

УДК 519.711.2

Щур І. – ст. гр. ПМм – 51

Тернопільський державний технічний університет імені Івана Пулюя

МАТЕМАТИЧНА МОДЕЛЬ НОСОВИХ ПРИГОЛОСНИХ МОВНИХ СИГНАЛІВ

Науковий керівник: к.т.н., Шадріна Г.М.

Основна проблема, яка перешкоджає впровадженню мовних технологій в численні сфери діяльності людини, це недостатня стійкість процесу розпізнавання і розуміння мови.

Про потрібність мовного інтерфейсу засвідчує постійно зростаюче число комерційних розробок систем, які використовують мовний інтерфейс. Наприклад NaturallySpeaking, фірми Dragon Sytem, дозволяє редагувати і формувати текст за допомогою власного текстового процесора без використання клавіатури і мишки. Компанія IBM розробила аналогічну програму, яка дозволяє здійснювати мовний ввід і форматування тексту в текстовому процесорі MS Word. На практиці ці програми показують недостатньо високі результати.

Основним недоліком існуючих систем є їх орієнтація на використання статистичних даних без врахування особливостей мовотворення, які зв'язані з індивідуальними особливостями диктора. Точність результату роботи таких систем залежить від повноти і адекватності використаних баз даних голосів, введення яких є дуже трудомістким і дорогим процесом. Ще одна проблема – адаптація відомих систем для мов, які відрізняються від базових (для яких розроблялася система).

Проблема розпізнавання мови полягає в визначеності послідовності слів, які містяться в голосовому повідомленні, на основі потоку акустичної інформації, тобто автоматичному розумінню мови, виявленню змісту сказаного диктором.

Основним із методів розпізнавання мови є метод еталонного розпізнавання. Проте, як для еталону, так і для спонтанної мови потрібно чітко вимовляти звуки та слова. Чітка вимова потребує спеціального навчання диктора.

Існують також інші методи розпізнавання мови. Зокрема, проблема розпізнавання звуків може вирішуватися, шляхом використання математичних моделей замість шаблонів в більшості відомих програм для перетворення мови в текст. Математичні моделі повинні відображати характерні особливості конкретного звуку одного диктора і бути достатньо точними для розпізнавання. Вплив попередніх та наступних звуків в слові може змінювати деякі параметри носового звуку, тому він повинен бути передбачений в моделі. Крім того, математична модель повинна містити особливості конкретного звуку для вирізнення його з групи подібних.

Для побудови математичної моделі носових приголосних звуків „м” та „н” можна використати емпіричні дані про процес їх утворення, а також їх спектральні характеристики. Звуки „м” та „н” є вокалізованими, тобто в їх утворенні бере участь голос та шуми. Тому, можна розраховувати на присутність в спектрі високочастотних (більше 1 кГц) складових одночасно з низькочастотними шумами.

Для побудови математичної моделі доцільно використати опис звуку у вигляді періодично корельованого випадкового процесу (ПКВП), який дає змогу у своїй структурі поєднати властивості повторності з випадковістю.