

ГОЛОСОВОЕ УПРАВЛЕНИЕ КАК СПОСОБ ВЗАИМОДЕЙСТВИЯ ЧЕЛОВЕКА С МАШИНОЙ

Гелетто В.М.

Научный руководитель – Филиппенко И.В.

Харьковский национальный университет радиоэлектроники

(61166, Харків, пр.Науки, 14, каф. АПВТ, (057)702-13-54)

e-mail: violetta.heletto@nure.ua

A voice-user interface (VUI) makes spoken human interaction with computers possible, using speech recognition to understand spoken commands and answer questions, and typically text to speech to play a reply. A voice command device (VCD) is a device controlled with a voice user interface. Voice user interfaces have been added to automobiles, home automation systems, computer operating systems, home appliances like washing machines and microwave ovens, and television remote controls.

Голосове управління вже не новинка, а цілком усвідомлена реальність. Подібні технології все більше і більше проникають в наше життя. Цю функцію вже мають смартфони, телевізори та інша техніка, яка керується пультами з вбудованою обробкою голосу.

Людський голос, що подається на пристрій, спочатку оцифровується і передається на пульт або інший пристрій управління, у вигляді цифрових сигналів, а потім уже на кінцевий пристрій. Для того щоб це зробити, виконується декілька операцій. Спочатку подається аналоговий сигнал, який дискретизується за допомогою періодичної вибірки. Після цього, аналоговий сигнал квантується шляхом приписування аналоговій величині найближчого цифрового значення. Також їх фільтрують, найчастіше використовують фільтр для обмеження нижніх частот, так як основна інформація передається на частоті до 4000 Гц. Далі оцифровані дані потрапляють на подальшу обробку.

Системи розпізнавання голосу приймають оцифровані дані та виділяють з них голосові команди, які заздалегідь були додані в пам'ять пристрою. При цьому програма повинна вилучати смислові слова зі всього тексту. Методи за допомогою яких, визначають смислову одиницю мови на даний момент не є надійними, їх точність зворотно пропорційна кількості слів у тексті. Окремо потрібно подбати про шумозаглушення, бо завжди нас оточують шуми. Крім того, багато людей мають індивідуальну промову, специфіку вимови, акценти.

На даний момент більшість систем використовують модульну архітектуру, яка обладнана блоком шумозаглушення (видаляє шуми та викривлений звук для покращення якості переданої інформації), перетворювача сигналу у оцифровані вектори, детектора розпізнавання голосу та головного модуля, що містить пам'ять з записаними ключовими словами. Існує ряд проблем, які повинні бути вирішені для комфортного

використання даного способу взаємодії з пристроєм. На даний момент пристрої розпізнавання голосу не пристосовані до всіх фраз, які вимовляє людина. Також з'являється проблема з обмеженістю словникового запасу програми, так як досить складно завантажити в неї всі слова, які можуть знадобитися користувачеві. Крім того, не слід забувати про те, що у користувача з пристроєм повинен бути діалог, де воно буде інформувати про результат (або його відсутність) після виконання команди. Не варто забувати, що не у всіх випадках доречно управління голосом. Повинна бути виконана якась дія з боку користувача для повідомлення пристрою, що йому слід включитися в роботу, інакше можуть бути постійні збої в роботі за рахунок того, що функції будуть включатися і вимикатися випадковим чином. Також виникає багато проблем при спробах підлаштувати програми розпізнавання голосу для всіх мов світу. Найкраще вони працюють на англійській мові.

На даний момент йде активна розробка пристроїв і програм розпізнавання голосу. Окремо існують концепції пристроїв, які виступають в ролі посередника між людиною і, наприклад, розумним будинком (розумне кільце, годинник, окуляри і т.д.). Так само окремим видом розробок є програмні системи від таких провідних розробників як Apple, Google, Amazon, Microsoft. У цьому році на виставці технологій у Лас-Вегасі CES 2020 була представлена система від Samsung, названа NEON. Це штучний інтелект, який крім того, що управляється голосом так ще й до того ж має зовнішність. Це віртуальні істоти, які можуть демонструвати свої емоції і вчитися, спілкуючись зі своїм власником. При цьому вони здатні підтримати розмову і можуть стати своєрідним віртуальним другом. Така технологія, може дуже гарно виправити проблему з відсутністю діалогів. За допомогою штучного інтелекту вони можуть стати чудовими помічниками для користувачів.

Можливість управління голосом дуже важлива в розвитку технологій людства для спрощення наших життів і побуту. На даний момент вирішені такі проблеми як адаптація програм на інші мови (якість розпізнавання голосу в неідеальних умовах становить 95-98%). Але в найближчих планах у пріоритеті є рішення проблем з природною мовою, відсутність діалогу між користувачем і програмою, недостатнього словникового запасу машини.

Список використаних джерел:

1. Ахаді С. Ефективний інтерфейс для автоматичного розпізнавання мови //IEEE Trans. про обробку мови і звуку. – 2003 рік.
2. Роуз Р. Надійні методи розпізнавання мови, що застосовуються до мови проти шуму: Європейська конференція з мови Комунікація та технології, Ольборг, Данія, 3-7 Вересень 2001 року.