

АНАЛІЗ ГОЛОСОВИХ ПОВІДОМЛЕНЬ ПРИ ВИЗНАЧЕННІ НАСТРОЮ КОРИСТУВАЧА МОБІЛЬНОГО ТЕЛЕФОНУ

Кравець К. І., бакалавр, kiber@dgma.donetsk.ua;

Шевченко Н. Ю., к.е.н, доцент, natalygustav@gmail.com

Донбаська державна машинобудівна академія, Краматорськ, Україна

Емоції зазвичай визначають поведінку людини, впливають на рішення, які вона приймає. Саме тому визначення, а в деяких випадках і підтримка гарного емоційного стану користувача мобільного пристрою може відігравати важливу роль при використанні деяких додатків, а особливо тих, що пов'язані з комерцією. У зв'язку з цим, автоматична і об'єктивна діагностика емоційного стану людини представляє практичний інтерес, а розробка мобільного додатку, здатного розпізнавати наявність певної емоції за голосом користувача мобільного пристрою, наприклад, задля покращення його емоційного стану, є актуальним питанням сьогодення.

Пропонується наступна схема розпізнавання емоцій (рис. 1).

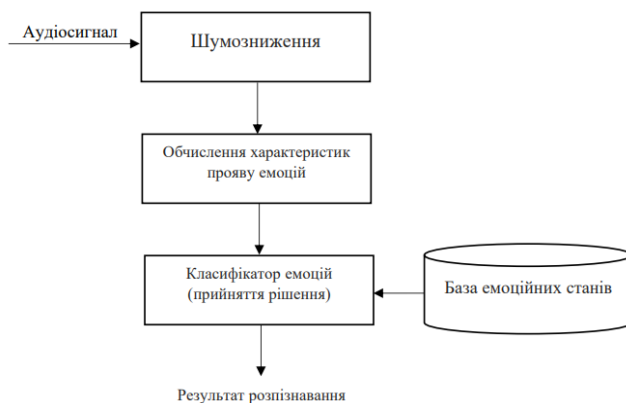


Рисунок 1. Структурна схема системи розпізнавання емоцій користувача за індивідуальними характеристиками голосу

До основних етапів системи розпізнавання емоцій користувача на основі аналізу його голосу через відповіді на певні запитання можна віднести:

1. Шумозниження [1]. У кожному звуковому фрагменті виконується розпізнавання наявності та зменшення амплітуди адитивних стаціонарних шумів за допомогою спектрального віднімання смуг, у яких не фіксується людський голос.

2. Обчислення характеристик прояву емоцій за наборами відрізків відповідей користувача, формування векторів ознак для подальшої класифікації [1]. Простір ознак містить сім акустичних, просодичних та екстралінгвістичних характеристик: висота голосу, сила голосу, відстань між форматами у голосних звуках, тривалість вимови складів, тривалість пауз, наявність кашлю, зітхань, плачу, сміху, зміна інтонації у реченні.

3. Для отримання акустичних характеристик голосу використовуються згладжені спектри широких фонетичних класів, за якими розраховується висота голосу (основний тон), сила голосу та відстань між форматами у голосних звуках [1]. В основі спектрів широких фонетичних класів лежать Вейвлет-перетворення [2]. Характеристики тривалості вимови складів та пауз між словами обчислюються усереднено по реченню, а зміна інтонації у реченні є динамічною просодичною характеристикою. Для спрощення задачі

класифікації інтонаційних особливостей різних мов, часова характеристика інтонації (висота тону) сегментується на наступні ділянки для кожного речення: підвищення тону до кінця речення; підвищення тону на початку речення; спад тону усередині речення; підвищення тону усередині речення. Екстралінгвістичними характеристиками є наявність у фрагментах пауз (шуму) однієї з акустичних подій, притаманних обраним емоціям: кашлю, зітхань, плачу, сміху.

4. На етапі прийняття рішення про емоційний стан користувача використовується сформований набір векторів ознак, який оцінюється за допомогою вирішального правила. Простим класифікатором виступає метод сумішей Гауса, який виконує оцінку приналежності набору векторів ознак до еталонної моделі користувача. Емоційна модель користувача представляє собою дані суміші Гауса, отримані за сьомою характеристиками голосу, і п'ять базових емоційних станів. Відповідність певного набору характеристик до одного з базових емоційних станів визначається за допомогою нейронної мережі, а саме багатошарової нейронної мережі прямого розповсюдження з однорідними нейронами. За методом навчання – навчання з вчителем, за типом вхідної інформації — двійкова мережа. На вході нейронна мережа отримує двійкову інтерпретацію звукової доріжки, а на виході — відповідність звукової доріжки певному емоційному стану.

Для навчання алгоритмів розпізнавання використані дві мовні бази даних, що складаються із записів мовлення, котрі анотуються до передбачуваних емоцій мовця та / або емоцій, сприйнятих групою слухачів:

- Emo_DB (Burkhardt, Paeschke, Rolfes, Sendlmeier & Weiss 2005), яка містить записи десяти акторів, кожен з яких виголошує кілька речень із нейтральним семантичним змістом, вироблених з кожною із семи передбачуваних емоцій: щастя, смутку, гніву, страху, нейтралі, огиди та нудьги;

- Surrey Audio-Visual Expression Emotion (Savee) (Haq, Jackson & Edge 2008), яка містить записи чотирьох акторів, кожен з яких виконує 120 висловлювань, розподілених на сім передбачуваних емоцій: щастя, смутку, гніву, страху, нейтральності, огиди та здивування.

У відповідності до наведеного алгоритму мобільний додаток визначення настрою користувача мобільного телефону має наступний функціонал (з позиції інструмента боротьби зі стресом та поганим настроєм):

- аналіз емоційного стану людини за голосом;
- надання мотиваційних повідомлень протягом дня, базуючись на визначеному настрої (емоційному стані);
- надання гумористичного побажання на день, опираючись на визначений емоційний стан користувача.

Додаток розроблений для ОС Android 5.0 з використанням інструментального засобу розробки додатків Android Studio спільно з Android SDK на мові Kotlin.

У момент першого запуску додатку «Daily mood» користувач потрапляє до вікна реєстрації, де йому пропонується ввести ім'я для персоніфікації взаємодії з додатком (рис. 1).

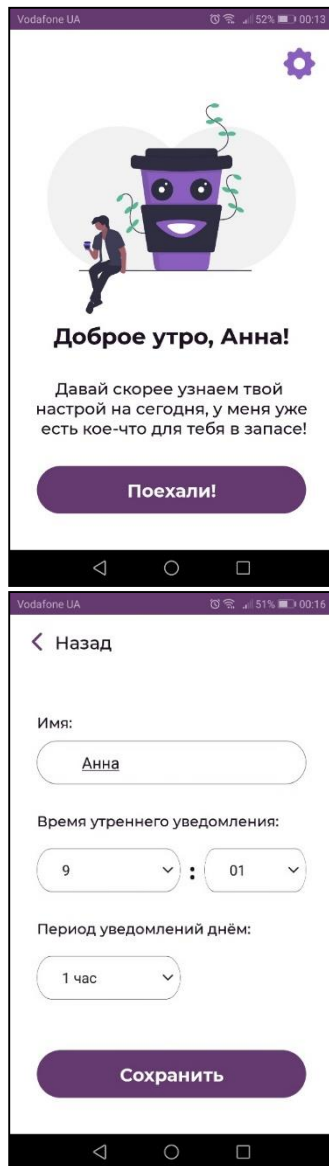


Рисунок 1. Вікна вітання та налаштувань

Після авторизації та кожного наступного разу при запуску додатку користувач потрапляє до вікна вітання, що пропонує розпочати аналіз настрою або перейти до налаштувань.

На наступному етапі користувач потрапляє до вікна питань, де додаток пропонує йому відповісти на 3 різних запитання, послідовно даючи відповідь за допомогою голосу (рис. 2а).

Після аналізу відповідей користувача, за допомогою нейронної мережі, додаток робить висновок про настрій користувача та оголошує його у вікні результату (рис. 2б) та пропонує переглянути побажання на день (рис. 2в).

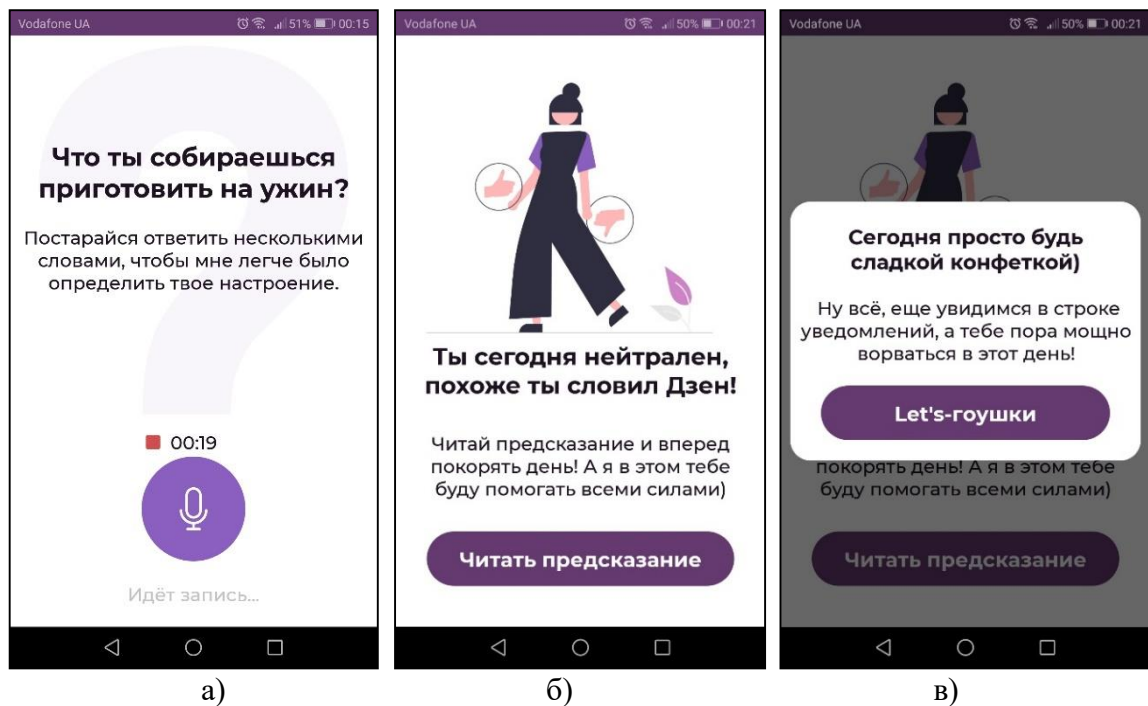


Рисунок 2. Вікна додатку:
а) питання; б) аналізу настрою; в) побажання на день

Протягом дня після проходження опитування користувачеві будуть надсилатися повідомлення з підбадьорюючими висловами згідно заданого часового інтервалу та поточного настрою користувача. Кожного дня додаток «Daily mood» у визначений ранковий час пропонуватиме власнику мобільного телефону пройти опитування та дізнатися свій настрій.

Запропонований мобільний додаток може стати початком широкого спектру програмних комплексів з аналізу та обробки голосових даних.

Технологія розпізнавання емоційного стану може бути використана у різноманітних сферах, таких як: соціальні мережі, маркет-плейси, медицина, а також відігравати роль потужного інструменту для збору даних про користувача (Big Data).

Література

1. Клименко М. С. Розробка структури системи розпізнавання емоційного стану диктора / М. С. Клименко, Ф. В. Фомін // Штучний інтелект. — 2016. — № 1. — С. 17–26.
2. Вейвлет-преобразование [Електронний ресурс]. — URL: Вейвлет-преобразование (gwyddion.net).

Анотація

Запропонована структурна схема системи розпізнавання емоцій користувача за індивідуальними характеристиками голосу та розглянуті основні етапи цієї системи. Наведений опис розробленого мобільного додатку, здатного розпізнавати наявність певної емоції за голосом користувача мобільного пристрою.

Ключові слова: розпізнавання емоцій, голосові повідомлення, користувач мобільного телефону, мобільний додаток.

Аннотация

Предложена структурная схема системы распознавания эмоций пользователя по индивидуальным характеристикам голоса и рассмотрены основные этапы этой системы. Приведено описание разработанного мобильного приложения, способного распознавать наличие определенной эмоции по голосу пользователя мобильного устройства.

Ключевые слова: распознавание эмоций, голосовые сообщения, пользователь мобильного телефона, мобильное приложение.

Abstract

The flow diagram of the recognition system of user's emotions is offered on individual descriptions of voice and the basic stages of this system are considered. Description over of the worked out mobile application able to recognize the presence of certain emotion on voice of mobile device user is brought.

Keywords: recognition of emotions, vocal reports, user of mobile telephone, mobile application.