КОМП'ЮТЕРНІ ЗАСОБИ ДЛЯ ЕКСПЕРИМЕНТАЛЬНИХ ДОСЛІДЖЕНЬ МОВНОГО СИГНАЛУ

Микола Сажок

Міжнародний науково-навчальний центр інформаційних технологій та систем
40 просп. Академіка Глушкова, Київ 252022
Електронна пошта: mykola@uasoiro.freenet.kiev.ua

Program complex, presented in this paper, is a set of tools required for speech recognition, understanding, synthesis methods modelling and system architecture tests. All cycle of applied tasks is supported. Among them are speech signal recording, collecting, segmentation and natural languages knowledge accumulation. Implemented features allow setting experiments for speech signal transformations, pre-processing, recognition, understanding, and synthesis. The ways to make developed software open for custom research features are discussed as well as its compatibility with popular tools for scientific research.

1 Вступ

Комп'ютерне моделювання — необхідний етап у розробленні методів та тестуванні архітектур систем розпізнавання, розуміння та синтезу усної мови. Висновки щодо придатності, а також шляхи вдосконалення того чи іншого методу або архітектури системи випливають з результатів їх натурного моделювання [1].

Тому розроблення засобів експериментальних досліджень є життєво необхідним для розвитку усномовних технологій. На щастя, сучасні комп'ютерні технології якраз і дозволяють створювати складні, інтелектуалізовані дослідні станції. Отже, з'являється можливість зробити ці станції достатньо універсальними, придатними не лише для задач тої чи тої дослідницької групи, але й для будь-якого дослідника в ділянці усномовних технологій.

Програмний комплекс складається з двох основних підкомплексів: студія дослідника усномовного сигналу (ДУМС) та студія дослідника природніх мов (ДПМ).

Студія ДУМС — інтегроване робоче місце для опрацювання мовного сигналу, лінгвістично-фонетичних знань та індивідуальну інформацію щодо диктора. З іншого боку, за допомогою цього модуля проводяться експерименти з попереднього оброблення, розпізнавання та синтезу мовного сигналу.

Студія ДПМ забезпечує введення, накопичення та оброблення знань про природні мови, а також проведення експериментів з автоматичного розуміння мови та перекладу з однієї мови на іншу. Одна з компонент цієї студії відповідає за формування бази знань про конкретну природну мову, інша компонента задає певну предметну область, що є частиною моделі зовнішнього світу, яка вводиться для розуміння природної мови [2].

2 Студія дослідника усномовного сигналу

Студія ДУМС задумана як універсальне робоче місце для науковця-дослідника усномовниго сигналу. Студія дозволяє оперувати як самим акустичним сигналом, так і його лінгвістично-фонетичними атрибутами, враховувати індивідуальні особливості диктора.

Запис мовного сигналу проводиться в діалоговому режимі, частинами. Якщо якість певного блоку сигналу не задовольняє експерта, або ж про це сигналізують закладені в студію алгоритми, такий блок позначається як "поганий" з тим, щоби його перезаписати.

Наступним важливим кроком є автоматизована сегментація сигналу на синтагми, ритмогрупи, слова, склади, фонеми, мікрофонеми тощо. Експерт приймає, підправляє або відміняє результати автоматичного сегментування. Для підвищення ефективності роботи експерта забезпечується спрощений доступ до сегментів, напрклад, фонем (див. рис. 1). Отже, експерт може легко виділяти, проглядати та прослуховувати кожну ділянку сигналу, що відповідає певній фонемі.

Результати сегментації, коротка інформація про диктора зберігаються в стандартному звуковому (wave) файлі як окремий блок файлу (chunk), не пошкоджуючи, таким чином сам файл, залишаючи його доступним для інших програм. Більш докладна індивідуальна інформація може зберігатися в окремому файлі, причому посилання на нього зберігається в звуковому файлі.

Щойно усномовний сигнал відсегментовано, експерт накопичує сегменти фонем і оперує з ними як з окремими лінгвістино-акустичними одиницями, що економно задаються модельними елементами, якими є, наприклад, одно-квазіперіодні сегменти [3]. В бібліотеці сегментів фонем містяться посилання на акустичні дані, що зберігаються в звуковому файлі. Також передбачена можливість запису звукових даних безпосередньо в бібліотеку сегментів фонем.

Обслуговування всієї необхідної лінгвістично-фонетичної інформації відбувається наступним чином: кожному сегменту фонеми приписується ім'я фонеми, її фонемне оточення, що автоматично обчислюються за результатом сегментації. Крім того, запам'ятовуються фраза, позиція фонеми у фразі, найближчі границі синтагм, ідентифікатор диктора тощо.

© Ukrobraz-98

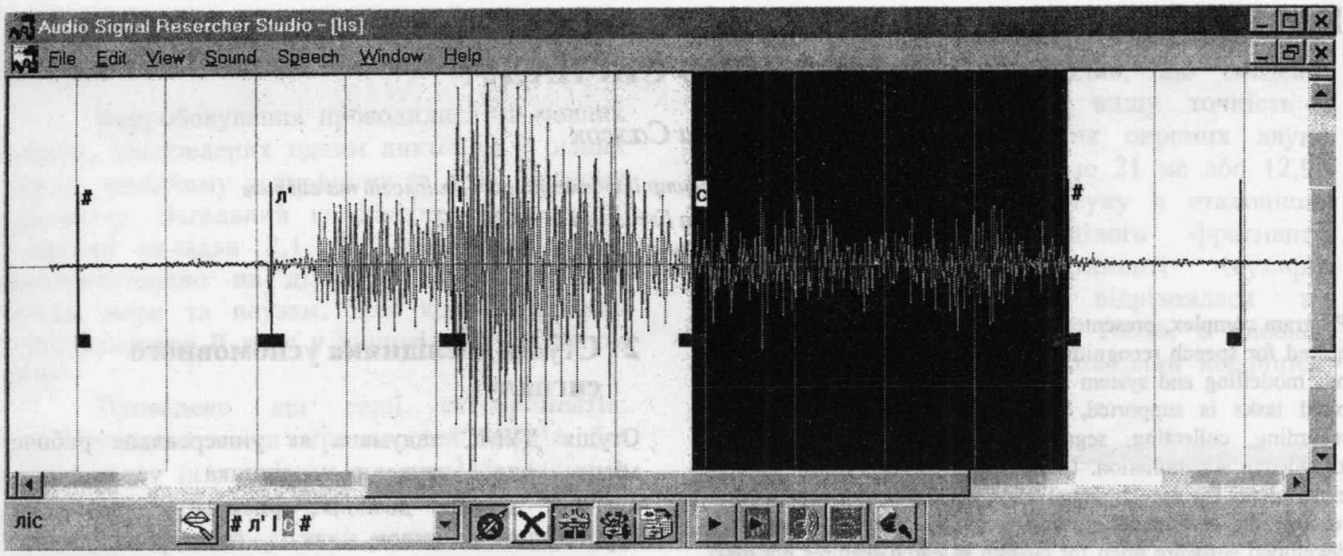

Рис. 1. Результат сегментації українського слова *ліс*. Кожний сегмент фонеми-трифона (ФТ) може бути виділений як за допомогою подвійного клацання мишкою на відповідній ділянці сигналу, так і шляхом вибору фонеми у вікні редагування (в даному випадку *с*). Потім експерт прослуховує виділене, коригує границі фонем, поповнює бібліотеку сегментів ФТ.

Експерт може вільно переміщати звукові файли до будь-якого іншого директорію, оскільки студія ДУМС за допомогою автоматизованої процедури поновлює зв'язки зі звуковими файлами.

Звукові сигнали можна редагувати, а саме: переставляти, вилучати, розтягувати в часі тощо.

Для обчислення границь одно-квазіперіодних сегментів (мікрофонем) використовується стійкий алгоритм, запропонований в [3]. Сегментація ж на ділянки, що відвідають окремим фонемам проводиться за аналізом одно-квазіперіодних сегментів, при цьому враховується інформацію про періодичність, тривалість, спосіб і місце творення фонеми.

Роль фонеми-трифона може виконувати не лише одна фонема в контексті з оточенням, а і послідовності фонем: склади, префікси, суфікси тощо.

Введена можливість одночасного візуального спостереження двох пар різних ділянок звукового сигналу, комбінуючи при цьому його різні представлення. Додатково забезпечується фільтрування сигналу, порівняння ділянок сигналу за різними мірами схожості, формантний аналіз, "зашумлення" сигналу тощо.

3 Студія дослідника природних мов

Студія ДПМ забезпечує введення, накопичення та оброблення знань про природні мови. Студія дозволяє також ставити експерименти в ділянці автоматичного розуміння мови та перекладу з однієї мови на іншу. Одна з компонент цієї студії відповідає за формування бази знань про конкретну природну мову, що фактично є словником з описом всіх словоформ, вживання слова в контесті з іншими тощо.

Кожній підтримуваній студією ДПМ мові відповідає динамічна бібліотека, що забезпечує функційні інтерфейси граматики та словника, роботу з текстами – всі операції, безпосередньо залежні від конкретної мови, а також взаємозв'язок з моделлю зовнішнього світу (МЗС). На даний момент лише для української мови існує версія такої динамічної бібліотеки.

За допомогою студії ДПМ експерт формує словник природної мови, зіставляючи слова зі складовими МЗС. Розроблено автоматизовану процедуру для знаходження в заданому тексті невідомих слів та додавання їх до словника в діалоговому режимі з користувачем.

МЗС породжує всі можливі канонічні форми заданої предметної області, за якими генеруються тексти вже для конкретної природної мови [1, 2].

4 Підсумок

Архітектура створеного програмного забезпечення дозволяє його широке використання в дослідженнях, тестах. Розробляється можливість додавати користувацькі функції, що дозволить будь-якій науково-дослідницькій групі проводити дослідження та тестувати свої алгоритми, методи.

Література

1. Т.К. Винцюк. *Анализ, распознавание и смысловая интерпретация речевых сигналов.* — Киев: Наукова думка, 1987, 264 с.
2. T.K. Vintsiuk. *Two Approaches to Create a Dictation/Translation Machine.* - Pros. of the 2nd Intern. Workshop "Speech and Computer", Cluj-Napoca, 1997, pp 1-6.
3. Taras K. Vintsiuk. *Optimal Joint Procedure for Current Pitch Period Discrimination and Speech Signal Partition into Qausi-Periodic and Non-Periodic Segments.* — Proc. Of the First Workshop on Text, Speech, Dialogue (TSD'98), Brno, 1998, pp. 135—140