

13. Giacobazzi R. and Scozzari F. A Logical Model for Relational Abstract Domains. ACM Transactions on Programming Languages and Systems, 20(5):1998. – P.1067–1109.

14. Шекета В.І. Модифікаційні предикатні запити, як інструмент підтримки діалогу з користувачем в інформаційних системах на основі баз даних і знань // Вісник Тернопільського держ. технічн. ун-ту. –Серія "Математичне моделювання". – 2003.

З. Шпак

Національний університет "Львівська політехніка"

УДК 681.84.087.4

РЕГУЛЮВАННЯ ТРИВАЛОСТІ ТОНАЛЬНИХ ЗВУКІВ ЗІ ЗБЕРЕЖЕННЯМ СТРУКТУРИ МОВНОГО СИГНАЛУ

© Шпак З., 2003

Розроблено алгоритм синхронного з основним тоном часового перетворення мовних сигналів тональних звуків для задач високоякісного сповільненого відтворення голосових записів. Розтягування тривалості звуку здійснюється узгодженим повторенням оригінальних квазіперіодів основного тону, чим забезпечується практична незмінність спектральної структури звуку для широкого діапазону значень коефіцієнтів зміни темпу.

The algorithm of pitch-synchronized time-scale modification of voiced phonemes is presented. The time-domain expansion of sounds duration is performed by introducing in speech wave the additional segments consisted of transitional and original pitches signal. In order to keep the phone structure regularity, the weighted overlap-add technique has been used while composing the transitional pitches. Suggested method preserves the spectral structure of voiced phonemes and high naturalness of slowdown reproduced speech.

Вступ

Щоб досягти високої розбірливості й натуральності звучання голосових повідомлень, відтворених у зміненому темпі, процеси часового масштабування мовних сигналів повинні базуватись на природних темпоральних закономірностях. Визначальними для збереження натуральності голосу мовця є темпоральні перетворення тональних звуків, які у звичайному мовленні займають близько 74% часу загальної тривалості звукового потоку (без пауз). Часове масштабування тональних звуків у разі зміни темпу

є найбільшим, зокрема, перехід від швидкого до протяжного мовлення пов'язаний зі зростанням тривалості голосних звуків приблизно в 4 рази [1]. Тому якість сповільнено/прискорено відтвореного мовного запису значною мірою залежить від адекватності часового перетворення сигналів тональних звуків до їх природних трансформацій при зміні темпу мовлення.

Запропонована в статті схема збільшення тривалості тональних звуків є складовою частиною методу диференційованого темпорального перетворення елементів голосових записів і подальшим розвитком алгоритмів синхронної з основним тоном часової модифікації мовних сигналів [2,3]. Масштабування тональних звуків виконується у часовому полі, що дозволяє найповніше зберегти їх оригінальні параметри та загальну структуру. Тривалість звуку збільшується шляхом введення додаткових квазіперіодів основного тону (КПОТ), максимально узгоджених з поточною структурою мовного сигналу; кількість додаткових КПОТ розраховується окремо для перехідних та стаціонарних ділянок так, щоб сумарно забезпечити заданий коефіцієнт зміни тривалості звуку.

Часові перетворення тональних звуків

Група тональних звуків об'єднує голосні, сонорні та дзвінки приголосні звуки, які характеризуються рядом спільних артикуляційно-акустичних властивостей, пов'язаних з голосовим джерелом їх формування.

Основною характеристичною відмінністю мовних сигналів тональних звуків є присутність квазіперіодичних коливань з частотою основного тону, зміна якої відображає просодичну лінію мовного повідомлення. Другою важливою рисою є наявність формантних екстремумів, розподіл яких відрізняє звуки між собою. Дослідженнями встановлено, що природне сповільнення швидкості мовлення не викликає помітних відхилень ні у загальній формі мелодичного контуру фраз, ні у значеннях частот основного тону в окремих ділянках тональних звуків [3]. Практично незмінним залишається і розподіл формантних параметрів. Водночас для всіх тональних звуків зазначено істотне збільшення кількості КПОТ.

З огляду на природні закономірності темпоральних модифікацій тональних звуків найбільш адекватним способом їх часового перетворення визнано доповнення мовного сигналу звуку додатковими квазіперіодами, які відтворюють поточну спектральну структуру мовної хвилі. Варіант звичайного повторного відтворення одного або декількох послідовних КПОТ не відповідає цій умові через порушення гладкості сигналу в точках повторення. Тому запропоновано в місцях доповнення вводити перехідні композитні квазіперіоди, сформовані з двох сусідніх, сигнали яких треба взаємоузгодити.

Зміна швидкості мовлення викликає неоднакові перетворення стаціонарної і перехідних ділянок звуків. Якщо зміни тривалості стаціонарних є достатньо регулярними і мають визначену залежність від значення загального коефіцієнта сповільнення [1], то трансформації перехідних ділянок характеризуються великою варіативністю навіть для однакових звуків. У більшості випадків змінюються тривалості як початкової, так і кінцевої перехідних ділянок, проте зміна їх тривалості є усереднено в 1,4 (а в разі значних сповільнень – в 1,7) раз меншою, ніж відповідна відносна зміна стаціонарних частин звуків.

Оскільки перехідні ділянки відіграють важливу роль у передаванні через мовний сигнал корисної інформації, а їх темпоральні зміни є невеликими і характеризуються нестабільністю, то залишаємо коротку початкову і кінцеву перехідні ділянки звуку без змін. Відповідно коректуємо значення коефіцієнта збільшення тривалості стаціонарної частини звуку $k_v^{(st)}$:

$$k_v^{(st)} = \frac{k_v t_v - (t_{beg}^{(v)} + t_{end}^{(v)})}{t_v - (t_{beg}^{(v)} + t_{end}^{(v)})}, \quad (1)$$

де k_v – коефіцієнт зміни тривалості даного тонального звуку; t_v – загальна тривалість цього звуку; $t_{beg}^{(v)}, t_{end}^{(v)}$ – тривалості початкової та кінцевої перехідних ділянок, що встановлюються на етапі звукоподілу залежно від місця звуку в слові та темпорального класу попереднього (наступного) звуків.

Для максимального збереження природної структури мовної хвилі тонального звуку алгоритм часового перетворення будемо так, щоб, з одного боку, мінімізувати кількість додатково введених композитних КПОТ, а з іншого – уникнути довгих ділянок повторення, що також негативно впливають на природність звучання відтвореного сигналу. Виділяємо три вітки алгоритму: перша використовується для коефіцієнтів зміни тривалості $k_v^{(st)} \leq 1.5$, друга – для значень $1.5 < k_v^{(st)} \leq 3$, а третя – для значень $k_v^{(st)} > 3$.

Алгоритм перетворення у разі $k_v^{(st)} \leq 1.5$.

Кількість додаткових квазіперіодів у цьому випадку не перевищує третини від загальної кількості КПОТ, тому мовний сигнал доповнюємо тільки узгоджувачими перехідними квазіперіодами. Основні кроки алгоритму:

1. Встановлюємо початкові значення параметрів перетворення. Праворуч по часовій осі від початку стаціонарної частини звуку знаходимо точку $t_1^{(max)}$ максимуму першого КПОТ ділянки стаціонару (рис. 1а), від якої будемо відраховувати всі наступні КПОТ даного тонального звуку. Приймаємо:

$$t_{or}^{(0)} = t_1^{(max)} - t_{beg}^{(v)}; \quad \Delta t_{add}^{(0)} = 0; \quad i = 0, \quad (2)$$

де i – номер останнього відтвореного оригінального КПОТ; $t_{or}^{(i)}$ – тривалість поточного сегмента оригінального сигналу після відтворення i -го КПОТ; $\Delta t_{add}^{(i)}$ – поправка, яка враховує похибку значення перетвореної тривалості, що залишилась після виведення попереднього додаткового сегмента.

2. Відтворюємо наступний оригінальний КПОТ, збільшивши значення i та додаємо його тривалість $T_p^{(i)}$ до значення тривалості поточного оригінального сегмента: $t_{or}^{(i)} = t_{or}^{(i-1)} + T_p^{(i)}$. Обчислюємо тривалість додаткового

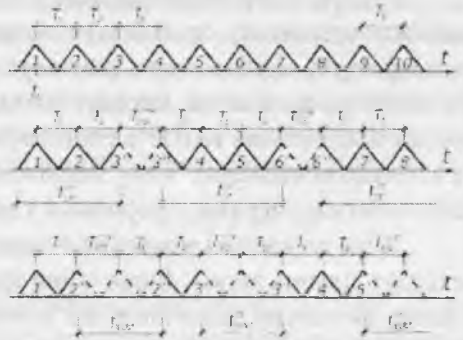


Рис. 1. Схеми синхронного з основним тоном збільшення тривалості тональних звуків: а – оригінальна структура звуку; б – схема доповнення композитними КПОТ у разі $s_p = 1,3$; в – схема введення додаткових сегментів у разі $s_p = 2,5$.

сегмента, яким треба доповнити вже відтворений оригінальний, щоб забезпечити заданий коефіцієнт зміни тривалості звуку $k_v^{(st)}$, враховуючи похибку, що нагромадилась на попередніх кроках виведення додаткових КПОТ:

$$t_{add}^{(i)} = t_{or}^{(i)} (k_v^{(st)} - 1) + \Delta t_{add}^{(i-1)} \quad (3)$$

3. Перевіряємо умову введення додаткового КПОТ. Якщо $t_{add}^{(i)} > 0.5T_P^{(i)}$, то виконуємо дії наступних кроків для формування та виведення композитного КПОТ; інакше, якщо не досягнуто кінця звуку, то повертаємось до кроку 2, щоб відтворити ще один квазіперіод оригінального мовного сигналу.
4. Формуємо та відтворюємо додатковий композитний КПОТ тривалістю

$$T_{Pcmp}^{(i)} = (T_P^{(i)} + T_P^{(i+1)}) / 2. \quad (4)$$

Спосіб формування перехідних композитних квазіперіодів описано нижче.

5. Визначаємо різницю між розрахунковою тривалістю додаткового сегмента та реально відтвореною: $\Delta t_{add}^{(i)} = t_{add}^{(i)} - T_{Pcmp}^{(i)}$, що буде становити похибку тривалості перетвореного сигналу після відтворення i -го оригінального КПОТ. Встановлюємо $t_{or}^{(i)} = 0$ і переходимо до кроку 2 для виведення наступної ділянки оригінального мовного сигналу.
6. Коли досягнуто кінця стаціонарної частини тонального звуку, завершуємо його часове перетворення виведенням кінцевої ділянки звуку тривалістю $t_{end}^{(v)}$.

Рис. 1б графічно ілюструє описану схему розтягування мовного сигналу тонального звуку додатковими перехідними КПОТ для випадку, коли коефіцієнт збільшення тривалості $s_v^{(st)} = 1,3$. Композитний період (позначений штриховими лініями) вводиться перший раз після відтворення двох оригінальних квазіперіодів, а в наступній частині стаціонарної ділянки – регулярно після кожних трьох або чотирьох КПОТ.

Часові перетворення у разі $1.5 < k_v^{(st)} \leq 3$

Масштабування тривалості звуку більше ніж у півтора рази пов'язане з введенням у мовний сигнал великої кількості додаткових КПОТ (так, у випадку $k_v^{(st)} = 3$ після кожного оригінального квазіперіоду треба відтворювати два додаткові). Якщо доповнення сигналу виконувати кількарізним повторенням композитних КПОТ, то це призводить до певних змін у спектральній структурі сигналу, що проявляються у частковій втраті чистоти і натуральності відтвореного голосу.

Щоб зменшити негативний вплив композитних КПОТ, скорочуємо вдвічі кількість додатково введених сегментів, компенсуючи це скорочення збільшенням їх тривалості. Кожен додатковий сегмент формуємо з двох КПОТ: композитного, що узгоджує мовні сигнали двох оригінальних квазіперіодів – $(i + 1)$ -го та $(i - 1)$ -го та поточного i -го, що повторюється. Тим самим доповнюємо наступний оригінальний сегмент ще одним квазіперіодом, а кількість композитних КПОТ зменшуємо в два рази.

Основні кроки описаного алгоритму часового перетворення мовного сигналу тонального звуку залишаються таким же, тому наведемо тільки необхідні зміни та доповнення.

1, 2. Виконуємо без змін.

3. Коректуємо умову виведення композитного КПОТ та квазіперіоду повторення:

$$t_{add}^{(i)} > T_P^{(i)}.$$

4. Композитний КПОТ формуємо, накладаючи на сигнал наступного $(i + 1)$ -го квазіперіоду мовну хвилю попереднього $(i - 1)$ -го КПОТ. Тривалість композитного квазіперіоду дорівнює півсумі тривалостей компонентних КПОТ

$$T_{Pcmp}^{(i)} = (T_P^{(i-1)} + T_P^{(i+1)}) / 2.$$

5. Після відтворення композитного КПОТ виводимо повторно i -й оригінальний квазіперіод. Визначаємо похибку тривалості перетвореного сигналу:

$$\Delta t_{add}^{(i)} = t_{add}^{(i)} - T_{Pcmp}^{(i)} - T_P^{(i)}. \quad (5)$$

6. Після виведення останнього квазіперіоду стаціонарної частини звуку додатково перевіряємо умову $t_{add}^{(i)} > 0.5T_P^{(i)}$. Якщо вона виконується, то формуємо і відтворюємо останній композитний квазіперіод на основі $(i+1)$ -го та i -го оригінальних КПОТ.

На рис. 1в проілюстровано схему розтягування тривалості звуку для випадку $k_v^{(st)} = 2.5$. Композитний КПОТ та квазіперіод повторення виводяться регулярно після відтворення одного або двох оригінальних квазіперіодів.

Часові перетворення в разі $k_v^{(st)} > 3$

Тестовими аудіовипробуваннями встановлено, що у випадках, коли коефіцієнт зміни тривалості звуку перевищує 3 (таке збільшення є характерним для голосних звуків у разі значного сповільнення темпу [1]), найкращу якість звучання промасштабованих голосових записів забезпечує схема перетворення, яка містить повторне виведення додаткових сегментів, складених з композитного і оригінального КПОТ. Це дозволяє уникнути довгих ділянок повторення, що можуть сприйматись на слух як реверберація звуку і водночас незначно збільшити кількість перехідних квазіперіодів.

Таке перетворення повністю базується на попередньому алгоритмі; доповнення вводяться після кроку 5, що виконує виведення додаткових сегментів і корекцію похибки відтворення. Якщо залишкова похибка $\Delta t_{add}^{(i)} > T_P^{(i)}$, то знову відтворюємо вже сформований композитний квазіперіод тривалістю $T_{Pcmp}^{(i)}$ і поточний оригінальний КПОТ. Зменшуємо значення похибки:

$$\Delta t_{add}^{(i)} = \Delta t_{add}^{(i)} - T_{Pcmp}^{(i)} - T_P^{(i)}, \quad (6)$$

і переходимо до кроку 2 для відтворення наступного оригінального КПОТ.

Після відтворення останнього КПОТ мінімізуємо залишкову похибку виведенням сегмента з одного, двох або трьох додаткових квазіперіодів.

В основу розробленого алгоритму часового масштабування тональних звуків покладено процедуру розбиття мовного сигналу звуку на послідовність квазіперіодів основного тону. В [4] запропоновано алгоритм швидкого поділу мовного сигналу звуку на послідовність КПОТ, що встановлює границі квазіперіодів у ділянках їх амплітудних максимумів для найкращого збереження природних параметрів сигналу в перехідних композитних КПОТ.

максимумів для найкращого збереження природних параметрів сигналу в перехідних композитних КПОТ.

Формування композитних КПОТ

Іншою важливою процедурою алгоритму зміни тривалості звуків є формування додаткових перехідних КПОТ, параметри яких повинні відповідати поточній структурі мовного сигналу і взаємоузгоджувати сусідні оригінальні КПОТ у точках повторень.

При формуванні перехідних КПОТ необхідно враховувати можливу різницю у значеннях тривалостей сусідніх КПОТ, з яких формується композитний: $\Delta T_p^{(i)} = |T_p^{(i+1)} - T_p^{(i)}|$ (для випадків $k_v^{(st)} \leq 1.5$), чи $\Delta T_p^{(i)} = |T_p^{(i+1)} - T_p^{(i-1)}|$ (для випадків $k_v^{(st)} > 1.5$), відносне значення якої може досягати 12% [3]. Тому приймаємо, що тривалість композитних КПОТ дорівнює півсумі тривалостей компонентуючих квазіперіодів (4).

Щоб забезпечити плавність зміни мовного сигналу композитного КПОТ і його узгодженість з навколишніми оригінальними квазіперіодами, на компонентуючі сигнали накладаємо лінійні функції зважування:

$$w_1^{(i)}(t) = w_i(t) = t / T_{Pcmp}^{(i)}; \quad w_2^{(i)}(t) = 1 - w_i(t), \quad 0 \leq t < T_{Pcmp}^{(i)}. \quad (7)$$

Композитний КПОТ, що відтворюється після i -го оригінального квазіперіоду, формуємо так:

$$x_{cmp}^{(i)}(t) = w_1^{(i)}(t)x_1^{(i)}(t) + w_2^{(i)}(t)x_2^{(i)}(t) = w_i(t)x\{t_0^{(x_1)} + t - w_i(t)\Delta T_p^{(i)} / 2\} + [1 - w_i(t)]x\{t_0^{(x_2)} + t + w_i(t)\Delta T_p^{(i)} / 2\}, \quad 0 \leq t < T_{Pcmp}^{(i)}, \quad (8)$$

де $x_1^{(i)}(t)$ – сигнал оригінального КПОТ, що передує композитному: i -го або $(i - 1)$ -го; $x_2^{(i)}(t)$ – сигнал $(i+1)$ -го оригінального КПОТ, наступного за композитним; $t_0^{(x_1)}$, $t_0^{(x_2)}$ – моменти початків компонентуючих квазіперіодів.

У практичних реалізаціях процесу зміни тривалості звуку ми оперуємо з про- дискретизованим сигналом, тобто обчислення відліків композитного КПОТ виконуємо за формулою

$$x_{cmp}^{(i)}(n) = w_i(n)x_1^{(i)}(n) + [1 - w_i(n)]x_2^{(i)}(n), \quad n = 0, N_{Pcmp}^{(i)} - 1, \quad (9)$$

де $N_{Pcmp}^{(i)} = T_{Pcmp}^{(i)}F_s$; F_s – частота дискретизації мовного сигналу.

Це вимагає приведення розмірів оригінальних квазіперіодів, з яких формується композитний, до встановленого середнього значення $N_{Pcmp}^{(i)}$, тобто інтерполяції проміжних відліків компонентуючих КПОТ. Позначимо:

$$\alpha_1 = T_p^{(x_1)}F_s / N_{Pcmp}^{(i)}; \quad m_1 = [n\alpha_1]; \quad \alpha_2 = T_p^{(x_2)}F_s / N_{Pcmp}^{(i)}; \quad m_2 = [n\alpha_2],$$

де $[]$ – функція цілої частини.

Тоді

$$\begin{aligned}
 x_1^{(i)}(n) &= x_1^{(i)}(m_1)(1 - n\alpha_1 + m_1) + x_1^{(i)}(m_1 + 1)(n\alpha_1 - m_1) ; \\
 x_2^{(i)}(n) &= x_2^{(i)}(m_1)(1 - n\alpha_2 + m_2) + x_2^{(i)}(m_2 + 1)(n\alpha_2 - m_2) , \quad n = 0, \overline{N_{\text{смп}}^{(i)} - 1}.
 \end{aligned}
 \tag{10}$$

Підставляючи (10) в (9), отримуємо значення n -го відліку поточного композитного квазіперіоду.

Висновки

Зазначимо основні переваги розробленого методу темпорального перетворення мовних сигналів тональних звуків:

- незмінність перехідних ділянок звуків, що є важливим для розпізнавання мови;
- забезпечення заданої тривалості результуючого сигналу – похибка перетворення кожного звуку не перевищує $\pm 0,5T_p$ для всіх значень коефіцієнта зміни тривалості; оскільки значення похибок є випадковими величинами, то в процесі відтворення послідовності тональних звуків відбувається взаємна компенсація похибок масштабування окремих звуків;
- формування перехідних композитних КПОТ виконується за схемою, що дозволяє максимально зберегти оригінальні параметри обох компонентів квазіперіодів, оскільки найменших змін зазнають ділянки амплітудних максимумів сигналу, які мають визначальний вплив на якість відтворення звуку;
- кількість композитних КПОТ не перевищує третини від кількості оригінальних квазіперіодів звукового сигналу, а максимальний розмір сегмента повторення дорівнює двом квазіперіодам;
- запропонований підхід можна успішно використати для скорочення тривалості тональних звуків шляхом вилучення з мовного сигналу сегментів, що складаються з двох або трьох послідовних КПОТ, і введення на їх місце перехідних

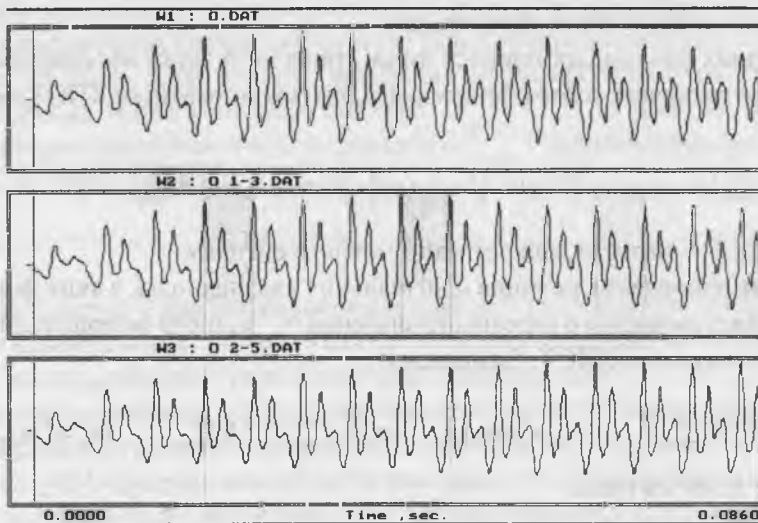


Рис. 2. Перетворення мовного сигналу звуку о : а – оригінальний мовний сигнал; б – результат збільшення тривалості в 1,3 раза; в – результат розтягування звуку в 2,5 раза

Темнішим кольором виділені сегменти додаткових КПОТ – їх кількість і точки введення у мовну хвилю повністю відповідають схемам рис. 1.

Аудиторські випробування підтвердили, що запропонований алгоритм синхронного з основним тоном часового масштабування тональних звуків дозволяє зберегти оригінальну спектральну структуру мовних сигналів і тим самим забезпечити натуральність та чистоту звучання відтворених голосових записів, навіть якщо коефіцієнт зміни темпу досягає значення 3.

1. Шпак З., Рашкевич Ю. Збереження природних темпоральних закономірностей у процесах сповільненого відтворення мовної інформації // Вісник Нац. ун-ту "Львівська політехніка", 2003. – № 481. – С. 94–100.
2. Рашкевич Ю.М. Перетворення часового масштабу мовних сигналів. – Львів: Академічний експрес, 1997. – 140 с.
3. Шпак З.Я. Синхронне з основним тоном збільшення часового масштабу мови // Вісник Держ. ун-ту "Львівська політехніка". – 1998. – № 351. – С. 177–185.
4. Шпак З. Виділення квазіперіодів основного тону в процесі часового масштабування мовних сигналів // Технічні вісті. – 2003. – № 1(16), 2(17). – С. 61–63.

О. Різник, Б.Балич

Національний університет "Львівська політехніка"

УДК 621.382:519.101

СИНТЕЗ МОНОЛІТНИХ КОДІВ НА ОСНОВІ ЧИСЛОВИХ ЛІНІЙОК-В'ЯЗАНОК СПЕЦІАЛЬНОГО ТИПУ

© Різник О., Балич Б., 2003

Розглядається метод дослідження комбінаторних конфігурацій за допомогою цілочислових послідовностей – в'язанок як зручних математичних моделей для синтезу та оптимізації систем з нееквідистантною структурою.

There is method for research of combinatorial configurations is considered using integer sequences - bundles as useful mathematical models for synthesis and optimization of systems with non-uniform structure.

Вступ

При обробці інформації необхідно враховувати наявність суттєвого недоліку двійкової системи числення, який полягає в нездатності виявляти та виправляти помилки. Актуальною проблемою є створення кодів з максимальною кодовою відстанню та помірною надлишковістю, які спроможні виявляти та виправляти помилки. Одним з таких кодів є монолітний код; будується монолітний код методом, заснованим на