­­МІНІСТЕРСТВО ОСВІТИ ТА НАУКИ УКРАЇНИ

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ “ЛЬВІВСЬКА ПОЛІТЕХНІКА”

­­­

Лабораторна робота №4

з курсу «**Проблемно-орієнтовані методи та засоби**

**інформаційних тенологій**»

для студентів базового напрямку 6.08.04 "Комп’ютерні науки"

(заочна форма навчання)

Варіант 10

Виконав студент гр. КНз-3

Чалий Михайло

­­Львів 2015

## МЕТА РОБОТИ

Мета роботи – вивчити формати цифрового представлення звукових сигналів і методів їх обробки для аналізу та синтезу звукових повідомлень.

## ТЕОРЕТИЧНІ ВІДОМОСТІ

При однобайтовому представленні максимальна кількість квантів буде 256 (від 0 до 255, причому значення 128 відповідає нульовому рівню сигналу). Похибка при однобайтовому представленні не перевищує 1/256\*100 %≈0.4 %. При двобайтовому представленні кількість квантів рівна 65536, а похибка не первищує 1/65536\*100 %. Кількість квантів визначає не тільки точність (похибку) представлення, але й динамічний діапазон зміни сигналу, який визначається співвідношенням максимального і мінімального його рівнів. Динамічний діапазон визначається в децибелах по формулі D=20\* lg\*(Umax/Umin) dB. Наприклад, при Umax=1 Вольт і Umin=10 мікроВольт, динамічний діапазон буде рівний 100 dB.

Частота дискретизації і кількість рівнів квантування визначають затрати пам’яті, необхідної для збереження відліків дискретизованого аналогового сигналу. Наприклад, для для двохканального стререофонічного запису музики при двобайтовому представленні відліків і частоті дискретизації 44 кГц затрати пам’яті становитимуть M=44 кГц\*2 байти\*2 канали=176 Кбайт/сек= 633.6 Мбайт/год. При цьому, для неперервного запису чи відтворення потік байтів чи бітів оцифрованого сигналу (бітрейт) має бути не меншим ніж 176 Кбайт/сек (без використання стиснення).

Системи комп’ютерної генерації (синтезу) голосових повідомлень базуються на двох підходах:

* побудова фізичної моделі мовної системи людини (біонічний підхід), який ще називається артикулярним синтезом;
* моделювання та формування акустичного сигналу методами компілятивного та формантного синтезу.

Класифікація систем розпізнавання голосових повідомлень по призначенню:

* командні системи;
* системи диктування тексту.

Класифікація систем розпізнавання мови по споживчих якостях:

* диктороорієнтовані;
* дикторонезалежні;
* розпізнаючі окремі слова;
* розпізнаючі розмовну (неперервну) мову.

Основними форматами для запису і відтворення аудіосигналів є WAV та MP3. Формат WAV є метаформатом для даних будь-якого типу. Має стандатний заголовок і описи областей даних (однієї чи кількох). Метод кодування вказується в заголовку і розпізнається операційною системою для підключення та запуску кодека.

Структура WAV-файлу містить п’ять розділів: заголовок файлу, заголовок опису файлу, заголовок фактичних даних, заголовок даних файлу, дані відліків. Опис полів кожного розділу представлений у таблиці.

|  |  |  |
| --- | --- | --- |
| Заголовок файлу | | |
| Назва поля | Ім’я поля | Розмір |
| Тип файлу (для WAV-файлу – ‘RIFF’) | FileType | DWORD |
| Розмір файлу в байтах | FileSize | DWORD |
| Заголовок опису файлу | | |
| Назва поля | Ім’я поля | Розмір |
| Назва заголовку опису файлу (‘WAVEFMT ‘) | Descriptor | 8 BYTE |
| Розмір заголовку в байтах | DescSize | DWORD |
| Стиснення (1 – відсутнє) | Compression | WORD |
| Кількість каналів (1 – MONO, 2 – STEREO) | NonberChannel | WORD |
| Частота дискретизації | DescrFreq | DWORD |
| Бітрейт | Bitrate | DWORD |
| Представлення (1 – 8 біт) |  | WORD |
| Розрядність (8, 16) |  | WORD |
| Резерв | Reserved | WORD |
| Заголовок фактичних даних | | |
| Назва поля | Ім’я поля | Розмір |
| Назва заголовку фактичних даних (‘FACT’) | Descriptor | DWORD |
| Розмір заголовку в байтах | DescSize | DWORD |
| Кількість байтів даних | DataLenth | DWORD |
| Заголовок даних | | |
| Назва поля | Ім’я поля | Розмір |
| Назва заголовку даних (‘DATA’) | Descriptor | DWORD |
| Кількість байтів даних | DataLenth | DWORD |
| Дані відліків | | |
| Назва поля | Ім’я поля | Розмір |
| Відліки |  | BYTE  WORD  DWORD |

Формат MP3 (скорочення від MPEG Layer 3) – один із основних цифрових форматів для збереження стиснутих аудіоданих. Є частиною стандартів стиснення аудіо та відеоданих MPEG1 та MPEG2. Використовується для передачі аудіоданих в реальному часі по мережевих каналах і для кодування звукових CD-дисків. Забезпечує бітрейт 320 Кбіт/сек, який є максимальний для кодування звуку з характеристиками CD Audio (44 кГц, 16 біт Stereo). Вважається, що бітрейт 256 Кбіт/сек є достатнім по якості для більшості користувачів.

# **Завдання**

**Варіант 10**

* + - 1. За допомогою програми “Звукозапис” запишіть з мікрофона ряд звукових (мовних) повідомлень у вигляді WAV-файлу. При цьому за допомогою конвертора перетворіть отриманий сигнал в досліджуваний монофонічний сигнал з частотою дискретизації 8000 відліків за секунду при однобайтовому представленні (256 квантів).
      2. Порівняйте звучання мовних сигналів при частоті дискретизаціїї 44 кГц та 8 Кгц і дво- та однобайтовому представленні. Зробіть висновки щодо якості звучання.
      3. За допомогою програми графічного представлення сигналів виведіть на екран графіки простих повідомлень “Так”, “Ні”. Опишить властивості вказаних сигналів (форма, періодичність, зміна амплітуди і т.п.).
      4. На основі записаних мовних повідомлень синтезуйте нове повідомлення (“Так-“Так”- “Ні”) методом компілятивного синтезу і запишіть його у вигляді WAV-файлу. Відтворіть отримане повідомлення за допомогою програми “Звукозапис” і зробіть висновки про його особливості.

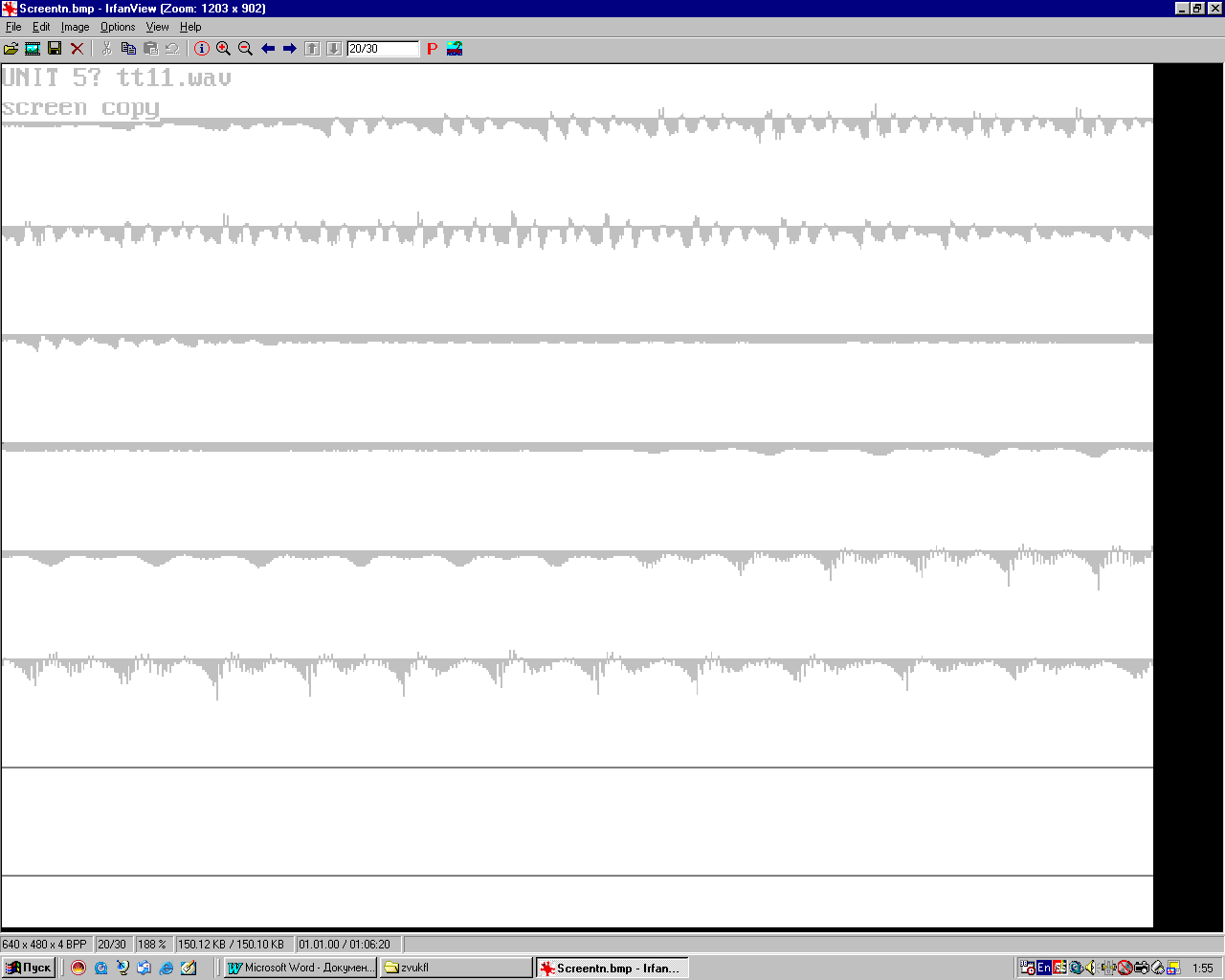
1. **Хід роботи**
2. ****Генерую за допомогою програми сигнал відповідного звукового файлу. На Рис.1 показаний графік сигналу слів ”Так-Ні”.

Рис.1 Сигнал слів ”Так-Ні”

1. Наступним кроком є синтез нового повідомлення , яке складається із двох слів: Ні-Ні. Графік синтезованого повідомлення показаний на Рис.2.

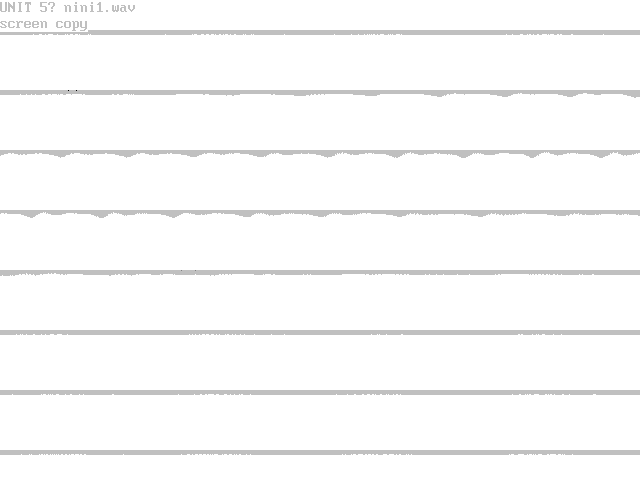
****

Рис.2 Графік сигналу результуючого повідомлення

## Висновок

Під час виконання даної лабораторної роботи я навчився синтезувати нові повідомлення із існуючих слів, графічно представляти їх, вивчив формати цифрового представлення звукових сигналів і методи їх обробки для аналізу та синтезу звукових повідомлень.