

РЕАЛИЗАЦИЯ ОСНОВНЫХ АЛГОРИТМОВ ОБРАБОТКИ РЕЧЕВОЙ ИНФОРМАЦИИ НА КОММЕРЧЕСКИХ СИГНАЛЬНЫХ ПРОЦЕССОРАХ МАЛОЙ ВЫЧИСЛИТЕЛЬНОЙ МОЩНОСТИ

Ю. В. Рождественский

Доклад посвящен изложению результатов исследования алгоритмической базы речевых технологий применительно к сигнальным процессорам малой вычислительной мощности.

Основные цели этой работы:

-определить круг задач в области речевых технологий, имеющих наиболее широкое практическое применение; определить набор основных алгоритмов обработки речи и их функциональный состав;

-оценить возможности распараллеливания вычислительных процедур, базирующиеся как на принципах физиологического восприятия звука, так и на особенностях математической реализации алгоритмов обработки речевой информации;

-исследовать возможности реализации этих алгоритмов на микроконтроллерах и сигнальных процессорах малой вычислительной мощности, ориентируясь на отечественные разработки специальных процессоров новых архитектур и высокой надежности с технологией параллельных вычислений.

В качестве базового был выбран набор задач речевой обработки в составе:

- сжатие голосовой информации при передаче по цифровым каналам;
- эхоподавление;
- распознавание голосовых команд;
- идентификация личности по голосу;
- синтез речевых сообщений.

Установлено, что естественная параллельность алгоритмов речевой обработки определяется физиологией человеческого слуха и рассматриваемым рабочим частотным диапазоном. Для большинства информационных каналов она равна 16.

В рассмотренных задачах речевой обработки 70-80 % алгоритмов и отдельных функций допускают 16-кратное распараллеливание вычислительных ресурсов.

Выбранные задачи были реализованы и отлажены в системе команд 16-битного сигнального процессора малой производительности и 8-битного микроконтроллера. Результаты представлены в табл. 1. Некоторое снижение качественных характеристик голосовой обработки, выполненной на микроконтроллере, для многих приложений несущественно. Вы

делен и сформулирован набор конкретных вычислительных задач для реализации на процессоре рекуррентной параллельной архитектуры.

Таблица 1.

Задача речевой обработки	Полученные характеристики	Потребные вычислительные ресурсы					
		Сигнальный процессор			Микроконтролл 8 бит		
		Производительность MOPS	RAM Кбайт	ROM Кбайт	Производительность MOPS	RAM Кбайт	ROM Кбайт
1. Сжатие голосовой информации MELP-4.8	Качество MOS=3.8	18					
2. Эхоподавление G.168	120 мс 32 мс	20 6	5.2	8			
3. Распознавание голосовых команд на 100 слов	Точность: без шумов: 98%	12	3		10		
	96% в шумах SNR = 15дб: 95%	12	3		10		
4. Идентификация личности по голосу: 64 диктора, 10 фраз по 4-7 слов	Точность: 1% 2%	10		26	9.4	1.5	16
5. Синтезатор речи на 200 слов	Качество: MOS=3.86 MOS=3.74	5.6	8		9.9	1.4	14.5