рассчитывается самостоятельно. Субъективная оценка сжатия производится по изображению «Отклонение».

7. Изменяя маску квантования, найти фактор качества, соответствующий субъективной оценке «5», при котором сжатие будет максимально. Данные добавить в табл. 5.1.

6. Содержание протокола

1. Тема и цель лабораторной работы.
2. Структурная схема JPEG-преобразований.
3. Ответы на ключевые вопросы.
4. Три маски квантования в соответствии с домашним и лабораторным заданиями.
5. Итоговая таблица 1 с заполненными *четырьмя* строками в соответствии с п.п. 2 и 3 лабораторного задания.
6. Выводы по проделанной работе.

7. Список литературы

1. Рудий Є.М. Технології передачі дискретних повідомлень. Т.1: Стиснення сигналів: підручник. – Одеса: ОНАЗ ім. О.С. Попова, 2011. – 170 с.

Лабораторна робота 2.3 ПОРІВНЯННЯ РІЗНИХ ВЕЙВЛЕТ-ПЕРЕТВОРЕНЬ ПРИ СТИСНЕННІ

1. Мета роботи

Порівняння вейвлетів Хаара, Добеши, симлетів та койфлетів при стисненні за допомогою їх зображень

2. Ключові положення

У практиці передачі інформації часто потрібно представити сигнал у вигляді сукупності його послідовних наближень. Наприклад, при перегляді й передачі зображень із вибіркою з деякої бази даних можна спочатку передати грубу його версію, а потім (при необхідності) послідовно її уточнювати. При стисненні зображень часто без втрати якості можна забирати із зображення незначущі дрібномасштабні деталі.

Інструментом розподілу (декомпозиції) сигналів на складові, аналізу їх порядку й реконструкції сигналів з певних складових (або з виключенням певних складових, наприклад шумів або малозначимих деталей) є кратномасштабний (багатомасштабний) аналіз. Він дозволяє одержати гарний дозвіл за часом (поганий за частотою) на високих частотах і гарний дозвіл за частотою (поганий за часом) на низьких частотах. Цей підхід стає ефективним, якщо сигнал має короткі високочастотні компоненти й протяжні низькочастотні компоненти. Саме такі сигнали й зустрічаються найчастіше.

Ідея кратномасштабного аналізу полягає в тому, що розкладання сигналу проводиться по ортогональному базису, утвореному зсувами й кратномасш-

табними копіями вейвлетної функції. Згортка сигналу з вейвлетами дозволяє виділити характерні риси сигналу в області локалізації цих вейвлетів, причому, чим більший масштаб має вейвлет, тим більш широка область сигналу буде впливати на результат згортки.

Вейвлети (wavelets) – це узагальнена назва часових функцій, що мають вигляд хвильових пакетів тієї або іншої форми, локалізованих по осі незалежної змінної (t або x) і здатних до зсуву по ній і масштабуванню (стисненню/розтягання). Вейвлети створюються за допомогою спеціальних базових функцій – прототипів, що задають їхній вид і властивості.

Вейвлети й засновані на них інтегральні вейвлет-перетворення були запропоновані на початку 90-х років минулого століття (хоча перший найпростіший тип вейвлета, по суті був описаний Хааром ще в 1909 році) і в наступний час інтенсивно розвивалися.

Набір вейвлетів може наближати складний сигнал або зображення, причому ідеально точно або з деякою погрішністю. Завдяки прекрасному поданню *локальних особливостей* сигналів, принципово відсутньому в рядах Фур'є, і множині видів вейвлети знайшли практичне застосування для аналізу тонких особливостей складних сигналів і зображень для їхнього стиснення й очищення від шуму.

Ряд Фур'є використовує в якості базисних функцій синусоїди. Вони гранично локалізовані в частотній області (перетворюючись на спектрограмі у вертикальну лінію), але дуже погано локалізовані (точніше, взагалі не локалізовані) у часовій області. Протилежний приклад – імпульсна базисна функція

$$\delta_k(t) = \begin{cases} 1, & k = t, \\ 0, & k \neq t. \end{cases}$$

Вона чітко локалізована в часовій області й тому ідеально підходить для подання розривів сигналу. Але ця базисна функція не несе інформації про частоту сигналу й тому погано пристосована для подання сигналів на заданому відрізку часу й тим більше періодичних сигналів.

Вейвлети займають проміжне положення між розглянутими крайніми випадками (синусоїдою й імпульсною функцією). Базисними функціями вейвлетів можуть бути різні функції, у тому числі, що нагадують модульовані імпульсами синусоїди, функції зі стрибками рівня і т.д. Це забезпечує легке подання сигналів з локальними стрибками й розривами наборами вейвлетів того або іншого типу. Майже всі вейвлети не мають аналітичного подання у вигляді однієї формули й можуть даватися ітераційними виразами.

Вейвлет-перетворення

Одна з основних ідей вейвлет-подання сигналів полягає в розбивці сигналу на дві складові – грубу (апроксимуючу) і витончену (ту, що деталізує) – з наступним їхнім дробленням з метою зміни рівня декомпозиції сигналу.

Число використовуваних при розкладанні сигналу вейвлетів задає *рівень декомпозиції* сигналу. При цьому за нульовий рівень декомпозиції ухвалюється

сам сигнал, а рівні декомпозиції утворюють спадаюче *вейвлет-дерево* того або іншого вигляду. Точність подання сигналу в міру переходу на більш низькі рівні декомпозиції знижується, але зате з'являється можливість вейвлет-фільтрації сигналів, видалення із сигналів шумів й ефективної компресії сигналів.

Пряме вейвлет-перетворення означає розкладання довільного вхідного сигналу на принципово новий базис у вигляді сукупності хвильових пакетів – вейвлетів, які характеризуються чотирма принципово важливими властивостями:

- мають вигляд коротких, локалізованих у часі (або в просторі) хвильових пакетів з нульовим значенням інтеграла;
- мають можливість зсуву за часом;
- здатні до масштабування (стиснення/розтягання);
- мають обмежений (або локальний) частотний спектр.

В основі *безперервного* вейвлет-перетворення лежить використання двох безперервних й інтегровальних по всій осі t (або x) функцій:

- вейвлет-функція $\psi(t)$ з нульовим значенням інтеграла $\left(\int_{-\infty}^{\infty} \psi(t) dt = 0 \right)$,

що визначає деталі сигналу, й функція, що породжує деталізуючі коефіцієнти:

$$d_{j-1} = \int s(t) \overline{\psi_{j-1}(t)} dt \quad (1)$$

- масштабуюча функція $\phi(t)$ з одиничним значенням інтеграла $\left(\int_{-\infty}^{\infty} \phi(t) dt = 1 \right)$, що визначає грубе наближення (апроксимацію) сигналу, й функцію, що породжує коефіцієнти апроксимації:

$$a_{j-1} = \int s(t) \overline{\phi_{j-1}(t)} dt. \quad (2)$$

ϕ -функції $\phi(t)$ властиві далеко не всім вейвлетам, а тільки тим, які належать до ортогональних, тобто таким, у яких інтеграл від добутку будь-яких двох функцій ряду дорівнює нулю.

Властивість ортогональності помітно полегшує аналіз, дає можливість реконструкції сигналів (повного й точного відтворення) і дозволяє реалізувати алгоритми швидких вейвлет-перетворень.

Вейвлети Хаара

Один з перших відомих ортогональних вейвлетів - *вейвлет Хаара* (haar). Функція ψ у нього має вигляд прямокутних імпульсів меандру (значення 1 в інтервалі $[0, 0.5]$ і -1 в інтервалі $[0.5, 1]$). Функція ϕ (див. рис. 2.1) має значення 1 в інтервалі $[0, 1]$ і 0 за межами цього інтервалу. Вейвлети Хаара добре локалізовані в просторі, але не дуже добре локалізовані в частотній області, оскільки меандр має широкий спектр частот (теоретично нескінченний).

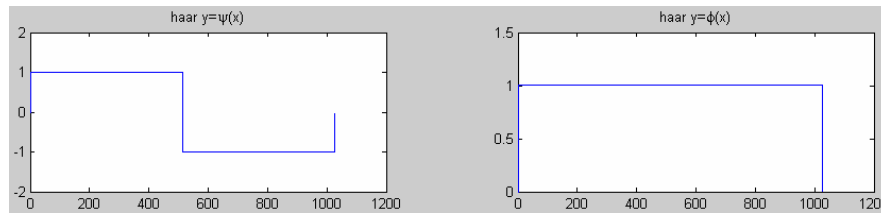


Рисунок 2.1 - Функції psi і phi вейвлета Хаара

Перетворення Хаара в загальному вигляді для одномірного сигналу (відліків) спрощується в порівнянні з формулами (1 і 2) і виглядає в такий спосіб. Нехай s є одномірний дискретний сигнал. Кожній парі елементів з індексами $2j$ і $2j + 1$, $j \in Z$, поставимо у відповідність два значення:

$$a_j = \frac{s_{2j} + s_{2j+1}}{2}, \quad d_j = \frac{s_{2j} - s_{2j+1}}{2}. \quad (3)$$

До сигналу a_j можна застосувати аналогічну операцію й так само одержати два сигнали, один з яких є огрубленою версією a_j , а інший містить інформацію, що деталізує, необхідну для відновлення a_j .

Зворотнє перетворення Хаара виглядає в такий спосіб:

$$s_{2j} = a_j + d_j, \quad s_{2j+1} = a_j - d_j. \quad (4)$$

Розберемо конкретний приклад: нехай s_i : (220, 211, 212, 218, 217, 214, 210, 202). Попарно усереднивши коефіцієнти сигналу одержимо огрублену версію a_1 , а інформація, що деталізує, d_1 виходить у результаті ділення попарних різниць навпіл. Для переходу на наступний рівень декомпозиції необхідно повторити операцію, розглядаючи a_1 як s_i . Рекурсивне повторення процедури для середніх значень дає нам повне розкладання:

Рівень декомпозиції	Огрублена версія	Інформація, що деталізує
1	(220, 211, 212, 218, 217, 214, 210, 202)	
2	(215.5, 215, 215.5, 206)	(4.5, -3, 1.5, 4)
3	(215.25, 210.75)	(0.25, 4.75)
4	(213)	(2.25)

Таким чином, вейвлет-перетворення даного сигналу має вигляд: (213, 2.25, 0.25, 4.75, 4.5, -3, 1.5, 4)

На прикладі перетворення Хаара добре видна структура вейвлет-перетворення дискретного сигналу. На кожному кроці перетворення сигнал розпадається на дві складові: наближення з більш низьким дозволом – апроксимацію, що й деталізує інформацію.

Двовимірне вейвлет-перетворення

Розглянемо двовимірний сигнал s -матрицю кінцевого або нескінченного розміру. Застосуємо до кожного рядка матриці один крок одномірного вейвлет-перетворення. У результаті вийде дві матриці, рядки яких містять апроксимовану складову, й складову, що деталізує, рядків вихідної матриці. До кожного стовпця обох матриць також застосуємо крок одномірного

перетворення. У результаті виходить чотири матриці. Перша є апроксимуючою складовою вихідного сигналу (огрубленою версією), інші три містять інформацію, що деталізує, – вертикальну, горизонтальну й діагональну. Таким чином, крок двовимірного перетворення звівся до композиції одновимірних перетворень. Тому реалізація двовимірного перетворення не вимагає ніяких додаткових операцій.

Наприклад, для зображення 512*512 пікселів одержимо після першого перетворення 4 матриці розміром 256*256 елементів (рис. 2.2).

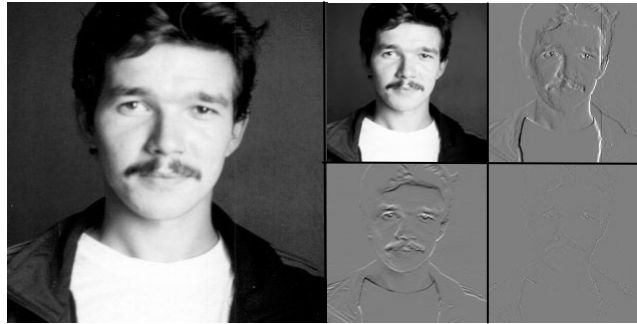


Рисунок 2.2 – Реалізація двовимірного вейвлет-перетворення

У першій, як легко догадатися, буде зберігатися зменшена копія зображення. У другій – усереднені різниці пар значень пікселів по горизонталі. У третій – усереднені різниці пар значень пікселів по вертикалі. У четвертій – усереднені різниці значень пікселів по діагоналі. За аналогією із двовимірним випадком ми можемо повторити наше перетворення й одержати замість першої матриці 4 матриці розміром 128*128. Повторивши наше перетворення втретє, ми одержимо в підсумку: 4 матриці 64*64, 3 матриці 128*128 і 3 матриці 256*256.

Наприклад, для матриці

j	0	1	2	3	a		та	d	
Рядок 1	15	18	12	16	$(15+18)/2$	$(12+16)/2$		$(15-18)/2$	$(12-16)/2$
Рядок 2	17	13	16	18	$(17+13)/2$	$(16+18)/2$		$(17-13)/2$	$(16-18)/2$
Рядок 3	16	14	17	17	$(16+14)/2$	$(17+17)/2$		$(16-14)/2$	$(17-17)/2$
Рядок 4	14	15	17	16	$(14+15)/2$	$(17+16)/2$		$(14-15)/2$	$(17-16)/2$

або

j	0	1		0	1
Рядок 1	16,5	14	та	-1,5	-2
Рядок 2	15	17		2	-1
Рядок 3	15	17		1	0
Рядок 4	14,5	16,5		-0,5	0,5

Далі застосовуємо перетворення Хаара до кожного стовпця матриць:

j	Стовпець 1	Стовпець 2	Стовпець 3	Стовпець 4
0	16,5	14	-1,5	-2
1	15	17	2	-1
2	15	17	1	0
3	14,5	16,5	-0,5	0,5

$$\begin{array}{c}
 j \quad a \\
 0 \left| \begin{array}{cc} (16,5+15)/2 & (14+17)/2 \end{array} \right| \text{та} \left| \begin{array}{cc} (-1,5+2)/2 & (-2-1)/2 \end{array} \right| \\
 1 \left| \begin{array}{cc} (15+14,5)/2 & (17+16,5)/2 \end{array} \right| \left| \begin{array}{cc} (1-0,5)/2 & (0+0,5)/2 \end{array} \right| \\
 d \\
 2 \left| \begin{array}{cc} (16,5-15)/2 & (14-17)/2 \end{array} \right| \text{та} \left| \begin{array}{cc} (-1,5-2)/2 & (-2+1)/2 \end{array} \right| \\
 3 \left| \begin{array}{cc} (15-14,5)/2 & (17-16,5)/2 \end{array} \right| \left| \begin{array}{cc} (1+0,5)/2 & (0-0,5)/2 \end{array} \right|
 \end{array}$$

або

Низькочастотна складова		та	Вертикальне відхилення	
15,75	15,5		0,25	-1,5
14,75	16,75		0,25	0,25
Горизонтальне відхилення		та	Діагональне відхилення	
0,75	-1,5		-1,75	0,5
0,25	0,25		0,75	-0,25

Відновлення вихідної матриці відбувається у зворотному порядку відповідно до формул (4).

Вейвлети Добеши і їх похідні

Тому що для повної реконструкції сигналу можуть бути застосовані тільки ортогональні вейвлети, а вейвлет Хаара має «негладкість», Інґрід Добеши запропонувала використовувати функції, що обчислюються ітераційним шляхом, які згодом були названі *вейвлетами Добеши*. Вони мають наступні властивості: ортогональність, компактний носій (тобто середнє значення функції дорівнює нулю й функція швидко убуває на нескінченності).

Вейвлети Добеши не можуть мати симетричність, що звужує їхнє використання. Однак можна спробувати наблизитися, наскільки можливо, до симетрії. Такі вейвлети, отримані з вейвлетів Добеши, називаються *симлетами*.

Питання про побудову вейвлетів, у яких нульові моменти має не тільки функція вейвлета $\psi(x)$, але, що й породжує вейвлет $\phi(x)$ був уперше поставлений Р. Койфманом, тому такі вейвлети називаються *койфлетами*. Наявність нульових моментів в вейвлетах, що породжують, полегшує аналіз і вейвлет-перетворення. Койфлети несиметричні, однак вони більш симетричні, ніж вейвлети Добеши.

Стиснення зображень за допомогою вейвлет-перетворення (JPEG-2000)

Алгоритм стиснення за допомогою вейвлет-перетворення ідентичний алгоритму JPEG. Відмінність полягає в застосуванні замість ДКП вейвлет-перетворення (Подумайте, як буде виглядати схема алгоритму). Вейвлет-перетворення має наступні переваги:

– при більших коефіцієнтах стиснення за допомогою вейвлет-перетворення зображення стає нечітким (розмитим), що сприймається оком людини набагато краще, ніж блокова структура JPEG;

– можливість використання різних функцій у якості базисних, а також створення нових вейвлетів для різних типів сигналів для більш точного наближення до них;

– можливість поступового перегляду зображення в процесі завантаження зображення по мережі.

Втрати якості в даному алгоритмі відбуваються за рахунок інформації, що деталізує, у такий спосіб. Усі значення коефіцієнтів, менші граничного, обнулюються, а інші округляються до цілого. Чим більше граничне значення, тим більше втрати інформації, що деталізує, і, отже, тим більш розмите зображення.

3. Ключові питання

1. У чому полягає ідея стиснення за допомогою вейвлетів?
2. Які загальні властивості мають вейвлети Добеши, симлети й койфлети?
3. У чому полягає одномірне перетворення Хаара?
4. Яким способом зберігається інформація про зображення при двовимірному перетворенні Хаара?
5. Назвіть переваги вейвлет-перетворення перед ДКП.

4. Домашнє завдання

1. За розділом 2 даного посібника вивчити принципи стиснення за допомогою вейвлет-перетворення.
2. Дайте відповіді на ключові питання.
3. Накресліть в протоколі лабораторної роботи структурну схему алгоритму стиснення за допомогою вейвлет-перетворення (у посібнику вона не представлена).

5. Лабораторне завдання

1. Провести розрахунок двовимірного перетворення Хаара для матриці вихідного зображення, що задано в лабораторній роботі 2.2.
2. Округлити коефіцієнти в отриманій матриці до цілих чисел.
3. Відновити вихідну матрицю з округлених коефіцієнтів та розрахувати для парних варіантів середньоквадратичне відхилення, для непарних – відношення сигнал/шум.
4. Запустити з робочого стола програму MATLAB. У командному рядку написати команду *guide*. У вікні, що з'явилося, перейти на вкладку Open Existing GUI і вибрати файл Wavelet_Compression. Натиснути кнопку  на панелі інструментів.

5. Для початку роботи натиснути кнопку «Пуск», вибрати вейвлет зі списку. Отримані на екрані дані звести в таблицю вигляду табл. 5.1.

Таблиця 5.1 – Підсумки виконання лабораторної роботи

Вейвлет	Розмір файлу, байт		Коефіцієнт стиснення	Оцінка якості	
	вхідного	стислого		СКВ	Суб'єктивна

Стиснення необхідно провести для всіх представлених у списку видів вейвлетів. Коефіцієнт стиснення розраховується самостійно. Суб'єктивна оцінка стиснення проводиться шляхом порівняння вхідного й стислого зображень.

6. Зміст протоколу

1. Тема й мета лабораторної роботи.
2. Структурна схема алгоритму стиснення зображень за допомогою вейвлет-перетворень.
2. Відповіді на ключові питання.
3. Підсумкова таблиця 1 із заповненими чотирма рядками відповідно до п. 2 лабораторного завдання.
4. Висновки за результатами роботи.

7. Список літератури

1. Рудий Є.М. Технології передачі дискретних повідомлень. Т.1: Стиснення сигналів: підручник. – Одеса: ОНАЗ ім. О.С. Попова, 2011. – 170 с.

Лабораторная работа 2.3 СРАВНЕНИЕ РАЗЛИЧНЫХ ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЙ ПРИ СЖАТИИ

1. Цель работы

Сравнение вейвлетов Хаара, Добеши, симлетов и койфлетов при сжатии с помощью них изображений

2. Ключевые положения

В практике передачи информации часто требуется представить сигнал в виде совокупности его последовательных приближений. Например, при просмотре и передаче изображений с выборкой из некоторой базы данных можно сначала передать грубую его версию, а затем (при необходимости) последовательно ее уточнять. При сжатии изображений часто без потери качества можно убирать из изображения незначимые мелкомасштабные детали.

Инструментом разделения (декомпозиции) сигналов на составляющие, анализа их порядка и реконструкции сигналов из определенных составляющих (или с исключением определенных составляющих, например шумов или малозначимых деталей) является кратномасштабный (многомасштабный) анализ. Он позволяет получить хорошее разрешение по времени (плохое по частоте) на высоких частотах и хорошее разрешение по частоте (плохое по времени) на низких частотах. Этот подход становится эффективным, если сигнал имеет короткие высокочастотные компоненты и протяженные низкочастотные компоненты. Именно такие сигналы и встречаются чаще всего.

Идея кратномасштабного анализа заключается в том, что разложение сигнала производится по ортогональному базису, образованному сдвигами и кратномасштабными копиями вейвлетной функции. Свертка сигнала с вейвлетами позволяет выделить характерные особенности сигнала в области локализации этих вейвлетов, причем, чем больший масштаб имеет вейвлет, тем более широкая область сигнала будет оказывать влияние на результат свертки.

Вейвлеты (wavelets) – это обобщенное название временных функций, имеющих вид волновых пакетов той или иной формы, локализованных по оси независимой переменной (t или x) и способных к сдвигу по ней и масштабированию (сжатию/растяжению). Вейвлеты создаются с помощью специальных *базовых функций* – прототипов, задающих их вид и свойства.

Вейвлеты и основанные на них интегральные вейвлет-преобразования были предложены в начале 90-х годов прошлого века (хотя первый простейший тип вейвлета, по существу был описан Хааром еще в 1909 году) и в последующее время интенсивно развивались.

Набор вейвлетов может приближать сложный сигнал или изображение, причем идеально точно или с некоторой погрешностью. Благодаря прекрасному представлению *локальных особенностей* сигналов, принципиально отсутствующему у рядов Фурье, и множеству видов вейвлеты нашли практическое применение для анализа тонких особенностей сложных сигналов и изображений для их сжатия и очистки от шума.

Ряд Фурье использует в качестве базисных функций синусоиды. Они предельно локализованы в частотной области (превращаясь на спектрограмме в вертикальную линию), но очень плохо локализованы (точнее, вообще не локализованы) во временной области. Противоположный пример – импульсная базисная функция

$$\delta_k(t) = \begin{cases} 1, & k = t, \\ 0, & k \neq t. \end{cases}$$

Она четко локализована во временной области и потому идеально подходит для представления разрывов сигнала. Но эта базисная функция не несет информации о частоте сигнала и потому плохо приспособлена для представления сигналов на заданном отрезке времени и тем более периодических сигналов.

Вейвлеты занимают промежуточное положение между рассмотренными крайними случаями (синусоидой и импульсной функцией). Базисными функциями вейвлетов могут быть различные функции, в том числе напоминающие модулированные импульсами синусоиды, функции со скачками уровня и т. д. Это обеспечивает легкое представление сигналов с локальными скачками и разрывами наборами вейвлетов того или иного типа. Почти все вейвлеты не имеют аналитического представления в виде одной формулы и могут даваться итерационными выражениями.

Вейвлет-преобразование

Одна из основополагающих идей вейвлет-представления сигналов заключается в разбивке сигнала на две составляющие – грубую (аппроксимирующую) и утонченную (детализирующую) – с последующим их дроблением с целью изменения уровня декомпозиции сигнала.

Число используемых при разложении сигнала вейвлетов задает *уровень декомпозиции* сигнала. При этом за нулевой уровень декомпозиции принимается сам сигнал, а уровни декомпозиции образуют ниспадающее *вейвлет-дерево* того или иного вида. Точность представления сигнала по мере перехода на более низкие уровни декомпозиции снижается, но зато появляется возможность вейвлет-фильтрации сигналов, удаления из сигналов шумов и эффективной компрессии сигналов.

Прямое вейвлет-преобразование означает разложение произвольного входного сигнала на принципиально новый базис в виде совокупности волновых пакетов – вейвлетов, которые характеризуются четырьмя принципиально важными свойствами:

- имеют вид коротких, локализованных во времени (или в пространстве) волновых пакетов с нулевым значением интеграла;
- обладают возможностью сдвига по времени;
- способны к масштабированию (сжатию/растяжению);
- имеют ограниченный (или локальный) частотный спектр.

В основе *непрерывного* вейвлет-преобразования лежит использование двух непрерывных и интегрируемых по всей оси t (или x) функций:

- вейвлет-функция $\psi(t)$ с нулевым значением интеграла $\left(\int_{-\infty}^{\infty} \psi(t) dt = 0 \right)$, определяющая детали сигнала и порождающая детализирующие коэффициенты:

$$d_{j-1} = \int s(t) \overline{\psi_{j-1}(t)} dt \quad (1)$$

- масштабирующая функция $\phi(t)$ с единичным значением интеграла $\left(\int_{-\infty}^{\infty} \phi(t) dt = 1 \right)$, определяющая грубое приближение (аппроксимацию) сигнала и порождающая коэффициенты аппроксимации:

$$a_{j-1} = \int s(t) \overline{\varphi_{j-1}(t)} dt. \quad (2)$$

Phi-функции $\varphi(t)$ присущи далеко не всем вейвлетам, а только тем, которые относятся к ортогональным, т.е. таким, у которых интеграл от произведения любых двух функций ряда равен нулю.

Свойство ортогональности заметно облегчает анализ, дает возможность реконструкции сигналов (полного и точного воспроизведения) и позволяет реализовать алгоритмы быстрых вейвлет-преобразований.

Вейвлеты Хаара

Один из первых известных ортогональных вейвлетов - *вейвлет Хаара* (haar). Функция ψ у него имеет вид прямоугольных импульсов меандра (значение 1 в интервале $[0,0.5]$ и -1 в интервале $[0.5,1]$). Функция ϕ (см. рис. 2.1) имеет значение 1 в интервале $[0,1]$ и 0 за пределами этого интервала. Вейвлеты Хаара хорошо локализованы в пространстве, но не очень хорошо локализованы в частотной области, поскольку меандр имеет широкий спектр частот (теоретически бесконечный).

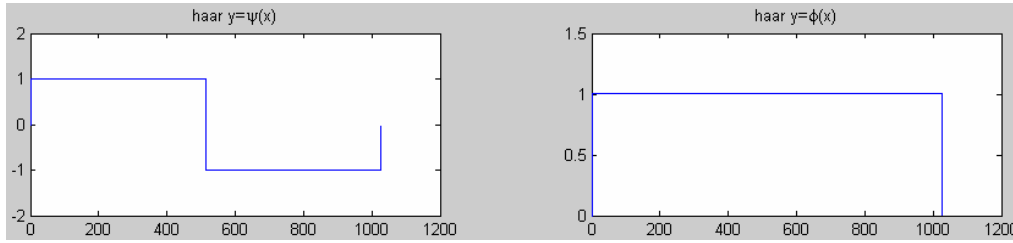


Рисунок 2.1 - Функции ψ и ϕ вейвлета Хаара

Преобразование Хаара в общем виде для одномерного сигнала (отсчетов) упрощается по сравнению с формулами (1 и 2) и выглядит следующим образом. Пусть имеется одномерный дискретный сигнал s . Каждой паре элементов с индексами $2j$ и $2j + 1$, $j \in \mathbb{Z}$, поставим в соответствие два значения:

$$a_j = \frac{s_{2j} + s_{2j+1}}{2}, \quad d_j = \frac{s_{2j} - s_{2j+1}}{2}. \quad (3)$$

К сигналу a_j можно применить аналогичную операцию и так же получить два сигнала, один из которых является огрубленной версией a_j , а другой содержит детализирующую информацию, необходимую для восстановления a_j .

Обратное преобразование Хаара выглядит следующим образом:

$$s_{2j} = a_j + d_j, \quad s_{2j+1} = a_j - d_j. \quad (4)$$

Разберем конкретный пример: пусть s_i : (220, 211, 212, 218, 217, 214, 210, 202). Попарно усреднив коэффициенты сигнала получим огрубленную версию a_1 , а детализирующая информация d_1 получается в результате деления попарных разностей пополам. Для перехода на следующий уровень декомпозиции необходимо повторить операцию, рассматривая a_1 как s_i . Рекурсивное повторение процедуры для средних значений дает нам полное разложение:

Уровень декомпозиции	Огрубленная версия	Детализирующая информация
1	(220, 211, 212, 218, 217, 214, 210, 202)	
2	(215.5, 215, 215.5, 206)	(4.5, -3, 1.5, 4)
3	(215.25, 210.75)	(0.25, 4.75)
4	(213)	(2.25)

Таким образом, вейвлет-преобразование данного сигнала имеет вид: (213, 2.25, 0.25, 4.75, 4.5, -3, 1.5, 4)

На примере преобразования Хаара хорошо видна структура вейвлет-преобразования дискретного сигнала. На каждом шаге преобразования сигнал распадается на две составляющие: приближение с более низким разрешением – аппроксимацию и детализирующую информацию.

Двумерное вейвлет-преобразование

Рассмотрим двумерный сигнал s -матрицу конечного или бесконечного размера. Применим к каждой строке матрицы один шаг одномерного вейвлет-преобразования. В результате получится две матрицы, строки которых содержат аппроксимированную и детализирующую составляющие строк исходной матрицы. К каждому столбцу обеих матриц также применим шаг одномерного преобразования. В результате получается четыре матрицы. Первая является аппроксимирующей составляющей исходного сигнала (огрубленной версией), остальные три содержат детализирующую информацию – вертикальную, горизонтальную и диагональную. Таким образом, шаг двумерного преобразования свелся к композиции одномерных преобразований. Поэтому реализация двумерного преобразования не требует никаких дополнительных операций.

Например, для изображения 512*512 пикселей получим после первого преобразования 4 матрицы размером 256*256 элементов (рис. 2.2).

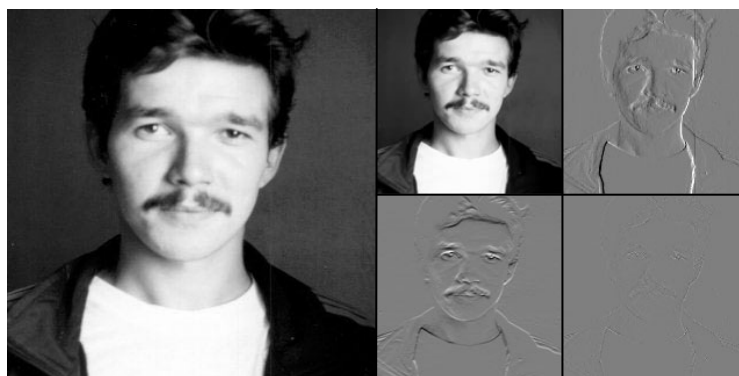


Рисунок 2.2 – Реализация двумерного вейвлет-преобразования

В первой, как л егко догадаться, будет храниться уменьшенная копия изображения. Во второй – усредненные разности пар значений пикселей по горизонтали. В третьей – усредненные разности пар значений пикселей по вертикали. В четвертой – усредненные разности значений пикселей по

диагонали. По аналогии с двумерным случаем мы можем повторить наше преобразование и получить вместо первой матрицы 4 матрицы размером 128×128 . Повторив наше преобразование в третий раз, мы получим в итоге: 4 матрицы 64×64 , 3 матрицы 128×128 и 3 матрицы 256×256 .

Например, для матрицы

$$\begin{array}{c} j \\ \text{Строка } 1 \\ \text{Строка } 2 \\ \text{Строка } 3 \\ \text{Строка } 4 \end{array} \begin{array}{c|c|c|c} 0 & 1 & 2 & 3 \\ \hline 15 & 18 & 12 & 16 \\ \hline 17 & 13 & 16 & 18 \\ \hline 16 & 14 & 17 & 17 \\ \hline 14 & 15 & 17 & 16 \end{array} \begin{array}{c} a \\ \hline (15+18)/2 \quad (12+16)/2 \\ (17+13)/2 \quad (16+18)/2 \\ (16+14)/2 \quad (17+17)/2 \\ (14+15)/2 \quad (17+16)/2 \end{array} \begin{array}{c} d \\ \hline (15-18)/2 \quad (12-16)/2 \\ (17-13)/2 \quad (16-18)/2 \\ (16-14)/2 \quad (17-17)/2 \\ (14-15)/2 \quad (17-16)/2 \end{array}$$

или

$$\begin{array}{c} j \\ \text{Строка } 1 \\ \text{Строка } 2 \\ \text{Строка } 3 \\ \text{Строка } 4 \end{array} \begin{array}{c|c} 0 & 1 \\ \hline 16,5 & 14 \\ \hline 15 & 17 \\ \hline 15 & 17 \\ \hline 14,5 & 16,5 \end{array} \begin{array}{c} 0 & 1 \\ \hline -1,5 & -2 \\ \hline 2 & -1 \\ \hline 1 & 0 \\ \hline -0,5 & 0,5 \end{array}$$

Далее применяем преобразование Хаара к каждому столбцу матриц:

$$\begin{array}{c} j \\ \text{Столбец } 1 \\ \text{Столбец } 2 \\ \text{Столбец } 3 \\ \text{Столбец } 4 \end{array} \begin{array}{c|c|c|c} 0 & 1 & 2 & 3 \\ \hline 16,5 & 14 & -1,5 & -2 \\ \hline 15 & 17 & 2 & -1 \\ \hline 15 & 17 & 1 & 0 \\ \hline 14,5 & 16,5 & -0,5 & 0,5 \end{array}$$

$$\begin{array}{c} j \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c|c} a \\ \hline (16,5+15)/2 \quad (14+17)/2 \\ (15+14,5)/2 \quad (17+16,5)/2 \\ \hline d \\ \hline (16,5-15)/2 \quad (14-17)/2 \\ (15-14,5)/2 \quad (17-16,5)/2 \end{array} \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c|c} d \\ \hline (-1,5+2)/2 \quad (-2-1)/2 \\ (1-0,5)/2 \quad (0+0,5)/2 \\ \hline (-1,5-2)/2 \quad (-2+1)/2 \\ (1+0,5)/2 \quad (0-0,5)/2 \end{array}$$

или

$$\begin{array}{c} \text{Низкочастотная составляющая} \\ 15,75 \\ 14,75 \\ \text{Горизонтальное отклонение} \\ 0,75 \\ 0,25 \end{array} \begin{array}{c} 15,5 \\ 16,75 \\ -1,5 \\ 0,25 \end{array} \begin{array}{c} \text{Вертикальное отклонение} \\ 0,25 \\ 0,25 \\ \text{Диагональное отклонение} \\ -1,75 \\ 0,75 \end{array}$$

Восстановление исходной матрицы происходит в обратном порядке в соответствии с формулами (4).

Вейвлеты Добеши и их производные

Так как для полной реконструкции сигнала могут быть применены только ортогональные вейвлеты, а вейвлет Хаара обладает «негладкостью», Ингрид Добеши предложила использовать функции, вычисляемые итерационным путем, впоследствии названные *вейвлетами Добеши*. Они обладают

следующими свойствами: ортогональностью, компактным носителем (т.е. среднее значение функции равно нулю и функция быстро убывает на бесконечности).

Вейвлеты Добеши не могут обладать симметричностью, что сужает их использование. Однако можно попробовать приблизиться, насколько возможно, к симметрии. Такие вейвлеты, полученные из вейвлетов Добеши, называются *симлетами*.

Вопрос о построении вейвлетов, у которых нулевые моменты имеет не только функция вейвлета $\psi(x)$, но и порождающий вейвлет $\varphi(x)$ был впервые поставлен Р. Койфманом, поэтому такие вейвлеты называются *койфлетами*. Наличие нулевых моментов в порождающих вейвлетах облегчает анализ и вейвлет-преобразование. Койфлеты несимметричны, однако они более симметричны, чем вейвлеты Добеши.

Сжатие изображений с помощью вейвлет-преобразования (JPEG-2000)

Алгоритм сжатия с помощью вейвлет-преобразования идентичен алгоритму JPEG. Различие состоит в применении вместо ДКП вейвлет-преобразования (Подумайте, как будет выглядеть схема алгоритма). Вейвлет-преобразование имеет следующие преимущества:

- при больших коэффициентах сжатия с помощью вейвлет-преобразования изображение становится нечетким (размытым), что воспринимается глазом человека гораздо лучше, чем блочная структура JPEG;
- возможность использования различных функций в качестве базисных, а также создания новых вейвлетов для различных типов сигналов для более точного приближения к ним;
- возможность постепенного просмотра изображения в процессе загрузки изображения по сети.

Потери качества в данном алгоритме происходят за счет детализирующей информации следующим образом. Все значения коэффициентов, меньше порогового, обнуляются, а остальные округляются до целого. Чем больше пороговое значение, тем больше потери детализирующей информации, и, следовательно, тем более размыто изображение.

3. Ключевые вопросы

1. В чем заключается идея сжатия с помощью вейвлетов?
2. Какими общими свойствами обладают вейвлеты Добеши, симлеты и койфлеты?
3. В чем заключается одномерное преобразование Хаара?
4. Каким образом хранится информация об изображении при двумерном преобразовании Хаара?
5. Назовите преимущества вейвлет-преобразования перед ДКП.

4. Домашнее задание

1. По разделу 2 данного пособия изучить принципы сжатия с помощью вейвлет-преобразования.
2. Ответьте на ключевые вопросы.
3. Начертите в протоколе лабораторной работы структурную схему алгоритма сжатия с помощью вейвлет-преобразования (в пособии она не представлена).

5. Лабораторное задание

1. Провести расчет двумерного преобразования Хаара для матрицы исходного изображения, заданной в лабораторной работе 2.2.
2. Округлить коэффициенты в полученной матрице до целых чисел.
3. Восстановить исходную матрицу из округленных коэффициентов и рассчитать для четных вариантов среднеквадратическое отклонение, для нечетных – отношение сигнал/шум.
4. Запустить с рабочего стола программу MATLAB. В командной строке написать команду *guide*. В появившемся окне перейти на вкладку Open Existing GUI и выбрать файл Wavelet_Compression. Нажать кнопку  на панели инструментов.
5. Для начала работы нажать кнопку «Пуск», выбрать вейвлет из списка. Полученные на экране данные свести в таблицу вида табл. 5.1.

Таблица 5.1 – Итоги выполнения лабораторной работы

Вейвлет	Размер файла, байт		Коэффициент сжатия	Оценка качества	
	исходного	сжатого		СКО	Субъективная

Сжатие необходимо провести для всех представленных в списке видов вейвлетов. Коэффициент сжатия рассчитывается самостоятельно. Субъективная оценка сжатия производится путем сравнения исходного и сжатого изображений.

6. Содержание протокола

1. Тема и цель лабораторной работы.
2. Структурная схема алгоритма сжатия изображений с помощью вейвлет-преобразований.
2. Ответы на ключевые вопросы.
3. Итоговая таблица 1 с заполненными четырьмя строками в соответствии с п. 2 лабораторного задания.
4. Выводы по проделанной работе.

7. Список литературы

1. Рудий Є.М. Технології передачі дискретних повідомлень. Т.1: Стиснення сигналів: підручник. – Одеса: ОНАЗ ім. О.С. Попова, 2011. – 170 с.

Лабораторна робота 2.4 ОЧИЩЕННЯ СИГНАЛІВ ВІД ШУМІВ ЗА ДОПОМОГОЮ ВЕЙВЛЕТІВ

1. Мета роботи

Вивчення принципів очищення сигналів від шумів за допомогою вейвлет-преобразований, оцінка якості відновленого сигналу і визначення найкращого порядку вейвлета Добеши, порогового методу, а також кількості рівнів розкладання.

2. Ключові положення

Типовий метод пригнічення шумів – видалення високочастотних складових із спектру сигналу. Стосовно вейвлетних розкладань це може бути реалізовано безпосередньо видаленням деталізуючих коефіцієнтів високочастотних рівнів. Але вейвлети мають в цьому відношенні ширші можливості. Шумові компоненти, і особливо великі випадкові викиди значень сигналів, можна також розглядати у вигляді безлічі локальних особливостей сигналів. Задаючи деякий поріг для їх рівня і зрізуючи по ньому деталізуючі коефіцієнти, можна не лише зменшувати рівень шумів, але і встановлювати порогові обмеження на декількох рівнях розкладання з урахуванням конкретних характеристик шумів і сигналів для різних типів вейвлетів. Це дозволяє створювати адаптивні системи очищення сигналів від шумів залежно від їх особливостей. Ці технології і розглядатимуться в даній темі.

Операція стиснення сигналів з видаленням малозначимих значень вейвлет-коефіцієнтів також виконується на основі певних порогових обмежень їх значень, і багато в чому практично тотожна операціям видалення шумів.

При очищенні сигналів від шумів і стисненні використовується швидко вейвлет-перетворення (ШВП). При цьому завжди слід враховувати, що якщо повний розмір сигналу складає M -відліків, а максимальний рівень розкладання рівний N , то для забезпечення нормальної роботи ШВП відношення $M/2^N$ має бути цілим числом, що забезпечує ціле число коефіцієнтів на останньому рівні розкладання. Якщо ця умова не виконується, рекомендується доповнювати масив відліків нульовими або будь-якими іншими значеннями.

Модель зашумленого сигналу зазвичай приймається аддитивною: $s(n) = f(n) + k \cdot e(n)$ з рівномірним кроком по аргументу n , де $f(n)$ – корисна інформаційна складова, $e(n)$ – шумовий сигнал, наприклад, білий шум певного рівня з середнім нульовим значенням. Процедура видалення шуму виконується з використанням ортогональних вейвлетів і включає наступні операції:

- **Вейвлет-розкладання сигналу $s(n)$ до рівня N .** Значення рівня N визначається частотним спектром інформаційної частини $f(n)$ сигналу, яку

бажано залишати в рядах коефіцієнтів апроксимацій. Тип і порядок вейвлета може істотно впливати на якість очищення сигналу від шуму в залежності як від форми сигналів $f(n)$, так і від кореляційних характеристик шумів.

– **Завдання типу і порогових рівнів** очищення за відомими апіорними даними про характер шумів або за певними критеріями шумів у вхідному сигналі. Порогові рівні очищення можуть бути гнучкими (залежно від номера рівня розкладання) або жорсткими (глобальними).

– **Модифікація коефіцієнтів деталізації** вейвлет-розкладання відповідно до встановлених умов очищення.

– **Відновлення сигналу** на основі коефіцієнтів апроксимації та модифікованих коефіцієнтів деталізації.

Вибір використовуваного вейвлета і глибини розкладання, в загальному випадку, залежить від властивостей конкретного сигналу. Можна дати лише декілька рекомендацій.

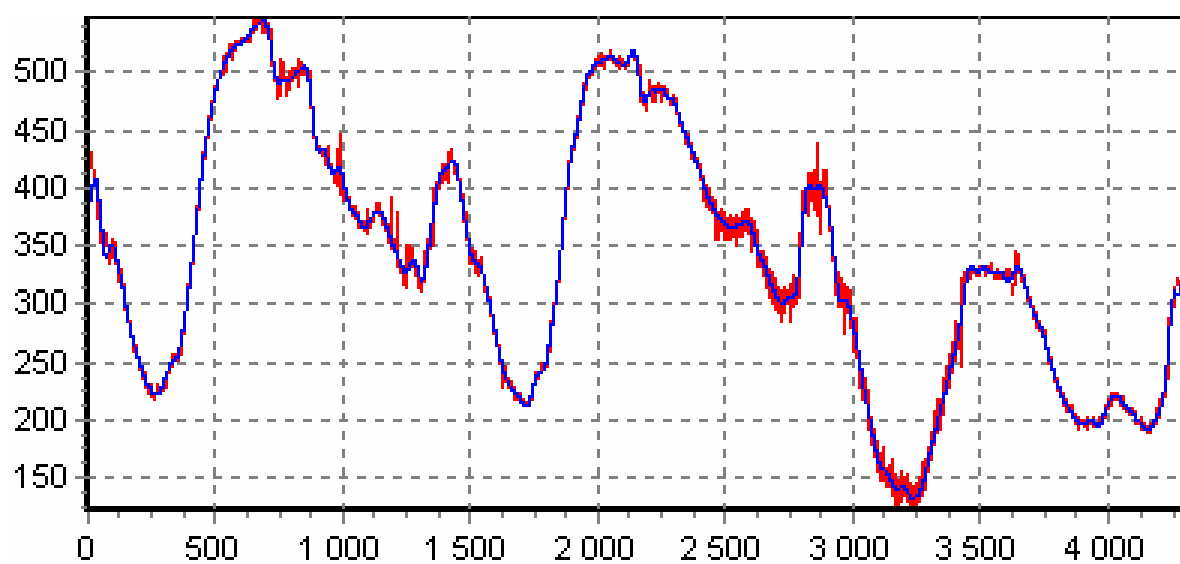
Гладші вейвлети створюють гладшу апроксимацію сигналу, і навпаки - "короткі" вейвлети краще відстежують списи функції, що апроксимується.

Глибина розкладання N впливає на масштаб деталей, що відсіваються. Іншими словами, при збільшенні глибини розкладання модель віднімає шум все більшого рівня, поки не настане "переукрупнення" масштабу деталей і перетворення почне спотворювати форму початкового сигналу. Цікаво, що при подальшому збільшенні глибини розкладання перетворення починає формувати згладжену версію початкового сигналу, тобто фільтрується не лише шум, але і деякі локальні особливості (викиди) початкового сигналу.

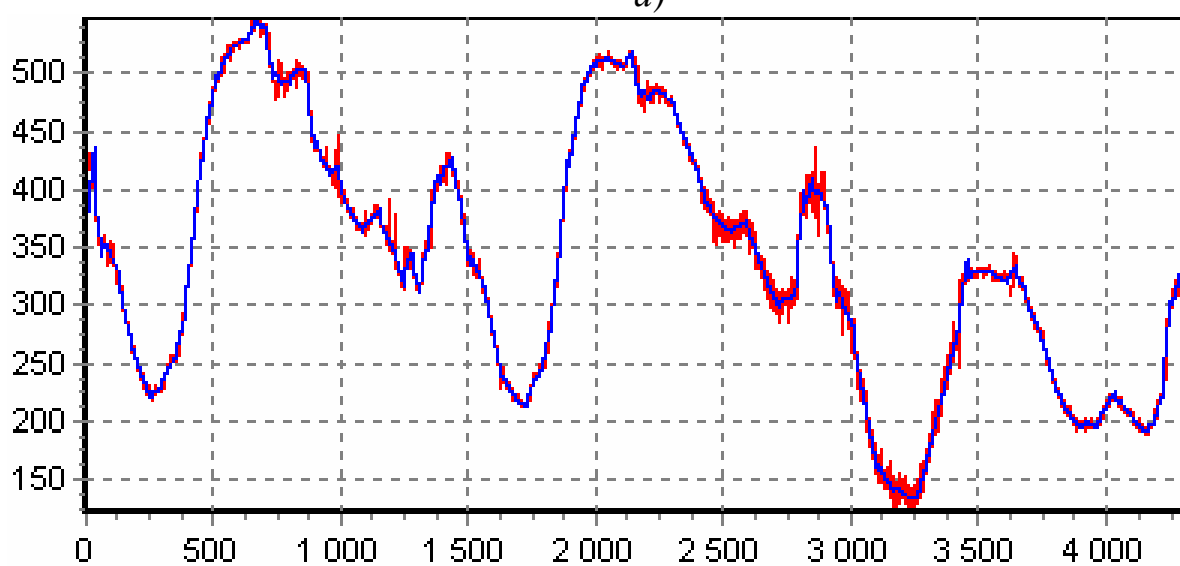
Проілюструємо сказане. На рис. 2.1,*а* показаний результат очищення деякого зашумленого сигналу за допомогою гладкого вейвлета (Добеші 7-го порядку, 5 рівнів розкладання), на рис. 2.1,*б* – те ж, але з використанням коротшого вейвлета (Добеші 2-го порядку, краще відстежуються списи сигналу), на рис. 2.1,*в* – результат "переукрупнення" сигналу (7 рівнів розкладання, згладжуються локальні особливості сигналу).

При виборі порогу шуму використовують, як правило, критерії, що мінімізують квадратичну функцію втрат для вибраної моделі шуму. У лабораторній роботі використовуються наступні порогові значення для обмеження коефіцієнтів розкладання:

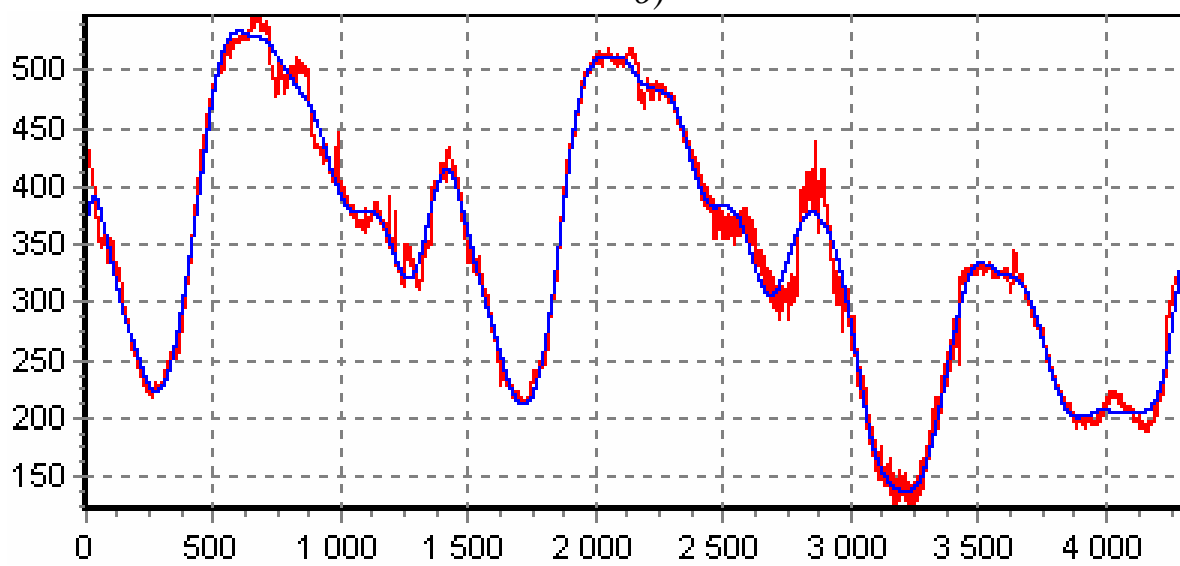
- адаптивний поріг на основі алгоритму Штейна незміщеної оцінки;
- евристичний варіант порогу по методу Штейна;
- універсальний поріг ;
- поріг за мінімаксною оцінкою.



a)



б)



в)

Рисунок 2.1 – Очищення сигналу від шуму за допомогою вейвлета Добеши

При гнучкому (чи м'якому) порозі задане порогове значення віднімається зі значень коефіцієнтів (по модулю), при жорсткому (чи твердому) порозі значення коефіцієнтів, менші порогових значень (по модулю), обнуляються. Жорсткий поріг зазвичай застосовується при компресії (стисненні) сигналів. Покажемо це на прикладі прямої лінії, обмеженої порогом 0,4, яка показана на рис. 2.2,а. На рис. 2.2,б ілюструється жорсткий пороговий метод, а на рис. 2.2,в – гнучкий.

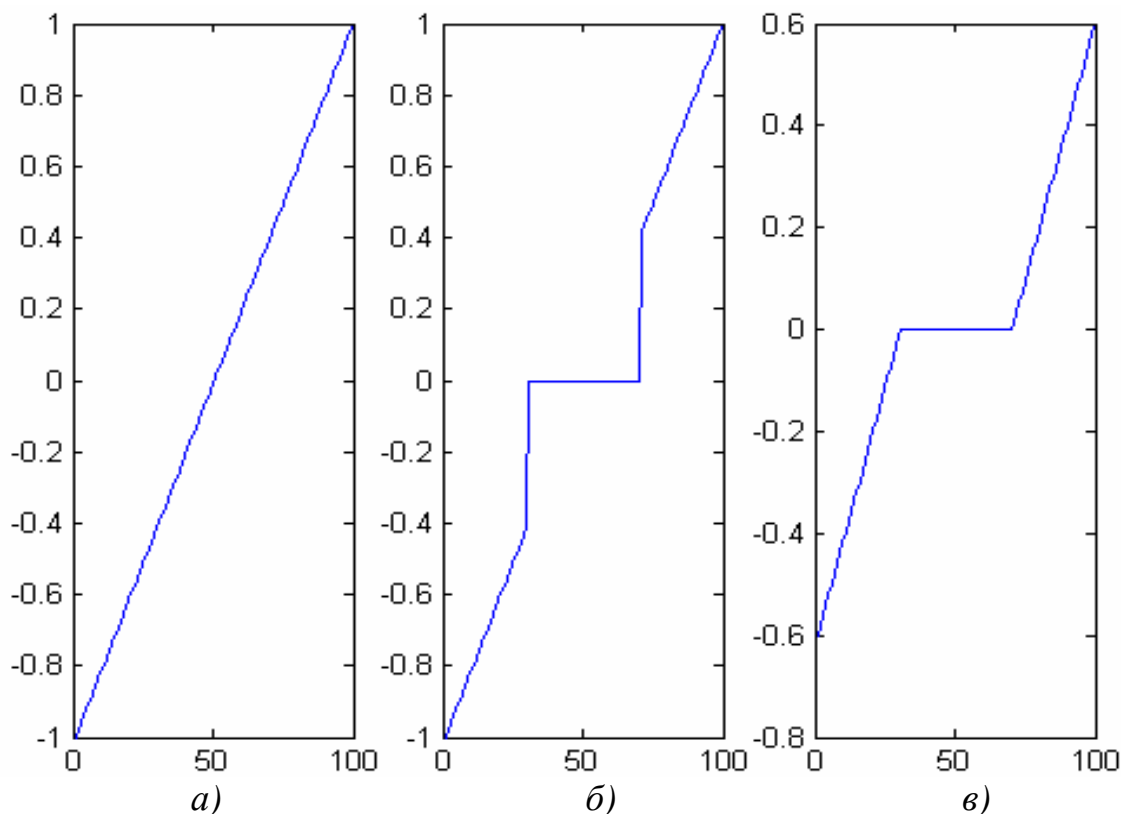


Рисунок 2.2 – Ілюстрація порогового метода

3. Ключові питання

1. В чому полягає метод очищення сигналу від шумів?
2. Яким чином метод пригнічення шумів реалізується за допомогою вейвлет-перетворень?
3. Назвіть необхідні операції для видалення шумів за допомогою вейвлетів.
4. Як впливає вибір вейвлета і глибини розкладання на міру очищення?
5. Які порогові методи використовуються в лабораторній роботі?
6. Для чого використовується м'який і жорсткий види порогів?

4. Домашнє завдання

За розділом 2 даного посібника вивчіть методи очищення від шумів і підготуйте відповіді на ключові питання.

5. Лабораторне завдання

1. Запустити з робочого столу програму MATLAB. У командному рядку написати команду *guide*. У вікні, що з'явилося, перейти на вкладку Open Existing GUI і вибрати файл Wavelet_Denoising. Натиснути кнопку  на панелі інструментів.

2. Для початку роботи натиснути кнопку "Пуск", вибрати кількість рівнів розкладання і порядок вейвлета Добеши зі списку. Змінюючи дані, виписати найкращі показники середньоквадратичного відхилення для кожної кількості рівнів розкладання.

3. Отримані на екрані дані звести в таблицю виду табл. 5.1:

Таблиця 5.1 – Підсумки виконання лабораторної роботи

Кількість рівнів розкладання	Порядок вейвлета Добеши	СКВ	Пороговий метод			
			адаптивний	еврістичний	універсальний	мінімаксний
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						

4. На підставі табл. 5.1 порівняти методи визначення порогового значення між собою, визначити найкращий порядок вейвлета, глибину розкладання і пороговий метод для представленого в лабораторній роботі сигналу.

6. Зміст протоколу

1. Тема й мета лабораторної роботи.
2. Відповіді на ключові питання.
3. Підсумкова таблиця 5.1
4. Висновки по виконаній роботі.

7. Список літератури

1. Рудий Є.М. Технології передачі дискретних повідомлень. Т.1: Стиснення сигналів: підручник. – Одеса: ОНАЗ ім. О.С. Попова, 2011. – 170 с.

Лабораторная работа 2.4

ОЧИСТКА СИГНАЛОВ ОТ ШУМОВ С ПОМОЩЬЮ ВЕЙВЛЕТОВ

1. Цель работы

Изучение принципов очистки сигналов от шумов с помощью вейвлет-преобразований, оценка качества восстановленного сигнала и определение наилучшего порядка вейвлета Добеши, порогового метода, а также количества уровней разложения.

2. Ключевые положения

Типовой метод подавления шумов – удаление высокочастотных составляющих из спектра сигнала. Применительно к вейвлетным разложениям это может быть реализовано непосредственно удалением детализирующих коэффициентов высокочастотных уровней. Но вейвлеты имеют в этом отношении более широкие возможности. Шумовые компоненты, и особенно большие случайные выбросы значений сигналов, можно также рассматривать в виде множеств локальных особенностей сигналов. Задавая некоторый порог для их уровня и срезая по нему детализирующие коэффициенты, можно не только уменьшать уровень шумов, но и устанавливать пороговые ограничения на нескольких уровнях разложения с учетом конкретных характеристик шумов и сигналов для различных типов вейвлетов. Это позволяет создавать адаптивные системы очистки сигналов от шумов в зависимости от их особенностей. Эти технологии и будут рассматриваться в данной теме.

Операция сжатия сигналов с удалением малозначимых значений вейвлет-коэффициентов также выполняется на основе определенных пороговых ограничений их значений, и во многом практически тождественна операциям удаления шумов.

При очистке сигналов от шумов и сжатии используется быстрое вейвлет-преобразование (БВП). При этом всегда следует учитывать, что если полный размер сигнала составляет M -отсчетов, а максимальный уровень разложения равен N , то для обеспечения нормальной работы БВП отношение $M/2^N$ должно быть целым числом, что обеспечивает целое число коэффициентов на последнем уровне разложения. Если это условие не выполняется, рекомендуется дополнять массив отсчетов нулевыми или любыми другими значениями.

Модель зашумленного сигнала обычно принимается аддитивной: $s(n) = f(n) + k \cdot e(n)$ с равномерным шагом по аргументу n , где $f(n)$ – полезная информационная составляющая, $e(n)$ – шумовой сигнал, например, белый шум определенного уровня со средним нулевым значением. Процедура удаления

шума выполняется с использованием ортогональных вейвлетов и включает в себя следующие операции:

– **Вейвлет-разложение сигнала** $s(n)$ до уровня N . Значение уровня N определяется частотным спектром информационной части $f(n)$ сигнала, которую желательно оставлять в рядах аппроксимационных коэффициентов. Тип и порядок вейвлета может существенно влиять на качество очистки сигнала от шума в зависимости как от формы сигналов $f(n)$, так и от корреляционных характеристик шумов.

– **Задание типа и пороговых уровней** очистки по известным априорным данным о характере шумов или по определенным критериям шумов во входном сигнале. Пороговые уровни очистки могут быть гибкими (в зависимости от номера уровня разложения) или жесткими (глобальными).

– **Модификация коэффициентов детализации** вейвлет-разложения в соответствии с установленными условиями очистки.

– **Восстановление сигнала** на основе коэффициентов аппроксимации и модифицированных детализационных коэффициентов.

Выбор используемого вейвлета и глубины разложения, в общем случае, зависит от свойств конкретного сигнала. Можно дать лишь несколько рекомендаций:

Более гладкие вейвлеты создают более гладкую аппроксимацию сигнала, и наоборот – «короткие» вейвлеты лучше отслеживают пики аппроксимируемой функции.

Глубина разложения N влияет на масштаб отсеиваемых деталей. Другими словами, при увеличении глубины разложения модель вычитает шум все большего уровня, пока не наступит «переукрупнение» масштаба деталей и преобразование начнет искажать форму исходного сигнала. Интересно, что при дальнейшем увеличении глубины разложения преобразование начинает формировать сглаженную версию исходного сигнала, т.е. отфильтровывается не только шум, но и некоторые локальные особенности (выбросы) исходного сигнала.

Проиллюстрируем сказанное. На рис. 2.1,а показан результат очистки некоторого зашумленного сигнала с помощью гладкого вейвлета (Добеши 7-го порядка, 5 уровней разложения), на рис. 2.1,б – то же, но с использованием более короткого вейвлета (Добеши 2-го порядка, лучше отслеживаются пики сигнала), на рис. 2.1,в – результат «переукрупнения» сигнала (7 уровней разложения, сглаживаются локальные особенности сигнала).

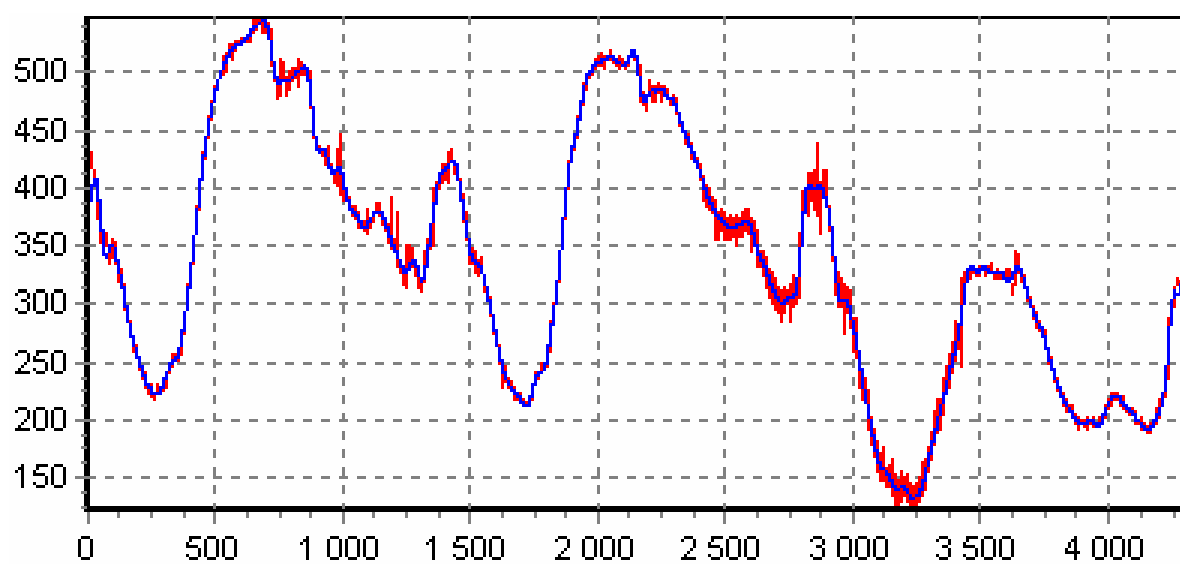
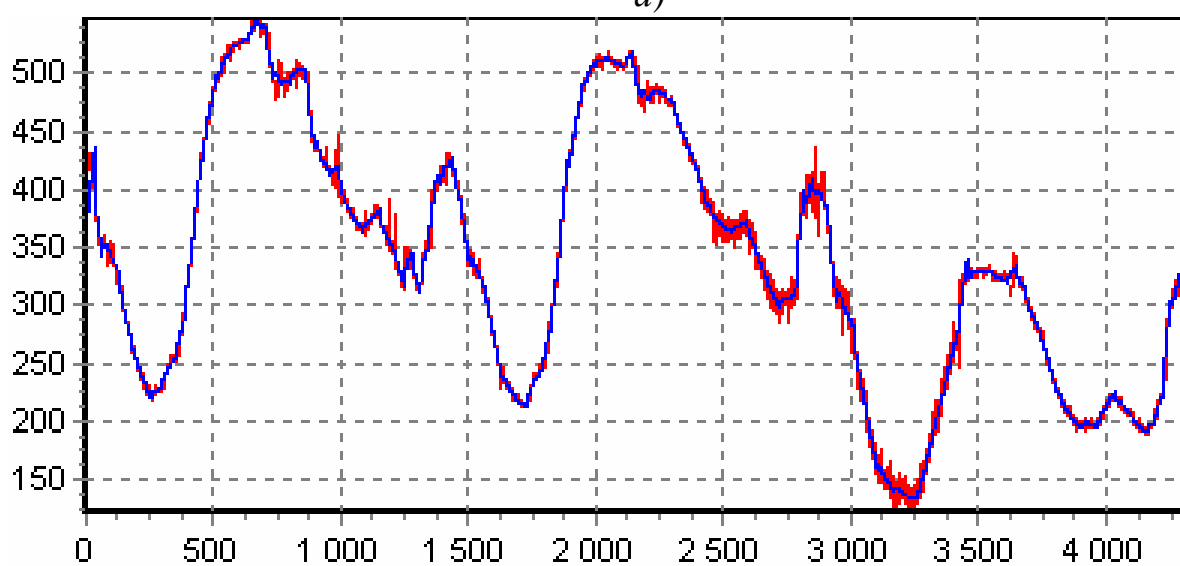
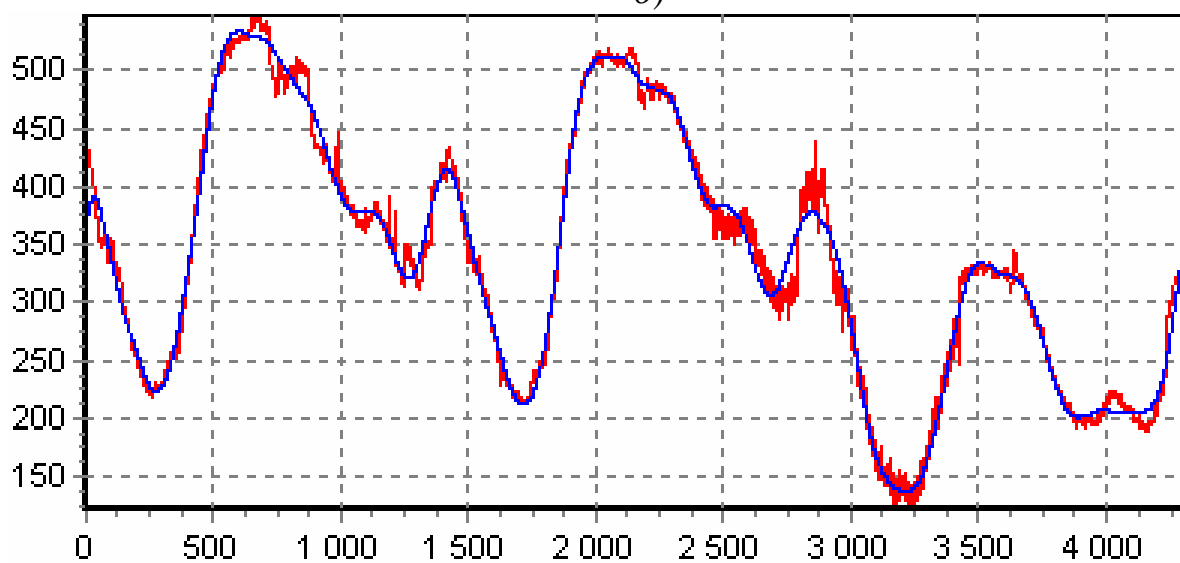
*a)**б)**в)*

Рисунок 2.1 – Очистка сигнала от шума с помощью вейвлета Добеши

При выборе порога шума используют, как правило, критерии, минимизирующие квадратичную функцию потерь для выбранной модели шума. В лабораторной работе используются следующие пороговые значения для ограничения коэффициентов разложения:

- адаптивный порог на основе алгоритма Штейна несмещенной оценки;
- эвристический вариант порога по методу Штейна;
- универсальный порог $\sqrt{2 \cdot \log(\text{length}(X))}$;
- порог по минимаксной оценке.

При гибком (или мягком) пороге заданное пороговое значение вычитается из значений коэффициентов (по модулю), при жестком (или твердом) пороге значения коэффициентов, меньшие пороговых значений (по модулю), обнуляются. Жесткий порог обычно применяется при компрессии (сжатии) сигналов. Покажем это на примере прямой линии, ограниченной порогом 0,4, которая показана на рис. 2.2,а. На рис. 2.2,б иллюстрируется жесткий пороговый метод, а на рис. 2.2,в – гибкий.

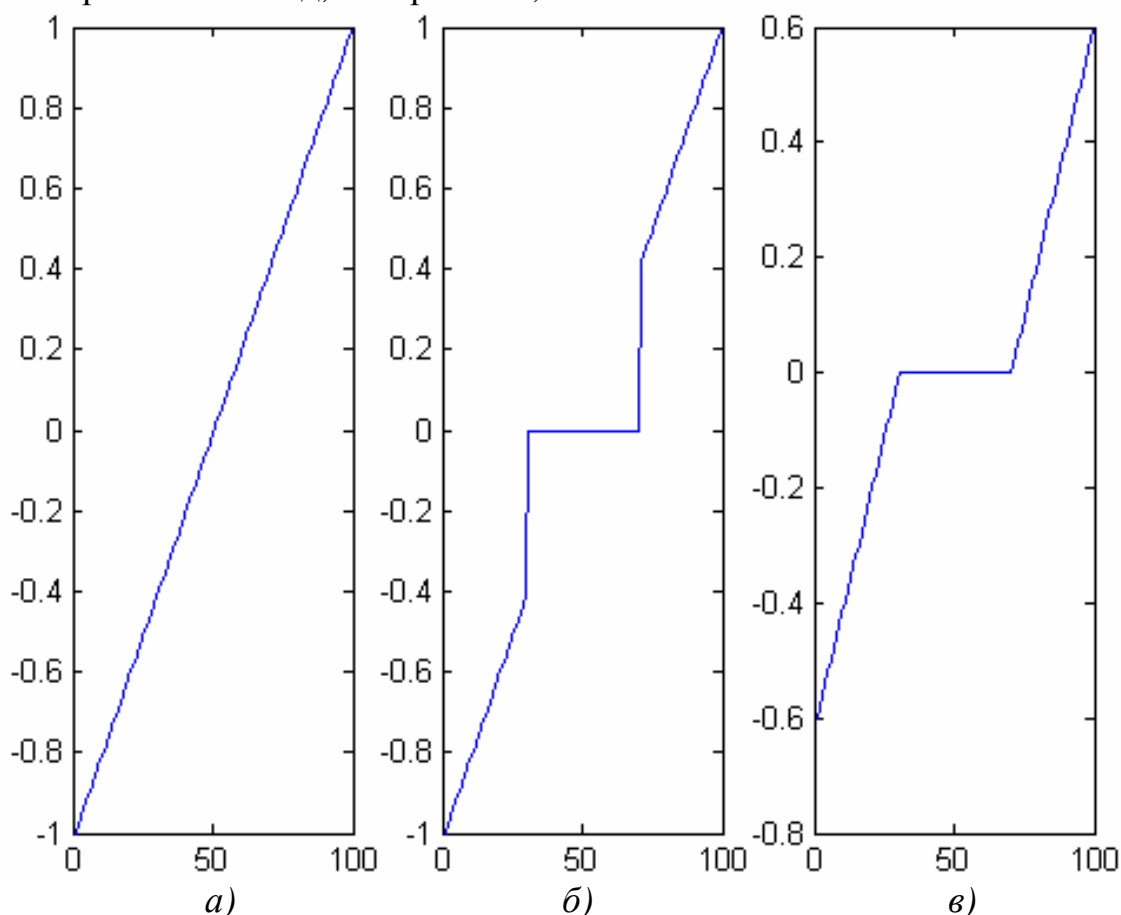


Рисунок 2.2 – Иллюстрация порогового метода

3. Ключевые вопросы

1. В чем заключается метод очистки сигнала от шумов?
2. Каким образом метод подавления шумов реализуется с помощью вейвлет-преобразований?

3. Назовите необходимые операции для удаления шумов с помощью вейвлетов.

4. Как влияет выбор вейвлета и глубины разложения на степень очистки?

5. Какие пороговые методы используются в лабораторной работе?

6. Для чего используется мягкий и жесткий виды порогов?

4. Домашнее задание

По разделу 2 данного пособия изучите методы очистки сигналов от шумов і підготуйте ответы на ключевые вопросы.

5. Лабораторное задание

1. Запустить с рабочего стола программу MATLAB. В командной строке написать команду *guide*. В появившемся окне перейти на вкладку Open Existing GUI и выбрать файл Wavelet_Denoising. Нажать кнопку  на панели инструментов.

2. Для начала работы нажать кнопку «Пуск», выбрать количество уровней разложения и порядок вейвлета Добеши из списка. Изменяя данные, выписать наилучшие показатели среднеквадратического отклонения для каждого количества уровней разложения.

3. Полученные на экране данные свести в таблицу вида табл. 5.1:

Таблица 5.1 – Итоги выполнения лабораторной работы

Количество уровней разложения	Порядок вейвлета Добеши	СКО	Пороговый метод			
			адаптивный	эвристический	универсальный	минимаксный
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						

4. На основании табл. 5.1 сравнить методы определения порогового значения между собой, определить наилучший порядок вейвлета, глубину разложения и пороговый метод для представленного в лабораторной работе сигнала.

7. Список литературы

1. Рудий Є.М. Технології передачі дискретних повідомлень. Т.1: Стиснення сигналів: підручник. – Одеса: ОНАЗ ім. О.С. Попова, 2011. – 170 с.