

# Керування швидкістю подання голосової інформації

Зореслава Шпак

Кафедра автоматизованих систем управління, Національний університет  
«Львівська політехніка», УКРАЇНА, м. Львів, вул. С. Бандери, 12;  
e-mail: zshpak@ukr.net

*Time-scale transformation regularity of the voice phonemes sounding at different speech rates has been investigated. A new time-domain method based on the differentiated modification of speech units durations has been developed. This method reproduces regularities of oral speech rate-change transition processes.*

Ключові слова – голосова комунікація, мовний сигнал, часове масштабування, темпоральні перетворення.

## I. Вступ

Потреба в регулюванні швидкості відтворення голосової інформації є актуальною для багатьох сфер усномовної комунікації. Зокрема, це навчальні системи, радіо, стенографування, логопедія тощо. Високу натуральність звучання прискореної чи, навпаки, сповільненої мови можна отримати тільки за умови, що часові модифікації голосових сигналів будуть близькими до природних темпоральних процесів.

## II. Темпоральна поведінка мовних елементів

Дослідження часових трансформацій елементів мовного потоку, викликаних зміною темпу мовлення, показали, що звуки різної артикуляційно-акустичної природи неоднаково змінюють свою тривалість. Найбільших змін зазнають міжслівні паузи та голосні звуки, а найменш варіативними є глухі приголосні [1].

Встановлено, що звуки з однорідними механізмами формування характеризуються подібністю відносної зміни тривалості. На основі аналізу темпоральних закономірностей всі мовні елементи поділено на шість груп: наголошені голосні, ненаголошені голосні, тональні приголосні, шумні приголосні, паузи зімкнення та міжслівні паузи. Елементи кожної групи вирізняються спільністю темпоральної поведінки, незалежно від їхньої абсолютної тривалості.

Підтверджено, що зміна темпу мовлення не впливає істотно на спектральні характеристики звуків, зокрема, практично незмінними залишаються розподіли формантних частот і мелодичні контури значень основного тону.

Таким чином, у процесі часового масштабування мовних записів має здійснюватись диференційована зміна тривалостей звукових елементів за умови збереження природної структури їхніх сигналів.

### III. Метод диференційованих часових перетворень

Процес відтворення голосових записів у заданому темпі включає декілька етапів. Спочатку виокремлюються паузи, які за тривалістю та звуковим оточенням розділяються на міжслівні та внутрішні.

На другому етапі у межах звукових ділянок визначаються квазістаціонарні сегменти з однорідним спектральним складом. Здійснюється аналіз параметрів виділених сегментів і формування з них ділянок звуків однієї з темпоральних груп. Процес класифікації звуків достатньо складний, він становить окрему наукову задачу.

Третій етап призначений для диференційованого розрахунку тривалості кожного мовного елемента відповідно до заданого коефіцієнта зміни темпу  $k$  ( $k < 1$  – прискорення,  $k > 1$  – сповільнення):

$$t_{\text{вух}}^{(i)} = t_{\text{ex}}^{(i,j)} \cdot s_j(k), \quad i = \overline{1, N}, \quad j = \overline{1, M}, \quad (1)$$

де  $s_j$  – коефіцієнт, що описує нелінійну залежність зміни тривалості звуків  $j$ -ої темпоральної групи від загального коефіцієнта зміни темпу  $k$ .

Якщо задано фіксований час  $T_{\text{вух}}$ , за який має відтворюватися голосове повідомлення, то значення  $k$  розраховується з рівності:

$$\sum_{j=1}^M \rho_j s_j(k) = T_{\text{вух}} / T_{\text{ex}}, \quad (2)$$

де  $\rho_j$  – частка звуків  $j$ -ої темпоральної групи.

На наступному етапі змінюється тривалість кожного з мовних елементів – проводиться модифікація структури звукового сигналу. Цей етап визначальний для якості та чистоти звучання промасштабованого повідомлення, оскільки часове скорочення чи розтягування звуку не повинно впливати на його частотні характеристики. Перетворення тональних звуків реалізовано синхронно з основним тоном через вилучення окремих періодів або доповнення сигналу перехідними композитними періодами [2]. Модифікації сигналів шумних звуків включають амплітудно-фазові узгодження.

### Висновок

Розроблений метод диференційованого часового перетворення звукових ділянок дає змогу змінювати у широких межах швидкість подання мовної інформації зі збереження оригінальності її звучання.

### Література

1. З. Шпак, Ю. Рашкевич. Збереження природних темпоральних закономірностей у процесах сповільненого відтворення мовної інформації // Вісник Львівської політехніки: – Львів, 2003. – №481. – С. 94-100.
2. З. Шпак. Регулювання тривалості тональних звуків зі збереженням структури мовного сигналу // Вісник Львівської політехніки. – Львів, 2003. – №496. – С. 228-235.