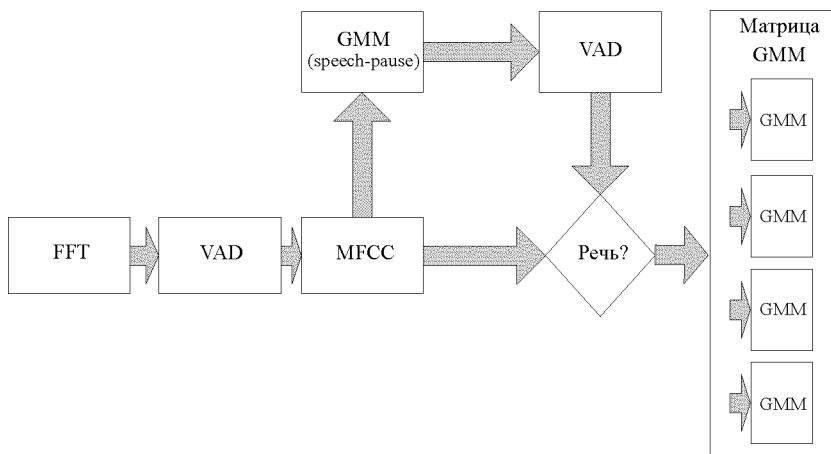


А. В. Максимов, Н. М. Чавчавадзе, М. А. Котов, С. Ю. Мельников, М. В. Федюкин (Москва, ТВП, «СТЭЛ КС»).
Способ реализации алгоритма идентификации дикторов с использованием ПЛИС.

Точность систем автоматической обработки речи, таких, как системы распознавания речи, системы идентификации дикторов и языков за последние годы возросла в несколько раз [1], [2]. Этот прогресс связан со значительным усложнением алгоритмов обработки речевых сигналов. Используемые алгоритмы характеризуются значительной вычислительной сложностью. Время их работы может в несколько раз превосходить длительность обрабатываемого речевого фрагмента. Одним из вариантов ускорения работы систем идентификации является их аппаратная реализация.

Рассмотрим аппаратную реализацию системы идентификации дикторов, основанной на использовании гауссовых смесей [3]. Ядром системы является плата-ускоритель с интерфейсом PCI-Express, выполненная на ПЛИС семейства Xilinx SXT. Архитектура системы приведена на рис.



Основные блоки системы: — VAD (Voice Activity Detector) — блок, отвечающий за выделение в поступающем потоке собственно речи;

— FFT и MFCC — блоки вычисления спектральных характеристик сигнала;

— GMM (speech-pause) — блок, выделяющий паузы в речевом сигнале;

— Матрица GMM — блок вычисления правдоподобия полученных данных с моделью диктора.

Плата позволяет обрабатывать одновременно три потока звуковых данных в формате 8кГц, 16 бит. Поддерживается до 200 моделей дикторов. Обеспечивается скорость обработки входного потока до 54 Мбайт/сек. Рекомендуемый объем данных для обучения диктора — от 30 секунд речи, рекомендуемая длительность тестового фрагмента — от 10 секунд речи.

СПИСОК ЛИТЕРАТУРЫ

1. Максимов А. В., Мельников С. Ю., Чавчавадзе Н. М., Федюкин М. В. Развитие систем автоматической текстонезависимой идентификации дикторов. — *Обзорные прикл. и промышл. матем.*, 2009, т. 16, в. 6, с. 1092–1094.
2. Максимов А. В., Мельников С. Ю., Чавчавадзе Н. М. Тенденции развития методов автоматической идентификации языка речевых и текстовых сообщений. — *Обзорные прикл. и промышл. матем.*, 2009, т. 16, в. 2, с. 365–367.
3. Kajarekar S. S., Scheffer N., Graciarana M., Shriberg E., Stolcke A., Ferrer L., Bocklet T. The SRI NIST 2008 speaker recognition evaluation system. — In: *Proceedings*

of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Taipei: 2009, p. 4205–4208.