

ОПТИМИЗАЦИЯ ВРЕМЕНИ ВЫЧИСЛЕНИЯ ПРИЗНАКОВ В РЕЧЕВЫХ СИСТЕМАХ НА КЛАСТЕРЕ IBM BLADE HS21 ИКИТ СФУ

М.С. Медведев, Г.С. Кирякова

В системах распознавания слитной речи время распознавания является главной характеристикой, определяющей качество работы системы. В настоящее время значительные усилия затрачиваются на разработку методов и алгоритмов позволяющих сократить время распознавания.

Система «Речь- текст» [1] включает БД фонем русского языка. Вычисление признаков фонем выполнено с использованием кратномасштабного вейвлет преобразования db8[2]. Эксперименты по распознаванию фонем и слов выполнялись на машине с процессором AMD Athlon 1800+, 512 Мб. Для формирования обучающей выборки использовались записи из речевой базы данных, в которой хранятся записи слов и текстов для каждого диктора (по 20 примеров - около 100 векторов признаков для каждой фонемы). В системе использовался словарь с грамматическими формами слов. По результатам экспериментов определялось время преобразования устного слова в его текстовое представление. При объеме словаря в 1210 слов время распознавания составило 288 мс. В таблице 1 приведены результаты распознавания слов в системе «Речь-текст», с использованием модели фонемы на основе вейвлет-преобразования и в системе SIRIUS [3], в которой вычисляются кепстральные характеристик для вычисления вектора признаков фонем.

Система	Модель фонемы	Среднее время распознавания слова
SIRIUS	Преобразование Фурье (кепстральные характеристики)	1160 мс
		670 мс
«Речь-текст»	Вейвлет-преобразование (базис Добеши 8)	288 мс

Таблица 1 – Сравнительный анализ времени распознавания слов

Средний коэффициент распознавания слов разработанной системой «Речь-текст» составил 90,0% [2]. Среднее время распознавания слова составило 288 мс, что удовлетворяет среднему темпу речи: одно слово за 500 мс [4]. Однако при увеличении размера словаря и процедуры формирования грамматической формы слова, процесс преобразования речи в текст требует больших временных затрат. Оценка времени T_k , которое затрачивается на предварительную обработку речевого сообщения и вычисление вектора признаков, составляет около 70% от общего времени принятия решений. Один из вариантов уменьшения T_k основан на том факте, что коэффициенты вектора признаков можно вычислять параллельно.

Для повышения эффективности обработки речевого сигнала, связанной с уменьшением временных затрат, анализировались возможности распараллеливания вычислительного процесса, используя ресурс, имеющийся в вычислительном центре ИКИТ СФУ - кластер IBM Blade HS21 (224 узла) производительностью 9287 TFlops/s и пакет rMatLab. Кластер ИКИТ СФУ работает под управлением ОС SUSE Linux Enterprise Server 10. Параллельные вычисления обеспечивают пакеты Distributed Computing Toolbox (DCT) и Computing Engine (MDCE) (Рис.1).

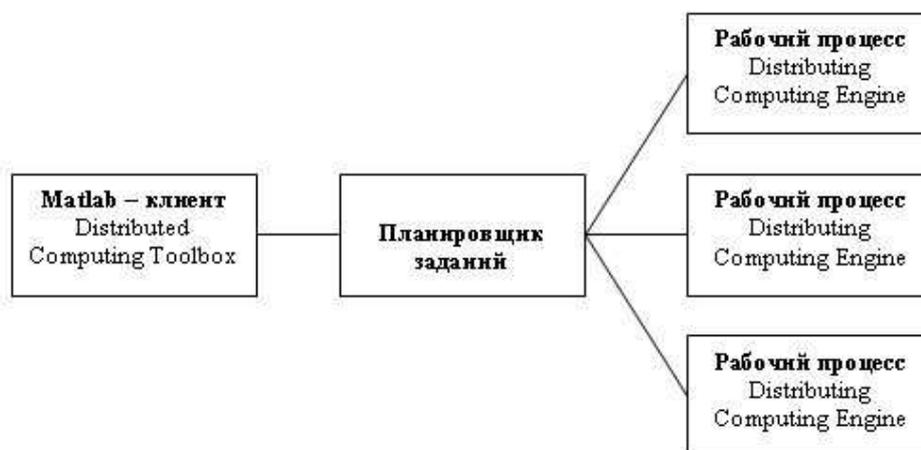


Рис. 1. Механизм распределенных вычислений Matlab

Для работы на кластере использовались клиент для удаленной работы на Unix-машинах Putty и порт сервера X Window System для операционной системы Microsoft Windows (XP/2003/Vista) и Xming, удаленный доступ через VPN. Оценивались время запуска, время решения задачи и время жизни стратегии с учетом динамики состояния среды.

Инструменты Distributed Computing Toolbox и MATLAB Distributed Computing Engine предназначены для реализации кластерных вычислений в среде MATLAB. С помощью первого из них пользовательская программа, написанная на языке MATLAB, разделяется на несколько независимых задач. MATLAB Distributed Computing Engine занимается выполнением этих задач в автономных сессиях MATLAB, запущенных на кластере.

Исходя из полученных данных, исследовались возможности связанные с

- разработкой методов обработки речевого сигнала в распределенных и параллельных средах.
- проведением вычислительных экспериментов в области обработки речевого сигнала,
- демонстрацией предлагаемых методов на задачах распознавания речи и адаптации технологии запуска программного обеспечения в области обработки речи через интерфейсы различного уровня с использованием параллельных вычислений для системы “Преобразование речи в текст” и “Скрытые Марковские модели в системах распознавание речи”.

Распараллеливание вычислительного процесса в алгоритмах системы «Речь-текст» позволяет сократить время вычислений для целого комплекса задач:

1. Преобразование в цифровую форму и предварительная обработка речевого сигнала, которое включает
 - выделение полезной информации из речевого потока;
 - подавление шума;
 - нормализация сигнала.
2. Вычисление вейвлет-коэффициентов вектора признаков фонем речевых единиц на этапах формирования обучающей выборки и распознавания;
3. Обучение нейронной сети и ее использование для классификации фонем;
4. Построение грамматической формы распознаваемых образов слов.

Результаты экспериментальных исследований с использованием параллельных вычислений для формирования вектора признаков речевого сигнала, используемого на этапе обучения и распознавания для системы «Речь-текст» приведены в таблице 2. Количество процессов N указывалось в matlabpool N. Параллелизм вычислений реализован с использованием механизма parfor.

Количество используемых процессов (ядер)	Среднее время, с
1	3,54
2	2,17
3	1,52
4	1,32
5	1,26
6	1,22
7	1,19
8	1,15

Таблица 2 – Зависимость времени решения задачи от числа заданных процессов

Результаты экспериментов показали сокращение временных затрат на вычисление признаков речевого сигнала при увеличении количества задействованных параллельных процессов. Минимальное время распознавания фразы из 8-ми слов составило 1150 мс, что удовлетворяет среднему темпу речи и составляет 143,75 мс

Ниже приведен фрагмент кода программы:

```
function [P] = wav_paral (Y)
matlabpool 8;
cca=0;
xds=Y'; % речевой сигнал
Fram=800; % длина кадра речевого
сигнала
Nfr=1;
Pt=[]; NRG=[];
```

```

parfor ff=1:(fix(length(xds)/Fram)-1)

    Nxds=premnmx(xds(Nfr*ff*Fram:Nfr*ff*Fram+Fram));
    [c,l]=wavedec(Nxds,10,'db8');           % вейвлет-разложение
        DCELL=detcoef(c,l,'cells');       % вычисление
коэффициентов детализации
        w_pars(DCELL);                     % формирование вектора
признаков

        Pt(:,ff)=[NRG]';
    end
    dd=length(Pt);
    cca=cca+1;
    if cca>1
        P=[P,Pt];
    else
        P=Pt;
    end;
end;

```

ЛИТЕРАТУРА:

1. Медведев М.С. Интеллектуальный комплекс для исследования речевого сигнала. / Кирякова Г.С., Ли А.С. // Информационные технологии в науке, образовании, телекоммуникации и бизнесе '09: материалы XXXVI Международн. конференц. молодых ученых. – Ялта-Гурзуф, 2009.
2. Медведев М.С. Распознавание речевых образов с использованием вейвлет-преобразования. / Модели и методы обработки изображений ММОИ-2007: Материалы Всерос. научн. конференц. - Красноярск: СФУ, 2007. – С. 53-59.
3. Ронжин, А.Л. Система распознавания русской речи SIRIUS/А.Л. Ронжин, А.А. Карпов, И.В. Ли; С.-Петерб. ин-т информатики и автоматизации РАН. – СПб., 2005. – 112 с.
4. Ракуш, В. В. Алгоритмы распознавания изолированных слов на основе скрытых Марковских моделей/В.В. Ракуш, А.С. Рылов //Вопросы криминологии, криминалистики и судебной экспертизы: сб. науч.тр.- Минск, 1998. - Вып.13.- С. 234-238.