

OSCAR上での音声認識の並列処理手法

5P-6

飯田 晴彦、笠原 博徳、成田 誠之助

早稲田大学理工学部電気工学科

1. はじめに

音声認識における特徴抽出の有効な一つの手段として、音声波形から調音の状態を推定する方法があり、これを大別すると、理想化された声道モデルを用いて音声波の音響的特徴量に等価な音響管を推定するものと、調音モデルを設定してその調音パラメータを推定するものがある^[1]。

そのうち、調音モデルに基づく推定法では、調音器官の構造によって拘束された声道形の枠組みの中で、モデルと音声波の間の音響的な適合基準を最適にする調音パラメータ、あるいは声道形が推定され、声道の損失などを含めたより正確な音声生成モデルを導入することができる^[1]。そして、この調音パラメータの推定では、重みつき2乗誤差の和の評価関数の下での非線形最適化問題となるために^[1]、評価関数の計算、連立一次方程式の求解、ヤコビ行列の計算などの繰り返し計算が必要となる。従って、計算機上ではこの反復計算部分の計算量が大きくなり、この部分を高速化することは特徴量抽出の高速化のために必要となる。特に、反復計算部分の中でも、ケプストラム係数という音響量はヤコビ行列の計算の中などで数多く計算されるので、この高速化は重要である。

本稿では、この反復計算部分において最も計算量の多いケプストラム係数計算の高速化のために並列処理技術を適用することを考え、マルチプロセッサ上で効率良い並列処理を行うための手法を提案する。

2. 調音パラメータ推定^[2]

調音パラメータからケプストラム係数を求めるには、音響管モデルの各段の断面積の計算に始まり、線形予測係数(α パラメータ)への変換を経て、ケプストラム係数への変換という手順を踏む。

音響管モデルの各段の断面積は、調音パラメータから線形調音モデルを構成し、この調音モデルの各段の体積より求める。線形調音モデルは円錐台を連ねた形状をしており、音響管モデルは高さが同一の円柱を連ねた形状をしている。調音モデルにおいて、音響管モデルの各段に対応する段の体積を補間により計算し、その結果を音響管モデルの1段の長さ(円柱の高さ)で割ることによって、音響管モデルの各段の断面積を求める。各段の断面積を s_i 、音響管モデルの1段の長さを Δ 、補正係数を κ 、着目する周波数 f_0 をとすれば、反射係数 μ_i と損失係数 ξ_i は以下のように求められる。

$$\mu_i = \frac{s_{i+1} - s_i}{s_{i+1} + s_i} \quad (1)$$

$$\xi_i = \exp(-2\xi_i\Delta) \quad (2)$$

$$\xi_i = 0.0001\kappa f_0/s_i$$

または

$$\xi_i = \exp(L \cdot \text{sqrt}(f_0/s_{i+1})) \quad (3)$$

$$L = -0.000103\kappa\Delta$$

線形予測係数 $\alpha_j^{(i)}$ は以下のように求められる。

$$\alpha_j^{(i)} = \alpha_j^{(i-1)} - \mu_i \xi_i \beta_{j-1}^{(i-1)} \quad (4)$$

$$\beta_j^{(i)} = \xi_i \beta_{j-1}^{(i-1)} - \mu_i \alpha_j^{(i-1)}$$

$$\alpha_0^{(i)} = 1, \beta_0^{(i)} = -\mu_i$$

$$i = 1, \dots, M$$

$\alpha_i = \alpha_i^{(M)}$ とすると、ケプストラム c_i は以下のように求められる。

$$c_i = -\alpha_i + \sum_{j=1}^{i-1} \frac{1-j}{1} c_{j-1} \alpha_j \quad (5)$$

3. 並列処理手法

ここで提案する調音パラメータ推定の並列処理手法は、タスク生成、タスクグラフ生成、タスク・スケジューリング、マシンコード生成の4つの部分からなる。

タスク生成では、処理すべき計算全体をタスクに分割する。ここで問題となるのがタスク・グラニュラリティ(タスクの大きさ)の決定である。グラニュラリティを細かくし過ぎると、タスク間同期やデータ転送などのオーバーヘッドが増大する。また、グラニュラリティが大きすぎると、並列性が引き出されない。調音パラメータの推定の場合、複数インスタレーションレベルのグラニュラリティを採用している^{[3], [4]}。図1にタスク生成の例を示す。このタスクの表記法は、当研究室でタスクを陽に記述するために導入した構造化中間言語である。図1において、反射係数の計算タスクと損失係数の計算タスクは、それぞれ(1)式と(3)式をもとに生成したものである。①~③が各タスクを表わしており、これらのタスクのグラニュラリティは反射係数の計算では代入文レベルであり、損失係数の計算では小ブロックレベルである。

生成したタスク間のフロー依存や出力依存といったデータ依存関係をタスクグラフという無サイクル有向グラフで表す^{[3], [4]}。ケプストラム係数計算においては、ノード数約1100、アーク数約3900のタスクグラフとなる。

A Parallel Processing Scheme for Speech Recognition on OSCAR

Haruhiko IIDA, Hironori KASAHARA, Seinosuke NARITA

Waseda University

```

【反射係数の計算】
① 
$$\begin{cases} - & v_{15} v_{16} & ; \mu_1 = (s_1 - s_2) / (s_1 + s_2) \\ + & v_{15} v_{16} \\ / & t - 2 t - 1 \\ := & v_{33} t - 1 \end{cases}$$

② 
$$\begin{cases} - & v_{16} v_{17} & ; \mu_2 = (s_2 - s_3) / (s_2 + s_3) \\ + & v_{16} v_{17} \\ / & t - 2 t - 1 \\ := & v_{34} t - 1 \end{cases}$$

③ 
$$\begin{cases} - & v_{17} v_{18} & ; \mu_3 = (s_3 - s_4) / (s_3 + s_4) \\ + & v_{17} v_{18} \\ / & t - 2 t - 1 \\ := & v_{35} t - 1 \end{cases}$$

:
:

【損失係数の計算】
① 
$$\begin{cases} \text{tif} & & ; \text{if}(s_1 = s_2) \xi_1 = 1.0 \\ = & v_{15} v_{16} \\ \text{then} & & \\ := & v_{48} c \ 2 \\ \text{eot} & & \\ \text{else} & & ; \text{else } \xi_1 = \exp(L * \text{sqrt}(f_0 / s_2)) \\ / & v_{32} v_{16} \\ +\text{sqrt} & t - 1 \\ * & v_{47} t - 1 \\ +\text{exp} & t - 1 \\ := & v_{48} t - 1 \\ \text{eoe} & & \end{cases}$$

② 
$$\begin{cases} \text{tif} & & ; \text{if}(s_2 = s_3) \xi_2 = 1.0 \\ = & v_{16} v_{17} \\ \text{then} & & \\ := & v_{49} c \ 2 \\ \text{eot} & & \\ \text{else} & & ; \text{else } \xi_2 = \exp(L * \text{sqrt}(f_0 / s_3)) \\ / & v_{32} v_{17} \\ +\text{sqrt} & t - 1 \\ * & v_{47} t - 1 \\ +\text{exp} & t - 1 \\ := & v_{49} t - 1 \\ \text{eoe} & & \end{cases}$$

:
:
    
```

図1. タスク生成の例

この様にして求めたタスク間の先行制約と各タスクの実行時間を考慮して、各プロセッサへのタスク割当てと一つのプロセッサに割り当てられたタスク間の実行順序とを最適にしていく。これがスケジューリングであり、この問題は強NP困難問題として知られているが、当研究室で開発したヒューリスティック・アルゴリズムCP/DT/MISF法を用いることにより、短時間で近似解を得ることができる。このアルゴリズムは、データ転送を考慮して処理時間の最小化を図っている^{[3]、[4]}。

以上のようにして求めた、各プロセッサに割り当てたタスク用のコードを生成し、同期やデータ転送のコードを付加することにより、OSCAR^[5]上の各プロセッサ用のマシン・コードが生成される。このとき、タスクの割当てや、実行順序に注目することにより、不必要なタスク間同期コードを除去し、同期のオーバーヘッドを大幅に減少させることができる。また、レジスタ利用の最適化もここで行う^{[3]、[4]}。

4. 手法の有効性評価

図2にケプストラム係数計算をOSCAR上で行う際の並列処理効果のシミュレーション結果を示す。この評価に用いたシミュレータではデータ転送を考慮しており、

高精度のシミュレーション結果を得ることができる。図2からプロセッサ数が3台のとき約2.3倍、9台で約5.3倍の高速化が可能であることが分かる。先にも述べたように、ケプストラム係数計算は、調音パラメータ推定においてはヤコビ行列の計算などで数多く行われるので、この手法の有効性が確かめられたといえる。

5. まとめ

本稿ではケプストラム係数計算の並列処理手法を提案した。今後、OSCAR上でのインプリメントをしていく予定である。

謝辞

本研究の遂行に関して貴重なアドバイスを頂きました、早稲田大学の白井克彦教授、法政大学の小林哲則助教授に深く感謝いたします。

参考文献

- [1] 白井、菅田、”音声波からの調音パラメータの推定”、電子通信学会論文誌、61-A、5、p.409(昭53-5)
- [2] 菅田、”調音モデルに基づく音声の特徴抽出に関する研究”、早稲田大学博士論文、1977
- [3] H.Kasahara and S.Narita, "Practical multiprocessor scheduling algorithms for efficient parallel processing", IEEE Trans.Comput., pp.1023-1029,Nov. 1984
- [4] H.Kasahara and S.Narita, "An approach to supercomputing using multiprocessor scheduling algorithms", in Proc. IEEE 1st Int. Conf. on Supercomputing, Dec. 1985
- [5] 笠原、成田、吉田、富沢、”汎用目的マルチプロセッサ・システムOSCARの概要”、昭62情報前全大、4Q-1

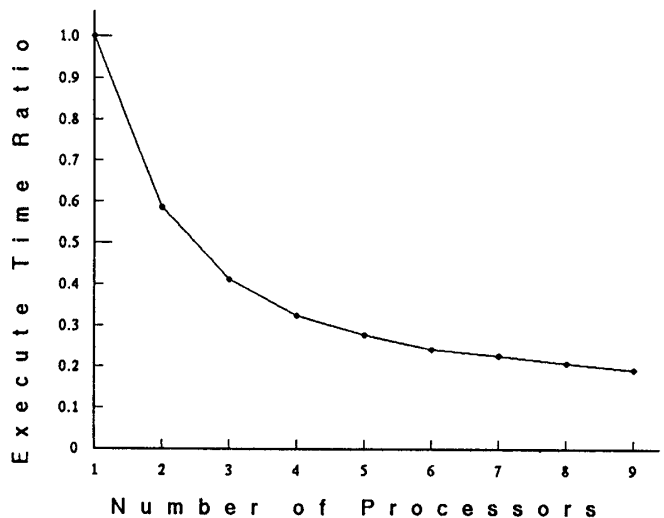


図2. シミュレーション結果