

НАПРЯМКИ УДОСКНАЛЕННЯ ПРОЦЕДУР ПОПЕРЕДНЬОЇ ОБРОБКИ ГОЛОСОВИХ СИГНАЛІВ У СИСТЕМАХ АВТЕНТИФІКАЦІЇ

Камені Нгалані Гі Берті, Кіщенко М.І., Пастушенко М.С.

Кафедра «Інфокомунікаційної інженерії ім. В.В. Поповського»,
Харківський національний університет радіоелектроніки, Україна

E-mail: nhalani.kameni@nure.ua,
marharyta.kishchenko@nure.ua,
mykola.pastushenko@nure.ua

Abstract

The actual scientific problem of carrying out preprocessing of a voice signal in authentication systems is considered. In contrast to the well-known works in which pre-processing of signals is carried out using amplitude-frequency information, it is proposed to carry out the specified processing using the phase space. The prerequisite for this is that the phase signal changes linearly within the specified limits. Research has shown that this approach has been successful. A technique for conducting preliminary processing of voice signals in the phase space has been developed.

Питання обробки голосових сигналів в даний час набувають особливої актуальності. Зумовлено це, насамперед, істотним застосуванням голосових даних у різних практичних додатках.

Не винятком є і голосова автентифікація, яка все ширше входить у повсякденну практику людини. Стосовно голосової автентифікації дослідження проводяться за такими основними напрямками:

- вдосконалення відомих і розробка нових методів вилучення мовних ознак;
- формування адекватних шаблонів користувача;
- розвиток процедур прийняття рішення на основі тих чи інших ознак.

Проте ефективність процедур автентифікації значною мірою залежить від якості оброблюваних матеріалів реєстрації. У багатьох практичних додатках якість матеріалів реєстрації має бути краще. У той же час це не скасовує необхідність формування рішень і щодо неякісних матеріалів реєстрації. Поліпшити якість матеріалів реєстрації дозволяють процедури попередньої обробки голосових сигналів. Аналіз літератури в області попередньої обробки голосових сигналів показує, що вказана обробка проводиться за такими напрямками:

- підхід, який називається "шумові ворота" (noise gate), а саме, їх конкретний спектральний різновид, що використовується після швидкого перетворення Фур'є, також віконні механізми згладжування сигналу та видалення його невеликих артефактів;
- рішення, засновані на рекурентних нейронних мережах з Long short-term memory (LSTM) блоками та на згорткових нейронних мережах;
- відомий метод спектрального віднімання (spectral subtraction);
- методи шумопридушення на основі Вінерівського оцінювання;
- нейромережеве шумопридушення та поліпшення мовного сигналу на основі згорткових нейронних мереж Conv-TasNet.

Спільним всім цих підходів є те, що обробка проводиться у просторі амплітудно-частотних характеристик голосового сигналу. Зауважимо, що формування шаблонів користувача системи автентифікації та прийняття рішень також здійснюється в цьому просторі.

Останнім часом з'явилися ряд робіт, які додатково використовують фазові дані (фазовий простір) голосового сигналу, наприклад [1-5]. Саме такий підхід був використаний у даній роботі, який, як буде показано нижче, дозволяє розширити можливості вирішення завдань попередньої обробки голосового сигналу. Зауважимо, що фазовий простір голосового сигналу безпосередньо пов'язаний з поняттям аналітичного сигналу, який ефективно і продуктивно використовується в радіолокації та радіоз-

в'язку.

Що стосується області голосових сигналів сформувати фазові дані можна програмно за допомогою перетворення Гільберта.

Як показали результати досліджень фазові дані голосового сигналу мають форму пилкоподібного сигналу, амплітуда якого змінюється за лінійним законом від 0 до 360 градусів, а тривалість невідома. Тут суттєвим є те, що амплітуда змінюється за лінійним законом. За наявності неточностей в реєстрації або зовнішнього шуму будуть з'являтися випадкові помилки, які призводитимуть до відхилення змін амплітуди фазового сигналу від лінійного закону. Цей факт можна використовувати при проведенні попередньої обробки голосового сигналу (досліджувана гіпотеза).

Для перевірки зазначеної гіпотези проводилася цифрова обробка голосового сигналу частотою дискретизації 64 кГц. У процесі досліджень сформульовані основні положення методики проведення попередньої обробки голосового сигналу, яка включає:

- формування фазових даних;
- виділення пилкоподібних фазових сигналів;
- проведення лінійної апроксимації пилкоподібного фазового сигналу та квадратичної апроксимації модуля вектору (огинаючої) аналітичного сигналу;
- уточнення матеріалів реєстрації та уявної складової аналітичного сигналу з урахуванням результатів розрахунків;
- перевірка правильності уточнення одержаних результатів розрахунку за допомогою траєкторії кінця вектору аналітичного сигналу.

Таким чином, наукова новизна отриманих результатів полягає в тому, що вперше розроблено методику та проведено експериментальні дослідження проведення попередньої обробки голосового сигналу користувача з використанням простору фазових даних, а також простору зміни модуля (огинаючої) та траєкторії вектору аналітичного сигналу.

Крім цього, розроблені нові методи апроксимації фазових даних і огинальної аналітичного сигналу на періоді пилкоподібного зміни фази.

Література

1. *Pastushenko, M., Faizulaieva, O.* (2016), "Employment of phase characteristics of user voice signal in authentication systems," Third International Scientific-Practical Conference Problems of Infocommunications Science and Technology (PIC S&T), Kharkiv, pp. 205-206. DOI: [10.1109/INFOCOMMST.2016.7905382](https://doi.org/10.1109/INFOCOMMST.2016.7905382)
2. *Пастушенко, Н.С., Педро, В.Г., Файзулаева, О.Н.* (2018), Исследование информативности фазовых данных голосового сигнала пользователя системы аутентификации. [Электронный ресурс] // Проблемы телекоммуникаций. № 1(22), pp. 67-74. DOI: <https://doi.org/10.30837/pt.2018.1.05>
3. *Pastushenko, M., Pastushenko, V., Pastushenko, O.* (2019), "Specifics of Receiving and Processing Phase Information in Voice Authentication Systems", International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), Kyiv, Ukraine, 2019, pp. 621-624. DOI: [10.1109/PICST47496.2019.9061260](https://doi.org/10.1109/PICST47496.2019.9061260)
4. *Pastushenko, M., Krasnozheniuk, Ya., Lemeshko, O.* (2020), Analysis of voice signal phase data informativity of authentication system // Zaporizhzhia, Ukraine, April 27-May 1, 2020. Proceedings of The Third International Workshop on Computer Modeling and Intelligent Systems (CMIS-2020). PP 1040-1053. URI: <http://openarchive.nure.ua/handle/document/11843>
5. *Pastushenko, M., Krasnozheniuk, Ya., Zaika, M.* (2020), "Investigation of Informativeness and Stability of Mel-Frequency Cepstral Coefficients Estimates based on Voice Signal Phase Data of Authentication System User," International Conference "Problems of Infocommunications. Science and Technology" (PIC S&T'2020), pp. 1-5. DOI: [10.1109/PICST51311.2020.9468083](https://doi.org/10.1109/PICST51311.2020.9468083)