

Автоматизація комп'ютерного документування мовної голосової інформації на основі обробки, сегментації і аналізу сигналів.

Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 01.05.02 – математичне моделювання та обчислювальні методи.

Здобувач – аспірант Інституту кібернетики імені В.М. Глушкова Національної академії наук України Загваздін Олександр Сергійович.

Науковий керівник – доктор фізико-математичних наук, професор кафедри моделювання складних систем факультету кібернетики Київського національного університету імені Тараса Шевченка Крак Юрій Васильович.

Анотація

Дисертацію присвячено створенню засобів автоматизації документування мовних голосових фонограм з використанням математичних моделей представлення мовних звукових сигналів і методів їх сегментації і цифрової обробки.

У роботі запропоновано метод адаптивного визначення пауз і голосової активності у мовному голосовому сигналі з адаптивним порогом шуму, який дозволяє з високою надійністю визначати паузи у голосових сигналах з високим рівнем стороннього шуму і нестационарним шумом. Запропоновано метод визначення позиції зміни диктора у мовному голосовому сигналі на основі байєсівського інформаційного критерію порівняння моделей диктора в околі паузи у голосовому сигналі, і запропоновано методи сегментації сигналу за ознакою зміни диктора. Покращено методи фільтрації мовного сигналу від сторонніх шумів з використанням інформації, отриманої про шум з ділянок без голосової активності, визначених за допомогою адаптивного методу визначення пауз. Покращено методи зміни швидкості відтворення голосового сигналу з використанням методики PSOLA без змін до акустичних характеристик сигналу.

На основі запропонованих моделей і методів створено інформаційну систему для автоматизованого розподіленого комп'ютерного документування мовних голосових фонограм, яка показала високу ефективність у підвищенні продуктивності роботи груп операторів-стенографістів.