The Parallel Algorithm for Human Tumor Growth Using Distributed Multiprocessor Computer System

Dr Norma Alias
Mathematic Department, Faculty of Science,
University Technology Malaysia
Johor, Malaysia

Norfarizan Mohd Said, Siti Nur Hidayah Khalid,
Dolly Sin Tien Ching, Phang Tau Ing
Mathematic Department, Faculty of Science,
University Technology Malaysia
Johor, Malaysia

Abstract—This paper focused on the implementation of parallel algorithm for the simulation of human tumor growth. The research emphasised on the grand challenge application of breast and brain tumors growth. Based on the present knowledge of the properties of tumors, some mathematical models have been developed by researchers to quantify the proliferation and invasion dynamics of tumors within anatomically accurate heterogeneous human tissue. This paper focuses on the implementation of parallel algorithm for the simulation of hyperbolic equation for breast pressure and parabolic equation of brain tumors growth. The platform for high performance computing of the parallel algorithms run on a distributed parallel computer system. The numerical finite-difference method is focused to design the discreatization of these partial differential equations. The result of finite difference approximation (explicit, Crank-Nicolson and fully implicit methods) will be presented graphically. The implementation of parallel algorithm based on parallel computing system is used to visualize the growth of human tumour. Parallel Virtual Machine (PVM) is emphasized as communication platform in parallel computer systems. The software system functions to enable a collection of heterogeneous computers to be used as synchronize and flexible concurrent computational resource. The parallel performance measurement will be analyzed from the aspect of speedup, efficiency, effectiveness and temporal performance.

Keywords-human tumor; parallel algorithm

I. INTRODUCTION

The human body is made up of many types of cells. Each type of the cells has special functions. Most of the cells in the body grow and then divide in an orderly way to form new cells as they are need to keep the body healthy and working properly. The cells will divide too often and without any order when they lose the ability to control their growth. The extra cells from a mass of tissue are called tumours.

Brain cancer is the prevalent cancers in the world and they are leading causes of death from cancer. Mathematical modelling of biomedical phenomena (Murray, 2003) can be extremely helpful in analyzing factors that may contribute to the complexity intrinsic in insufficiently understood developmental process disease. Based on present knowledge of the properties of tumors, a mathematical model has been developed to quantify the proliferation and invasion dynamics of tumors within anatomically accurate heterogeneous human tissue. The implications of the model would be considerable interest to those interests in the study of other diseases for which medical imaging plays a part of the assessment of the disease (other cancer as well as developmental diseases).

Breast cancer seems to be predominant among Chinese women with an incidence of 25 per 100000 population and Malay women with an incident of 16 per 100000 populations (Malaysia Medical Association internet site—http://www.mma.org.my/current_topic/woman.htm, March 2006). Breast cancer happens when cells in the breast begin to grow out of control and can then invade nearby tissue are called tumors. However, some tumors are not really cancerous because they cannot spread or threaten someone's life. Medical ultrasound is essentially a means of producing visual images based on echoes that occur in acoustic interface. This echoes contains information that can be used t study the various breast tissues. A lesion is localized if its echo pattern is different from that of the surrounding medium. A

II. MATHEMATICAL MODEL

The mathematical model has been developed to detect the growth and the extension of theoretical tumor cells in a matrix that accurately describes the human's anatomy to a resolution of 1 cu mm. By defining a virtual human brain and breast with the anatomical distribution of grey and white matter (the two primary tissue components of the human) to a resolution of 1 cu mm, the differential motility of tumors cells in grey and white can be modelled.

For one dimensional brain cancer, the tumor cells migrate more rapidly in white than in grey matter (Giese and Westphal, 1996; Silbergeld and Chicoine, 1997), the motility coefficient dependence on the local tissue composition. The model can be written mathematically as:

$$\frac{\partial c}{\partial t} = \nabla \cdot \left(D(x) \nabla c \right) + \rho c \qquad (1)$$

where $c(x,t)$ is the concentration of tumor cells at any position **x** and time t, ρ is the units of per day and represents the net rate of growth of tumour cells, including proliferation, loss and death, D denotes the units of cm^2 per day and represents the diffusion coefficient of cells in human tissue, $D(x) = D_g$ (constant for x in grey matter) and $D(x) = D_w$ (constant for x in white matter).

As noted, the diffusion coefficient in white matter is larger then that in grey matter: $D_w > D_g$. ∇ represents the spatial gradient. The diffusion term describes the active migration of the glioma cells using a simple Fickian diffusion (Murray 1993) where cells move from regions of higher to lower densities. Tumour cells are assumed to grow exponentially.

The boundary condition simply requires that glioma cells are not allowed to migrate outside of the human tissue. Assume that the tumour has grown to about 4000 cells as a local mass before it begins to diffuse. We used the growth rate, $\rho \approx 0.012$ /day (Alvord and Shaw, 1991; Swanson, 1999; Swanson *et al.,* 2000) and diffusion coefficient, $D \approx 0.0013\ cm^2$ /day (Swanson, 1999; Swanson *et al*, 2000) in the model as suggested for high-grade gliomas [Swanson, Alvord and Murray, 2000].

For two dimensional brain cancer growth,
$$\frac{1}{c^2}\frac{\partial^2 p(x,y)}{\partial t^2} + \gamma \frac{\partial p(x,y)}{\partial t} = \nabla.(c\nabla p(x,y))$$

The entire fixed computational domain $0 \leq x \leq 1$ and $0 \leq y \leq 1$ are each discretized using equally spaced meshes; the interface is a mesh point, corresponding to $x = 1$ and $y = 1$. The domain occupied by the tumor is embedded into a larger fixed, time-independent, computational domain D that is discretized using a uniform Cartesian mesh with $\Delta x = \Delta y = h$. The simulations of the two dimensional (2D) discrete model were carried out on a 10 x 10 grid. This involves approximating the continuous domain $[0,1] \times [0,1]$ in the usual way as a grid of discrete points (mesh size h), and time (t) by discrete increments (magnitude k). In this research, the initial value of the computation time unit Δt is equivalent to 1.5 days (Chaplain, 2001). The increment of the time Δt that refers to the intervals in the finite difference scheme is adaptive.

The hyperbolic equation used in one dimensional breast cancer growth to govern ultrasound pressure fields in lossy attenuated medium consists of a dissipative wave equation incorporating a frequency-dependent attenuation (A. Bounaïm *at el*, 2003).
$$\frac{1}{c^2}\frac{\partial^2 p(x,t)}{\partial t^2} + \gamma \frac{\partial p(x,t)}{\partial t} = \nabla.(c\nabla p(x,t))$$

Where x and t are the space and time variables, respectively, and c is the sound speed of the traversed tissue. Studies of bioacoustic models have shown that in viscous fluids the absorption coefficient γ is characterized by a power law frequency.

The breast tissue is initialized with the condition.
$$p(x,t_0) = p_{atm} \text{ and } \frac{\partial p}{\partial t}(x,t_0) = 0$$

where p – pressure
x – space variable
t – time variable
c – sound speed of the traversed tissue
γ– damping or attenuation parameter

$$\gamma = \gamma(f) = 2\frac{\alpha_0}{c}f^y$$

α – dependent material constant
f – frequency
y – frequency-power exponent varying from 0 to 2 and depending on tissue.

III. PARALLEL ALGORITHM

A. Gauss Seidel Red Black Iterative Method

To implement the parallel algorithm in solving the finite difference equation, Gauss Seidel Red Black (GSRB) algorithm is used. The iterative method contains 2 sub domain, Ω^R and Ω^B. There are a communication between Ω^R and Ω^B. The calculation of this method is show by the equation (5) and (6),

   i.   Grid calculation at Ω^R :
$$u_i^{(n+1)} = \frac{1}{2}r\{u_{i-1}^{(n+1)} - 2u_i^{(n+1)} + u_{i+1}^{(n)}\} + b_i,$$
$$i = 1,3,5,...,m. \qquad (5)$$

   ii.   Grid calculation at Ω^B :
$$u_i^{(n+1)} = \frac{1}{2}r\{u_{i-1}^{(n+1)} - 2u_i^{(n+1)} + u_{i+1}^{(n)}\} + b_i,$$
$$i = 2,4,6,...,m-1 \qquad (6)$$

The GSRB iterative using parallel computing is convergence very fast compares to the parallel algorithm using Gauss Seidel (GS) iterative. GSRB is a data decomposition approach that divides arrays among local processors to minimize the communication. The data structure has to be decomposed where given set of ranges assigned to particular processors must be physically sent to those processors for processing to be done. The result must be sent back to whichever processors responsible for coordinating the final result.

The simple GS update strategy may be appropriate in a sequential program and often preferred over Jacobi strategies because they allow solutions of comparable accuracy to be obtained using less iteration. However, for the solution of large sparse matrix problem, the GSRB parallel algorithm is more suitable to be implemented compare to the GS.

B. The Visualization of human tumor growth using Crank-Nicolson

The numbers of the cancer cells are computed by using parallel algorithm with PVM. The tumour is assumed growing to about 4000 cells as a local mass before it begin to diffuse. The boundary condition simply requires that glioma cells are not allowed to migrate outside of the brain tissue.

The concentration gradient of brain tumour cells is represented by the curves in Figure 1 for every 5 days within 1 month. The graph has shown that in the first 5 days, the tumour cells only proliferate to form a small dense lesion. After 20 days, the tumour cells become highly diffusive. The tumour cells have diffused to more than 160,000 after 30 days.
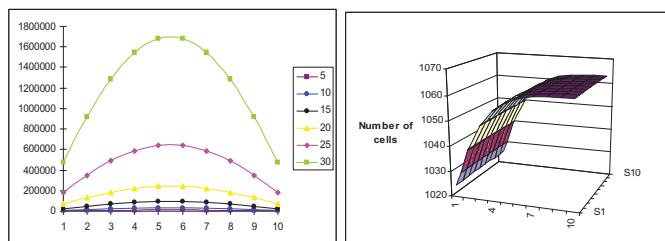

Figure 1. The growth rate of brain tumour 1D and 2D

Since that the glioma are diffuse tumour, only a small portion of actual tumour can be predicted. The length of the growth area that consist tumours cell is only 0.167cm within 30 days. From the graph of the growth of brain tumour, we also can conclude that the tumour growth slowly and quickly diffusing.

The results that is presented using PVM programming showing in Figure 1 is only a small area, about 2.78% of the tumour at death. For the detection of the growth of cancer cells in human brain using parallel computing, the area can be predicted broadly by using the different net proliferation rate, ρ and diffusion coefficient, D. The various area of the human brain has the different value for ρ and D.

C. The Visualization of breast tumor growth using Bounaïm

The pressure of the breast cancer computed using sequential algorithm with programming C of mathematical model in this research (A. Bounaïm *at el*, 2003). The pressure will be used to visualize the growth of the breast cancer within 30 days more obviously.
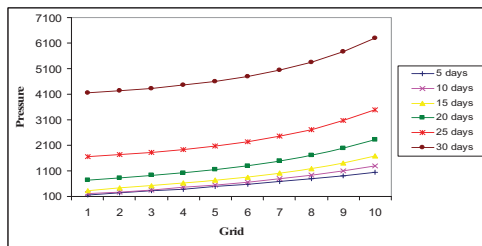

Figure 2. The growth rate of breast cancer

Figure 2 shows the pressure inside the breast for every 5 days within 1 month. The graph has shown that in the first 5 days, the pressure is low. However, after 20 days, the pressure increases drastically. The pressure becomes higher after 30 days. As a conclusion, the pressure is increasing with the tumor size. Therefore, the growth of breast cancer can be visualized from its pressure. The sound speed in the breast tissue and the tumor are extracted from clinical experiments together with the attenuation parameters (Weiwad et al, 2000), and are summarized in Table 4.2. The numerical scheme is then stable for a time step. $\Delta = \times - 26\ 10\ 9\ t\ s$

IV. PERFORMANCE ANALYSIS OF PVM

The parallel performance measurements such as time execution, speedup, efficiency, effectiveness and temporal performance are analyzed to proof parallel algorithms is significantly better than the sequential algorithms. The parallel implementation of an algorithm involves the division of total workload into a number of smaller tasks, which can be assigned to different processors and executed concurrently, which allows the large problems of iterative methods converge faster to the solution. The analysis of the executive time by parallel programming helps to understand the barriers of high performance computing and how improvement can be done by increasing the number of processors. As the cluster getting larger, it is become more important and efficient for solving the real-life problems.

In the PVM implementation of the modelling codes there is a master task and there are a large sparse problem number of worker tasks. The main job of master task is to divide the region domain into sub-domains and distribute them to worker tasks. The worker tasks perform execution computation and communication after each time step.

The execution time for parallel GSRB with PVM is about 10 times faster than sequential GS algorithm. This shows that parallel algorithm is significantly better than sequence algorithm. Besides that, parallel GSRB is more convergence compare to its sequential algorithm. The iteration for GSRB algorithm to be converging is only 64 while the GS with sequence algorithm perform 74 iterations for convergence.

The Amdahl's law state that the speed of a program is the time to execute the program while speedup is defined as the time it takes a program to execute in serial (with one processor) divided by the time it takes to execute in parallel (with many processors). The formula of speedup for a parallel application is given as

$$speedup(p) = \frac{Time(1)}{Time(p)},$$

where $Time(1)$ denotes execution time for a single processor and $Time(p)$ is the execution time using p parallel processors.
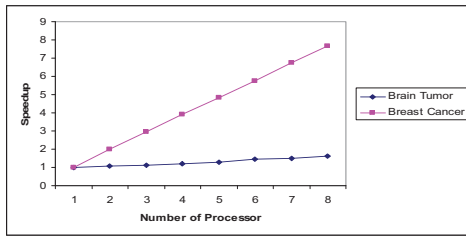
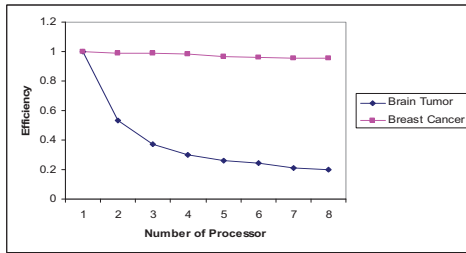Figure 2. The speedup vs. number of processors

Figure 3. Efficiency vs. number of processors

Figure 3 shows the speedup graph is the straight linear and as well as the number of processors increases, *p*. It is because the distributed memory hierarchy reduces the time consuming access to a cluster of workstations. According to Amdahl's Law, the speedup increases with the number of processors increase up to the certain level.

The efficiency of a parallel program is a measure of processor utilization. Figure 4 depicts the efficiency decreases with the increasing of number of processors, *p*. The decrease of efficiency contributed by the factors of poor load balance when imbalance workload distributed among the different processors. It is also contributed by the idle time, time startup and waiting time for all processors to complete the computations. Efficiency is the speedup divided by the number of processors used.

$$Efficiency = \frac{Speedup}{p} \quad , \; p = \text{number of processors.}$$

Effectiveness is used to calculate the speedup and the efficiency. The effectiveness is

$$Effectiveness = \frac{Speedup}{pTime(t)},$$

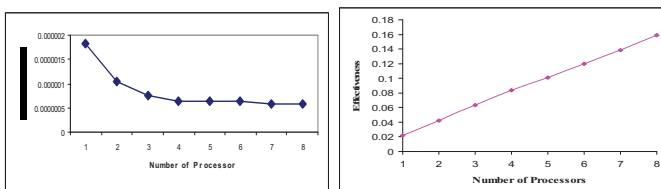where *p* is number of processors and *Time(t)* is execution time using *p* parallel processors.

Figure 5. Effectiveness vs. number of processors

Figure 5 shows that the effectiveness increase when the number of processors increasing. The formula of the effectiveness is depending on the speedup, when the speedup increases, the effectiveness will also increase. Furthermore, the using of the number of processors is added to make the graph increasing. The graph does not show the straight line because there is a communication between eight of the processors.
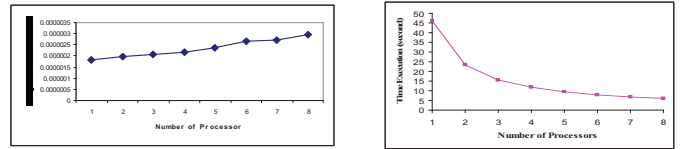

Figure 6.Temporal performance vs. number of processors

Temporal performance is a parameter to measure the performance of a parallel algorithm which is

$$Temporal = \frac{1}{Time(t)} .$$

Figure 6 shows that the temporal performance is increase with the increasing of the number of processors. The graph above has shows a straight line because of the execute time is decreasing extremely versus the number of processors.

CONCLUSION

A mathematical model using the one dimensional parabolic equation regarding the growth of human tumours has been presented. Three finite difference methods are studied and the Crank-Nicolson implicit method is shown to be the best method to solve the equation. In addition, this well known as highly stability and good accuracy method is then used to visualize the growth of human tumours for 30 days.

The results of the analysis for the performance measurements such as time execution, speed up, efficiency, effectiveness and temporal performance have proved that parallel algorithms is significantly better than the sequential algorithms especially in terms of the execution time. The Gauss Seidel Red Black is effective iterative method and found to be well suited for parallel implementation on the PVM where data decomposition is run synchronously and concurrently at every time level.

Finally, the development of mathematical model in visualizing the growth of human tumour can be easily extended into multiple dimensions parabolic equation. More number of processors could also be used in solving the mathematical model in order to improve the speed and performance on a distributed parallel computer systems

For all the problem statement, the parallel performance of finite difference of human tumor growths show that when the number of processor increase the speedup increased. The performance of the parallel computer had been analyzed using the graphs from the aspect of execution time, speedup, efficiency, effectiveness and temporal performance. From the analysis, parallel processing can be concluded as a better price

and performance tools in solving mathematical problem which can reduce time compare to sequence processing.

References

Al Geist, Adam Beguelin, Jack Dongarra, Weicheng Jiang, Robert Manchek, Vaidy Sunderam, *PVM: Parallel Virtual Machine, A User's Guide and Tutorial for Network Parallel Computing*. The MIT Press, Cambridge, Massachusetts, London, England.

C. Xavier, S.S. Iyengar. 1998. *Introduction to Parallel Algorithms.* John Wiley & Sons, Inc.

E.H.Twizell. 1984. *Computational Methods For Partial Differential Equations.* John Wiley & Sons, Inc.

G.D. Smith. 1985. *Numerical Solution of Partial Differential Equations*. Oxford University Press.

Hesham El-Rewini, Ted Lewis. 1997. *Distributed and Parallel Computing*. Manning Publications Co.

Kristin R. Swanson, Ellsworth C. Alvord, Jr, J. D. Murray. 2004. *Dynamics of A Model For Human tumours Reveals A Small Window For Therapeutic Intervention*. Volume 4, Number 1, pp. 289-295.

Kristin R. Swanson, Ellsworth C. Alvord, Jr, J. D. Murray. 2002. Virtual human tumours (gliomas) enhance the reality of medical imaging and highlight inadequacies of current therapy. *British journal of Cancer (2002) 86, the Cancer Research Campaign* : 14-18.

Kristin R. Swanson, Ellsworth C. Alvord, Jr, J. D. Murray. 2000. *A quantitative model for differential motility of gliomas in grey ad white matter*, *Cell Prolif*. 2000, 33, 317-329.

Kristin R. Swanson, Ellsworth C. Alvord, Jr, J. D. Murray, *Quantifying efficacy of chemotherapy of human tumours with homogeneous and heterogeneous drug delivery*.

Norma Alias. 2003. *Pembinaan dan Pelaksanaan Algoritma Selari bagi Kelas TTHS dan TTKS dalam Menyelesaikan Persamaan Parabolik pada Sistem Komputer Selari Ingatan Teragih*. Universiti Kebangsaan Malaysia: Tesis Ph.D

Norma Alias, Dr. Abdul Rahman Adullah, Dr. Mohd Salleh Sahimi Mohamed. *The Development and Parallel Performance of a Distributed Memory Machine Using PCs Based*.

Marc Snir, Steven Otto, Steven Huss-Lederman, David Walker, Jack Dongarra. 1998. "*MPI – The Complete Reference: Volume 1*, The MPI Core, second edition, Massachusetts Institute of Technology.

Michael Kofler. 1997. *Linux – Installation, Configuration, and Use*, Addison-Wesley.

Richard L. Burden, J. Douglas Faires. 1993. *Numerical Analysis*, Fifth Edition, PWS Publishing Company.

[1]