

УДК 681.51.011

Н.В. Палагнюк, Д.М. Михалик канд. техн. наук

Тернопільський національний технічний університет імені Івана Пулюя, Україна

АЛГОРИТМУ СИНТЕЗУ МОВЛЕННЯ І ЙОГО РЕАЛІЗАЦІЯ У ВИГЛЯДІ МОБІЛЬНОГО ЗАСТОСУНКУ ПІД ПЛАТФОРМУ IOS

N.V. Palahniuk, D.M. Mykhalyk Ph.D.

ALGORITHM OF SPEECH SYNTHESIS AND ITS IMPLEMENTATION AS A MOBILE APPLICATION ON IOS

На даний момент технологія синтезу мовлення не є дивиною для, наприклад, англійської мови, проте і синтез українського мовлення є достатньо розвинутою технологією, щоб синтезований голос звучав природно, а текст читався з інтонацією та потрібним темпом.

Озвучення текстів і сьогодні залишається важливою й актуальною задачею усномовної інформатики. Попри значні успіхи в синтезуванні мовних сигналів озвучення текстів досі не є розповсюдженою технологією. Це зумовлено зокрема тим, що підвищились вимоги до якості та натуральності звучання синтезованої мови. Разом з тим для впровадження синтезу в портативних пристроях повинна задовольнятися вимога на обмеження швидкодії та обсяги пам'яті.

Метою даної роботи є аналіз існуючих технологій синтезу мовлення та розробка і реалізація алгоритму синтезу мовлення на їх основі для портативних пристроїв на платформі iOS. На основі проведених досліджень буде розроблено мобільний застосунок під платформу iOS версій 7.0 і вище, який буде синтезувати українське мовлення.

Синтез мовлення – це процес перетворення друкарського тексту у мовний сигнал, в широкому сенсі – відновлення форми мовного сигналу за його параметрами [1].

Синтезатори мовлення ґрунтуються або на використанні моделей мовного тракту людини, або на конкатенації необхідних сегментів мовлення, що вже зберігаються в мовленнєвих базах даних (МДБ) [2].

На сьогодні, найбільш поширеним методом синтезу мовлення за текстом є конкатенативний метод, який полягає в поєднанні відрізків мовленнєвих сигналів, що зберігаються у МБД.

Для реалізації даної мети використано бібліотеку MVSpeechSynthesizer, яка вже містить алгоритми синтезу для багатьох мов, але, на жаль, й досі не містить української. Для виправлення цієї ситуації у МБД даної бібліотеки внесено елементи відрізків мовленнєвих сигналів, що відповідають таким одиницям, як фонемі (звуки), дифоні (відрізки мовленнєвого сигналу від середини однієї фонемі до середини наступної) та напівфоні (половини дифонів). Окрім того досліджено засоби генерування тональних (інтонаційних) та темпоральних (ритмічних) характеристик слів, речень та текстів і використано їх для генерування просодичних характеристик, що дозволяє синтезованому голосу звучати більш природно.

Література

1. Розпізнавання та синтез мовлення в Україні [Електронний ресурс] – Режим доступу: <http://speech.com.ua/>

2. Людовик Т.В., Сажок М.М., Селюх Р.А. Розроблення мовленнєвих баз даних для синтезу мовлення за текстом. // Автоматизовані системи управління і нові інформаційні технології. – Міжнародний науково-навчальний центр інформаційних технологій та систем НАН України та МОН України – К., 2005. - Випуск 3.