

# МОДЕЛЮВАННЯ ПЕРЕКЛАДУ УСНОМОВНИХ СЛІВ

Пилипенко Валерій

Міжнародний науково-навчальний центр інформаційних технологій та систем

40, просп. Академіка Глушкова, Київ 252022

Електронна пошта: pilipenk@uasoiro.freenet.kiev.ua

## Абстракт

Описується програмний комплекс, в який входять програми розпізнавання усної мови, перекладу її на іншу та озвучування відповіді. Комплекс дозволяє користувачу вимовляти в мікрофон слово рідною мовою, побачити переклад слова в текстовому редакторі і почути його. Також є сервісні функції, що дозволяють ввести нову мову для перекладу, словник і настроїтися на голос диктора. Окрім цього, є можливість контролювати процес введення звукового сигналу.

## Реалізація програми перекладача

Програма, що моделює переклад усномовних слів, реалізована у вигляді текстового редактора, який дозволяє користувачеві створювати текст не лише шляхом введення символів за допомогою клавіатури, але і промовляючи слова в мікрофон. Використовується стандартний текстовий редактор, вбудований в операційну систему Windows 95.

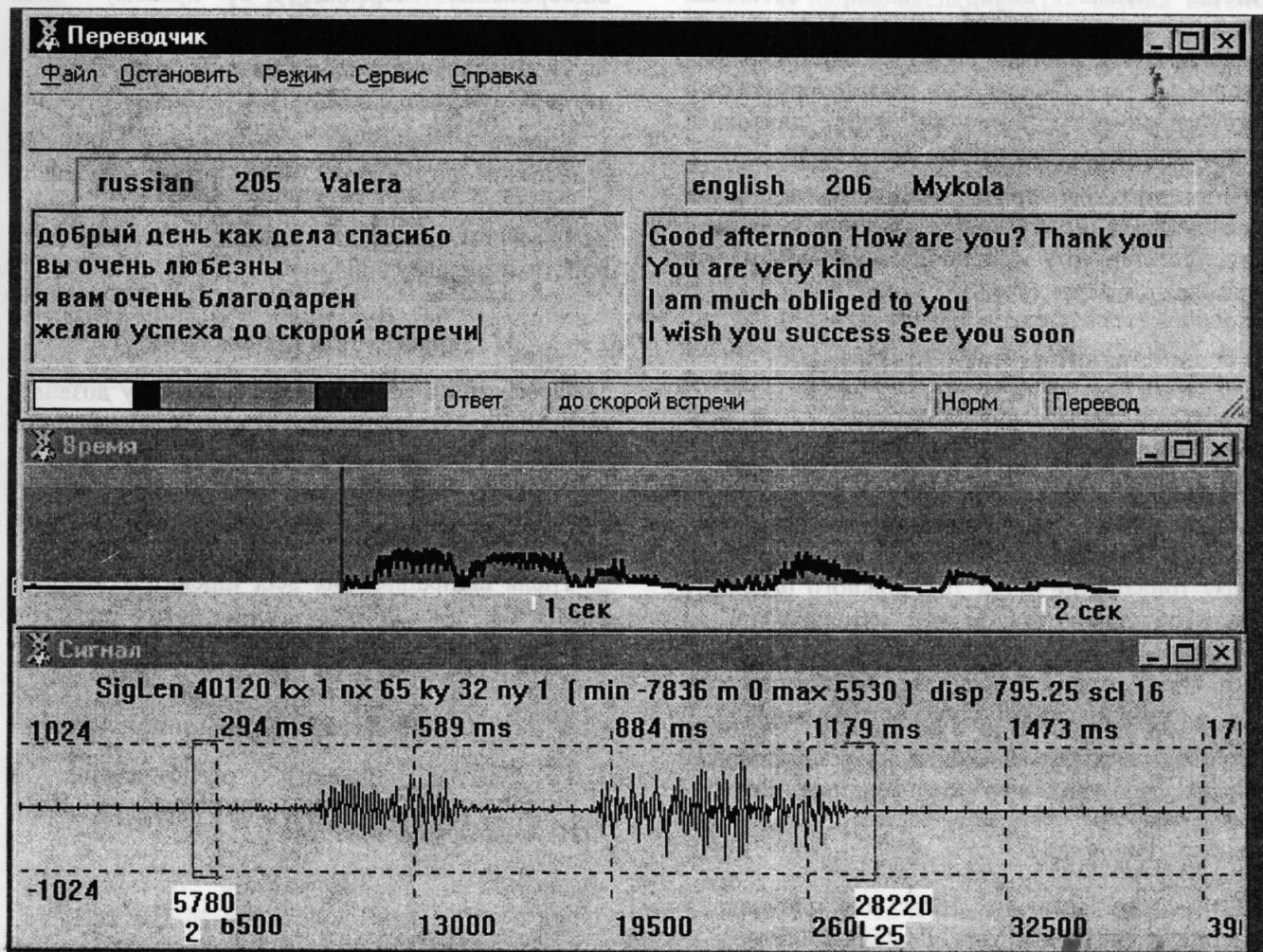


Рис. 1. Загальний вигляд програми перекладача усномовних слів. Панель текстового редактора поділена на дві частини – вхідний текст і переклад. Для контролю гучності вимовляння виведений індикатор гучності. При необхідності можна контролювати рівень гучності (середнє вікно). Також можна розглянути введений сигнал за допомогою нижнього вікна.

Вікно редактора поділене на дві частини — в лівій відображається текст фрази тією мовою, якою його вимовив користувач, в правій частині дається переклад на потрібну мову (рис. 1). В результаті, завжди можна водночас бачити текст на вхідній мові та перекладений текст.

Результат перекладу озвучується за допомогою заздалегідь записаних слів або синтезатора усної мови.

### **Настроювання на мову, словник і голос користувача**

Для введення в систему нової мови і нового словника достатньо набрати в довільному текстовому редакторі словник потрібною мовою і зберегти його в текстовому режимі під іменем, яке буде відображати мову і словник згідно погоджень, прийнятих у системі.

Програма настроюється на мови розпізнавання і перекладу, на робочий словник і на голос диктора. Після вибору мови, словника, користувача і кількості слів пропонується вимовити кожне вибране слово для настроювання на голос диктора. Образи слів запам'ятовуються в індивідуальному усномовному файлі диктора і використовуються при розпізнаванні мови.

Після того, як проведено настроювання на голос диктора, є можливість використати базу даних його голосу для відповідей в процесі розпізнавання-перекладу.

### **Модуль розпізнавання**

Програма реалізована мовою C/C++ в операційній системі Windows 95.

Вхідний сигнал, введений за допомогою Sound Blaster, перетворюється в цифровий вигляд (16 біт, 22 кГц) і рівномірно розбивається на інтервали аналізу тривалістю 15 мс. Для кожного інтервалу аналізу обчислюються вектор автокореляції та міри схожості з еталонними елементами з кодової книги. Після цього обчислюється найкраще слово згідно алгоритму динамічного програмування [1].

Проводилися експерименти на різноманітних вибірках, для яких програма показала наступні параметри:

1. Обсяг словника — до 4 тисяч слів.
2. Час розпізнавання 0.25 с для словника в одну тисячу слів для IBM PC Pentium-200 Pro.
3. Середня надійність розпізнавання не менше 90% для словника в одну тисячу слів.

### **Архітектура комплексу**

Моделювання процесу розпізнавання, перекладу з однієї мови на іншу та видачі відповіді проводиться за допомогою комплексу, що складається з таких програмних блоків:

1. Введення-виведення мовного сигналу, визначення моментів початку та кінця мовного сигналу, забезпечення оброблення мовного сигналу в реальному часі на фоні введення-виведення мовного сигналу.

2. Формування індивідуального усномовного файлу диктора за навчальними вибірками.

3. Формування бази даних для відповідей.

4. Автоматичне розпізнавання усномовного сигналу.

5. Відображення відповіді розпізнавання в лівому вікні текстового редактора та часового сигналу в окремих вікнах.

6. Переклад з вхідної мови на потрібну і відображення перекладу в правому вікні текстового редактора.

7. Відтворення або синтез перекладу у вигляді звукового сигналу.

Блочна структура комплексу дозволяє незалежно розробляти його різні частини і об'єднувати їх у відповідності з фіксованим інтерфейсом взаємодії модулів для перевірки роботи комплексу в цілому.

### **Прикінцеві положення**

Програмний комплекс для перекладу дозволяє відпрацьовувати різноманітні модулі у взаємодії з метою з'ясування вузьких місць і відпрацьовування нових алгоритмів розпізнавання мови, перекладу з однієї мови на іншу, а також синтезу усної мови.

Після відпрацювання всіх підсистем можлива реалізація комплексу для промислового використання.

### **Література**

1. Т.К. Винцюк. Анализ, распознавание и интерпретация речевых сигналов. — Киев: Наукова думка, 1987, 264 с.