

РОЗРОБЛЕННЯ РЕАБІЛІТАЦІЙНОЇ МОДЕЛІ ЗВУКОВИМОВИ ДЛЯ РОЗПІЗНАВАННЯ ЗВУКІВ „Р” ТА „Л” В СИСТЕМІ „ЛЮДИНА-КОМП’ЮТЕР”

Багато людей різного віку страждають від проблем, пов'язаних з неправильною вимовою звуків „Р” та „Л” (ротацізм та ламбдацизм). Щоб ліквідувати вищенаведені вади звуковимови необхідне систематичне тренування мовного апарату за допомогою комплексу фізичних вправ для зміцнення м'язів язика і щелепи, та набору слів і словосполучень, які багаторазово повторюються. Особливо важливою ця проблема є для людей з порушеннями слуху, які здатні сприймати інформацію лише на основі зорового аналізатора. В літературних джерелах не знайдено інформації про наявність комп'ютерних автоматизованих систем, які б дозволяли проводити тренування і оцінювати їх результативність. Тому, актуальною є задача розроблення реабілітаційної моделі звуковимови, придатної для тренування мовного апарату в системі „людина-комп'ютер”.

Етапи розроблення реабілітаційної моделі звуковимови для розпізнавання звуків „Р” та „Л” в системі „людина-комп'ютер” представлено так:

1. Мовний сигнал відбираємо від людини і передаємо у базу даних комп'ютера.
2. За допомогою кратномасштабного аналізу вейвлет-перетворення проводимо фонемну сегментацію мовного сигналу (слова) для виділення досліджуваного фрагменту (фонем /р/, /р'/, /л/, /л'/).
3. Виділяємо інформативні ознаки досліджуваних фонем (частоту основного тону, форманти).
4. Порівнюємо характеристики досліджуваних фонем з еталонними ознаками фонем /р/, /р'/, /л/, /л'/.
5. Виводимо результат розпізнавання.

Для здійснення фонемної сегментації мовний сигнал, оцифрований з частотою дискретизації 22 050 Гц, розбиваємо ковзаючим вікном, розміром 512 відліків на масиви з половинним перекриттям і формуємо з них єдиний масив. Отриманий масив даних розкладаємо на $j=6$ рівнів деталізації з використанням кратномасштабного вейвлет-перетворення в базисі Добеші. Для кожного j -го рівня будемо числову послідовність коефіцієнтів деталізації:

$$e_{ij} = 10 * \lg \sum_{k=0}^{n_j-1} d_{j,i+k}^2,$$

де i – номер ковзаючого вікна, $n_j = \frac{n}{2^j}$ - розмір ковзаючого вікна на j -му рівні, n - розмір вікна в початковому сигналі ($n=512$), $d_{j,i+k}$ - коефіцієнти деталізації.

Набір коефіцієнтів деталізації, дає змогу виявити міжфонемні переходи, оскільки на міжфонемних переходах сигнал зазнає змін і, відповідно, характеризується зростанням вейвлет-коефіцієнтів.

Тому, задача фонемної сегментації слова зводиться до знаходження міжфонемних меж, на основі яких виділяються фонемні для подальшого порівняння з еталонами.