

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ
БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ФАКУЛЬТЕТ РАДИОФИЗИКИ И КОМПЬЮТЕРНЫХ ТЕХНОЛОГИЙ
Кафедра информатики и компьютерных систем

Аннотация к дипломной работе
**«Разработка программного модуля для обработки и протоколирования
аудиоинформации»**

Тихонович Владислав Владимирович

Научный руководитель — к.т.н. Семенович С. Н.

Минск, 2025

РЕФЕРАТ

Дипломная работа 63 стр., 13 рис., 11 табл., 18 ист., 1 прил.

АУДИОИНФОРМАЦИЯ, БОРТОВОЙ РЕГИСТРАТОР, WHISPER.CPP, ПРОТОКОЛИРОВАНИЕ, РАСПОЗНАВАНИЕ РЕЧИ, ЦИФРОВАЯ ФИЛЬТРАЦИЯ, НЕЙРОСЕТЕВЫЕ МОДЕЛИ

Цель работы — разработка и реализация программного модуля, осуществляющего обработку, распознавание и протоколирование аудиоинформации, полученной с бортового регистратора CVR-120.

В работе рассмотрены особенности устройства CVR-120 и методика подключения к нему аппаратного модуля ССПА-2406. Разработан программный модуль, реализующий параллельную запись трёх аудиоканалов и их последующую конвертацию. Проведено исследование по выбору наиболее подходящей для проекта архитектуры распознавания речи. Для распознавания речи использована модель Whisper.cpp, предварительно протестированная на синтетических и реальных аудиоданных. Проведено сравнение качества распознавания различных вариантов моделей и реализаций.

Особое внимание уделено методам предварительной обработки аудиосигналов: медианному и полосовому фильтрам, спектральному вычитанию и вейвлет-фильтрации. Оценена их эффективность в улучшении точности распознавания. Также исследованы нейросетевые архитектуры (Spleeter, Open-Unmix, Demucs) для выделения речевых компонентов из шумового фона.

В заключительной части выполнена адаптация модуля для протоколирования под встраиваемые платформы (Raspberry Pi и др.), включая методы дистилляции и квантования моделей, что позволило снизить ресурсоёмкость и увеличить производительность без существенной потери точности.

Разработанный программный модуль протестирован и передан заказчику. Результаты работы могут быть использованы для повышения эффективности протоколирования аудиоинформации в авиационной и смежных отраслях.

РЭФЕРАТ

Дыпломная работа 63 стар., 13 мал., 11 табл., 18 крын., 1 дад.

АЎДЫЯІНФАРМАЦЫЯ, БАРТАВЫ РЭГІСТРАТАР, WHISPER.CPP, ПРАТАКАЛЯВАННЕ, РАСПАЗНАВАННЕ МАЎЛЕННЯ, ЛІЧБАВАЯ ФІЛЬТРАЦЫЯ, НЕЙРАСЕТКАВЫЯ МАДЭЛІ

Мэта работы — распрацоўка і рэалізацыя праграмнага модуля для апрацоўкі, распазнавання і пратакалявання аўдыяінфармацыі з бартавога рэгістратара CVR-120.

У працы разгледжаны асаблівасці ўстройства CVR-120 і метады падключэння да яго апаратнага модуля ССПА-2406. Распрацаваны праграмны модуль, які рэалізуе паралельную запіс трох аўдыяканалаў і іх наступную канвертацыю. Праведзена даследаванне па выбары найбольш прыдатнай для праекта архітэктury распазнавання маўлення. Для распазнавання маўлення выкарыстана мадэль Whisper.cpp, папярэдне пратэставаная на сінтэтычных і рэальных аўдыяданых. Праведзена параўнанне якасці распазнавання розных варыянтаў мадэляў і рэалізацый.

Асаблівая ўвага нададзена метадам папярэдняй апрацоўкі аўдыясігналаў: медыяннаму і паласавому фільтрам, спектральнаму адніманню і вейвлет-фільтрацыі. Ацэнена іх эфектыўнасць у паляпшэнні дакладнасці распазнавання. Таксама даследаваны нейрасецевыя архітэктury (Spleeter, OpenUnmix, Demucs) для вылучэння маўленчых кампанентаў з шумавога фону.

У заключнай частцы выканана адаптацыя модуля для пратакаліравання пад убудаваныя платформы (Raspberry Pi і інш.), уключаючы метады дыстыляцыі і квантавання мадэляў, што дазволіла знізіць рэсурсаёмістасць і павялічыць прадукцыйнасць без істотнай страты дакладнасці.

Распрацаваны праграмны модуль пратэставаны і перададзены заказчыку. Вынікі працы могуць быць выкарыстаны для павышэння эфектыўнасці пратакаліравання аўдыяінфармацыі ў авіяцыйнай і звязаных галінах

ABSTRACT

Thesis 63 pages, 13 figures, 11 tables, 18 references, 1 appendix.

AUDIO INFORMATION, FLIGHT RECORDER, WHISPER.CPP,
TRANSCRIPTION, SPEECH RECOGNITION, DIGITAL FILTERING,
NEURAL NETWORK MODELS

The goal of the work is to develop and implement a software module for processing, recognizing, and transcribing audio data from the CVR-120 flight recorder.

In the work, the features of the CVR-120 device and the method of connecting the SSPA-2406 hardware module to it are considered. A software module has been developed that implements parallel recording of three audio channels and their subsequent conversion. A study was conducted to select the most suitable speech recognition architecture for the project. The Whisper.cpp model was used for speech recognition, which was preliminarily tested on syn-thetic and real audio data. A comparison of the recognition quality of different model variants and implementations was carried out.

Special attention was paid to audio signal preprocessing methods: median and bandpass filtering, spectral subtraction, and wavelet filtering. Their effectiveness in improving recognition accuracy was evaluated. Neural network architectures (Spleeter, OpenUnmix, Demucs) were also investigated for extracting speech components from noisy backgrounds.

In the final part, the module was adapted for logging on embedded plat-forms (Raspberry Pi, etc.), including model distillation and quantization methods, which reduced resource consumption and increased performance without significant loss of accuracy.

The developed software module was tested and delivered to the customer. The results of the work can be used to improve the efficiency of audio logging in the aviation and related industries.