Acquisition and analysis of spectral image data by linear un-mixing, cluster computing and a novel spectral imager

P. R. Barber*, R. J. Edens, B. Vojnovic
Gray Cancer Institute, Mount Vernon Hospital, Northwood, Middlesex, HA6 2JR

SUMMARY

We describe how spectral imaging, linear un-mixing and cluster computing have been combined to aid biomedical researchers and allow the spatial segmentation and quantitative analysis of immunohistochemically stained tissue section images. A novel cost-effective spectral imager, with a bandwidth of 15 nm between 400 and 700 nm, allows us to record both spatial and spectral data from absorptive and fluorescent chemical probes. The linear un-mixing of this data separates the stain distributions revealing areas of co-localisation and extracts quantitative values of optical density. This has been achieved at the single-pixel level of an image by non-negative least squares fitting. This process can be computationally expensive but great processing speed increases have been achieved through the use of cluster computing. We describe how several personal computers, running Microsoft WindowsXP, can be used in parallel, linked by the MPI (Message Passing Interface) standard. We describe how the free MPICH libraries have been incorporated into our spectral imaging application under the C language and how this has been extended to support features of MPI2 via the commercial WMPI II libraries. A cluster of 8 processors, in 4 dual-Athlon-2600+ computers, offered a speed up of a factor of 5 compared to a singleton. This includes the time required to transfer the data throughout the cluster and reflects a processing efficiency of 0.62 (a Cluster Efficacy of 3.0). The cluster was based on a 1000Base-T Ethernet network and appears to be scalable efficiently beyond 8 processors.

Keywords: spectral imaging, immunohistochemistry, linear un-mixing, cluster computing.

1. INTRODUCTION

Spectral imaging generates a three-dimensional data set that represents the sample both spatially and spectrally. Traditional microscopy offers high spatial resolution by coupling a high-resolution imager, typically with a minimum of 500 x 500 pixels, to the output port of the microscope. The use of a colour imager provides some spectral information in the form of three broad, and partially overlapping, spectral bands covering the red, green and blue (RGB) regions. In contrast, spectral imaging can provide narrow bandwidth information at a high number of wavelengths spanning the visible spectrum. In the field of microscopy, the additional information available though this technique is not only a valuable tool for visualising previously invisible structures but is also of great advantage when the goals are automation and quantitative measurement.

Spectral Imaging can aid the histo-pathologist by segmenting the differently coloured dyes that stain thin (< 30 μm) sections of biological tissue. These histo-chemical markers can reveal specific cellular and tissue characteristics and are commonly used as a critical form of patient diagnosis. A spectral imaging microscope is capable of resolving spectral changes in optical density and can therefore be useful in segmenting these dyes, especially when their apparent colours are similar. Most dyes used in histology have very broad absorption spectra and it is this feature that limits the applicability of conventional RGB camera. In general these dyes have absorption spectra that vary slowly across the visible with structures at the scale of 20 nm or more. However, the shapes of the spectra are often specific to the dyes and spectral imaging exploits this characteristic[1,2]. This technique is also equally applicable to fluorescent stains where emission spectra overlap and so are not readily separated by traditional microscope filters and "fluorescence cubes".

Multiple band-pass filters placed immediately in front of the camera may be used to create a spectrally resolved imager. Although such a system may be simple to use and cost effective for screening a specific set of dyes it is inflexible when compared to a tunable device that allows multiple screening procedures, inexpensively, with a single unit.

Few practical methods for achieving tunable spectral imaging microscopy have been developed in recent years. The most promising of these involve the use of either a liquid crystal or acousto-optic spectral filter, or the use of a Sagnac interferometer. In summary, the use of an acousto-optic filter has advantages of speed (wavelength selection times of the order of microseconds), variable resolution (down to 1 nm) and a wide wavelength range (500 – 1000 nm).

* barber@gci.ac.uk; phone +44 1923 828-611; fax +44 1923 835-210; www.gci.ac.uk

Spectral Imaging: Instrumentation, Applications and Analysis III, edited by
Gregory H. Bearman, Anita Mahadevan-Jansen, Richard M. Levenson, Proc. SPIE
Vol. 5694 (SPIE, Bellingham, WA, 2005) · 1605-7422/05/$15 · doi: 10.1117/12.590547

However, this is often a specialised and expensive solution. A liquid crystal filter is more cost-effective but this technique suffers from low transmission (<30 %), limiting its application in fluorescence microscopy where signal intensities are low. A different approach, based on a Sagnac interferometer[3,4] is similarly specialised and expensive, offering a <10 nm resolution and <50 % transmission. It has the additional disadvantage that a full spectral scan must always be captured precluding the possibility of faster partial scans or different dwell times at specific regions of the spectrum. This is particularly problematic when the signal is photon-limited, where the signal-to-noise ratio can be improved by optimising the collection time. For a fuller review of techniques see the relevant book chapter by D.L. Farkas[5]. For this study, a system based around a Linearly Variable Filter (LVF) was used incorporating novel hardware and software to produce a cost-effective solution of reasonable performance compared to other techniques.

The spectra indicative of different markers that have been used to stain a sample can be separated out from the measured spectrum using a linear least squares algorithm. This technique is valid if the individual reference spectra add linearly to form the spectrum observed. This is obviously so for fluorescence signals, we must just take care to work within the dynamic range of the camera and that the response of the camera is linear. For trans-illumination experiments, where the spectra measured are of the samples absorption, not of the light transmitted, the linear relationship also hold true. Given this assumption, the separation, or "un-mixing", of biological markers from a spectral image is a relatively simple task of linear least-squares fitting. The proportions of the individual reference spectra are the fitted parameters given the reference and measured spectra. Since the problem is generally over-determined, i.e. the numbers of points that constitute the spectra are generally much greater than the number of fit parameters, the results are generally robust.

We have previously shown that this technique can be used to detect minor differences in spectra, by testing the system performance with combinations of coloured filters, and that this is a valid technique for use in determining clinical outcome[6,7,8].

The increased use of the linear least-squares un-mixing technique has led to concerns over the processing speed. The large amount of data to process has resulted in processing times of several minutes on desktop computers. A typical spectral image for this study consisted of 768 by 576 pixels each containing a 50-point spectrum. For our test samples there were 3 reference spectra leading to 442,000 spectral fits involving 200 points, all leading to the determination of 1.3 million free parameters. Thankfully, each pixel is independent and the calculation can be readily made parallel by splitting up the image. These facts lead us to pursue cluster computing as a cost effective solution to speed up to processing by utilising additional computing power. A computer cluster may consist of several ordinary desktop computers linked by a standard Ethernet network. With appropriate software the computers of the cluster can work in parallel to perform tasks in a fraction of the time required by a single desktop. This arrangement highly cost effective when compared to a multiple processor supercomputer.

There has recently been great interest in harnessing the power of a computing cluster and much work has been done by the "open source" software development community in developing "Beowulf" clusters[9] based on the Linux operating system. They have remained, in some ways, more of a developer's curiosity than usable tool, but not without exception; See the work of Dagget and Greenshields[10] for example. Further software development has lead to a greater degree of usability and the Condor Project[11] provides a usable "high-throughput" computing system. It is important to note that for our type of application we are more interested in "high-power" computing (where a single result is returned quickly) than "high-throughput" (where a high number of processes are left created and left for several hours to return many results) and, in addition to this, we are aiming for a highly usable system that requires minimal intervention from computer specialists for day-to-day running. For ease of development and use, we have traditionally used exclusively the Windows operating system and for this reason a cluster based on Windows was required. The time to reach the same level of instrumentation development under Linux would be excessive. Our aim was for the user, who has no cluster computing training, to start the usual software that drives the Spectral Imager and transparently experience much greater processing power provided by a cluster.

This paper describes our spectral imaging system and computing cluster and demonstrates the processing speed up achieved.

2. METHODOLOGY

2.1. Image capture

The spectral imaging device used for this study was developed and constructed in our Institute[6]. It uses a standard monochrome CCD camera (type 4912, Cohu Inc., USA) though other types of imagers, with increased sensitivity or resolution may be used. The spectrally selective element is placed between the camera and the microscope output port

using standard C-mount couplers and is based on a linearly variable dielectric band-pass filter together with novel drive hardware and acquisition software. The element has a resolution of 15 nm and covers the 400 nm to 700 nm band with a transmission of > 40 %. The system provides wavelength agility and allows for user-set dwell times. For this study, the spectrally resolved device was used with an upright microscope (Optiphot, Nikon, UK) equipped with a achromatic objective lens was used: 10x (0.25 NA).
Images were captured into a personal computer (PC) using a 1 GHz processor (Dell, UK, Precision 220, 256 Mbytes RAM) via a frame grabber (type PCI-1409, National Instruments Ltd., UK) and all software was written in the 'C' programming language under the LabWindows/CVI™ development environment (National Instruments Ltd., UK) and WindowsXP operating system (Microsoft Corp., USA).
The system is capable of capturing the spectral image of an emissive sample in ~10 s, spanning 400 to 700 nm in 6 nm steps (50 images). Optical density (OD) spectra were obtained by first capturing the spectral profile of the illumination through a blank part of the sample but of equivalent optical thickness and a 'black' image with no illumination. These images were used to correct the spectral image taken through the sample and to calculate the wavelength-dependant optical density (OD(λ)) according to the equation:

$$OD(\lambda) = -10\log\left(\frac{I(\lambda) - I_{black}}{I_{blank}(\lambda) - I_{black}}\right) \quad (1)$$

where λ represents wavelength and I the intensity at a single image pixel.
Linear un-mixing of the observed spectra was achieved by least squares fitting, as previously described[12]. In this study we chose to use a non-negative least squares algorithm[13] since this approach cannot give rise to negative contributions that clearly have no physical meaning. Distribution maps for each constituent can thus be obtained from this un-mixing process where the results represent proportions of the reference spectra, having units of OD normalised to the references. This normalised optical density is mapped to pixel intensity such that areas of heavy staining appear bright in the resulting image.

2.2. Computing cluster

The cluster consisted of 4 desktop computers connected by an Ethernet switch. It was usual to initiate processing from a fifth "master" computer, also connected to the switch. All computers run Microsft® Windows XP™ Professional operating system (Microsoft, CA, USA) and were built, in-house, using off the shelf components. They contained dual-processors based on the Athlon™ range of chips, manufactured by AMD (Sunnyvale, CA, USA); these were selected for their enhanced floating point performance. Figure 1 shows schematic diagrams of the cluster as a whole and the relevant parts of a slave computer.
The four "slave" computers were based on Athlon™ MP 2600+ processors with a 2.1 GHz clock speed and the master was based on Athlon™ MP 2200+ processors with a 1.8 GHz clock speed. The role of the master of the cluster was usually to distribute data to the slave machines and to collects the results after processing; these tasks were much less processor demanding. All cluster computers contained the Gigabyte 7DPXDW+ motherboard (Giga-Byte Technology Co., Ltd., Taiwan, R.O.C.) with built in 10/100 Ethernet adapter (IEEE 802.3 100Base-T) and based on the AMD 760MPX chipset (AMD, Sunnyvale, CA, USA). This chipset supplies fast 64-bit, 66 MHz, PCI slots that are directly connected to the AMD 762 "Northbridge"[14]. Gigabit Ethernet (IEEE 802.3 1000Base-T) 64-bit PCI V2.2 adapters were added via this connection. These were unbranded of type NE-320G-TX. Each slave computer had 1 Gb of random access memory (RAM) of type: 2 x 512 Mb ECC Registered DDR PC2100 (Corsair Memory, Fremont, CA, USA).
The Ethernet network connecting all computers was constructed using one of two Ethernet switches which supported either 100 Mbit/s (100Base-T) or 1 Gbit/s (1000Base-T) raw data transfer speeds. One was a 100Base-T 8-port 10/100 + 2-port gigabit switch (unbranded type: eTen EW5082V) and the second was a 1000Base-T 8-port gigabit switch (unbranded type: eTen GSW308T). These were connected to the computers using Ethernet cable of type Cat-5[15].
The computer program to drive the spectral imager and to perform all processing, including that done via the cluster was written in the C programming language under Labwindows/CVI development environment (National Instruments Ltd, UK.). Communications between the master and slave computers was driven using the so-called "Message Passing Interface" (MPI)[16] protocol. Two software libraries where used: initial development was done using the freely available implementation of the MPI version 1 standard, MPICH[17], further development was performed using the commercial WMPI II (Critical Software SA, Portugal) which is an implementation of the MPI version 2 standard.

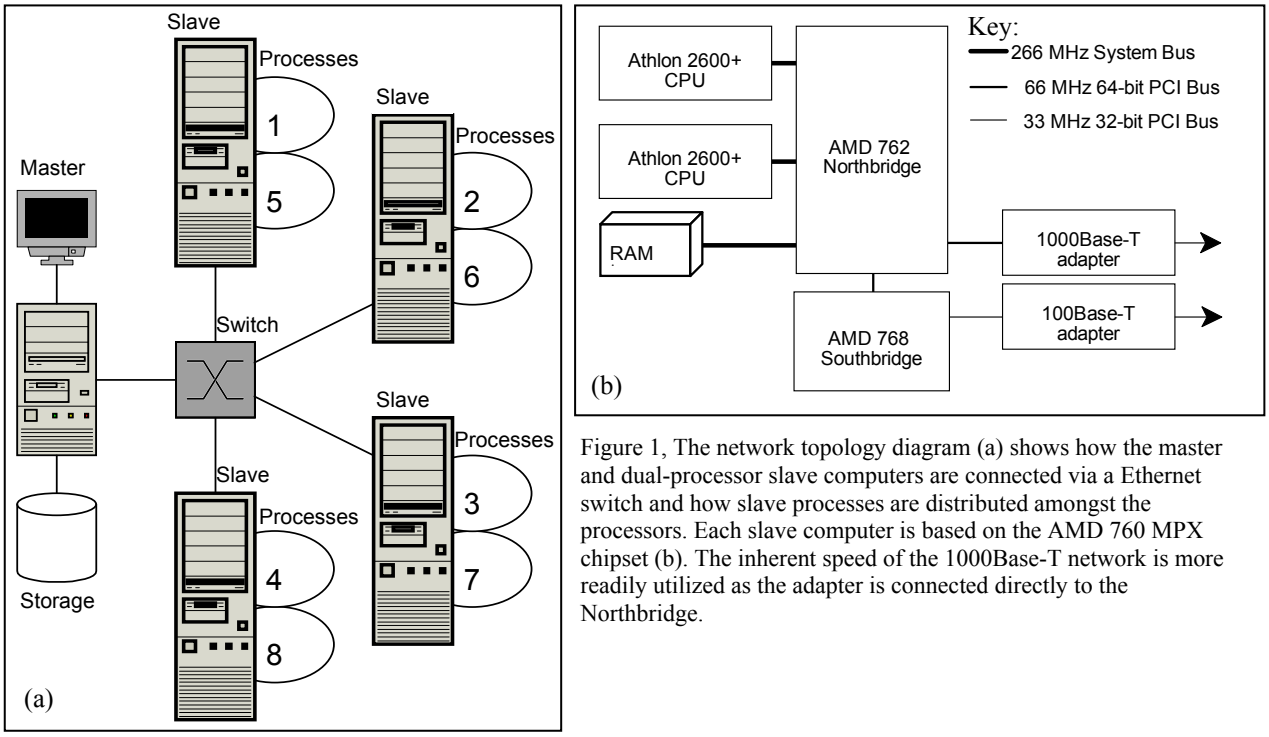

Figure 1, The network topology diagram (a) shows how the master and dual-processor slave computers are connected via a Ethernet switch and how slave processes are distributed amongst the processors. Each slave computer is based on the AMD 760 MPX chipset (b). The inherent speed of the 1000Base-T network is more readily utilized as the adapter is connected directly to the Northbridge.

MPI is an extremely flexible standard and a layer of code was written to abstract the salient parts of MPI and simplify them for the master/slave type of cluster and to provide and easier interface for programming.

The code structure was designed for ease of maintenance by compiling just one executable file which contains instructions for both the master and slaves. The code identifies itself as master or slave according to the "rank" of the process provided by MPI. Rank 0 being the master; any other being a slave. The cluster as a whole contains a number of processors, 8 in this case, and a decision must be made about how many software processes to start that will run concurrently on the available processors. With MPI version 1 this decision must be made by the user and initial tests showed that the optimum number of processes to start is equal to the number of processors. This provides the maximum speedup with least overhead. We can easily see that if we create too many then some processors may have more work than others (more than one process) and time may be wasted waiting for the processes to finish before the final result can be presented. With MPI version 2 the master process has the ability to create other processes; this enables the optimum number of processes to be created automatically.

The code execution proceeded as follows: slave processes would start and initialise and then wait for data from the master. The master process would also start and initialise and also proceed with normal events, such as acquire or load and display data. At the time that cluster processing is required the master would send sections of data to the slaves, the sections are equal in size, and then wait for the results to be returned. When a slave process had finished processing it returned a result to the master. The master collected all the results from the slaves and presented the combined result to the user. In the case of linear un-mixing with spectral image data the same processing was performed on each pixel independently and so the data could by simply portioned up into equal slices and sent to the slaves. This is a "single instruction - multiple data" arrangement.

Testing of the cluster was done by timing the linear un-mixing process five times, noting not only the total processing time but also the data transfer time. The spectral image used for testing was 41 Mbytes in size and this was split up and distributed, together with the reference spectra, to each of the slave processes. Each slave process then returned a result as sections of three floating point images representing the un-mixed stains.

Common cluster computing analysis metrics were used to determine the cluster performance. These were speedup (S), efficiency (ε) and efficacy (η). The speedup is defined as the ratio of the processing speed of the cluster with a number of processes (n) compared to a single process. Efficiency is defined below and values above 0.5 are usually considered good[18].

$$\varepsilon = \frac{S}{n} \tag{2}$$

In order to optimise the number of cluster processors, and give an indication of scalability, some kind of cost-benefit analysis is required as maximising ε would usually corresponds to $n = 1$. Maximising S would encourage the use of too many processors. If we define S as our benefit, at a cost of $1/\varepsilon$ we can define efficacy as[19]:

$$\eta = \frac{S}{1/\varepsilon} = \frac{S^2}{n} \tag{3}$$

2.3. Histology test samples

Human tumour biopsies were used to provide multiply stained histology sections. Formalin fixed and paraffin embedded tumours were used to provide 4-μm-thick tissue sections, stained for carbonic anhydrase IX (CA9), pimonidazole (PIMO), and with a haematoxylin (Hx) counter-stain. This was performed using a previously described protocol[6].

3. RESULTS

Most clinical tumours contain a significant sub-population of hypoxic tumour cells[20,21] which can be visualised immuno-histochemically through the intrinsic marker glucose carbonic anhydrase IX (CA9)[22] and extrinsic hypoxic markers such as pimonidazole (PIMO)[23]. An example of separating dyes from a multiply stained histological section is shown in figure 2. A section from a biopsy of bladder carcinoma was stained with PIMO (resulting in a brown colouration), CA9 (red colouration) and Hx (blue colouration). PIMO and CA-9 are known markers of hypoxia and are expected to show considerable overlap in their localisation in tumour tissue. Figure 2 (b, c and d) show the results of digitally extracting these dye distributions from the image taken with the spectral imager and the 10x objective. In this case the reference spectra were taken from separate singly stained sections of the same biopsy (figure 2 (e)).

The time taken to perform the linear un-mixing operation on the test image was measured in various cluster configurations. The average of 5 timings was taken and the results are presented in Table 1 and graphically in figure 3. In most cases the master computer was used to drive a number of cluster processes on the cluster computers. In the case of "zero" cluster processors, one of the slave machines was used as its own master and all operations were performed by a single process.

Table 1, Time spent by clusters of increasing size in linear unmixing the test spectral image using the 100Base-T and 1000Base-T Ethernet networks. The average and standard deviation of 5 measurements are shown.

	Number of Cluster Processors								
100Base-T	0	1	2	3	4	5	6	7	8
Total Time (s)	16.6 ± 0.03	21.2 ± 0.05	13.6 ± 0.03	11.3 ± 0.05	10.4 ± 0.08	9.54 ± 0.05	9.12 ± 0.06	8.76 ± 0.02	8.39 ± 0.06
Transfer (s)	-	4.01 ± 0.03	4.77 ± 0.02	5.16 ± 0.01	5.48 ± 0.06	5.48 ± 0.05	5.60 ± 0.04	5.59 ± 0.03	5.55 ± 0.04
1000Base-T	0	1	2	3	4	5	6	7	8
Total Time (s)	16.6 ± 0.03	17.6 ± 0.03	9.47 ± 0.03	6.88 ± 0.02	5.61 ± 0.03	4.78 ± 0.02	4.27 ± 0.01	3.89 ± 0.02	3.57 ± 0.02
Transfer (s)	-	0.85 ± 0.02	1.00 ± 0.03	1.11 ± 0.01	1.22 ± 0.01	1.22 ± 0.02	1.26 ± 0.03	1.21 ± 0.02	1.19 ± 0.02

From these data the speed-up achieved can be calculated and using equations 2 and 3 the efficiency and efficacy were also calculated and are presented in figure 4. In all cases the number of slave processes equalled the number of slave processors. It is worth noting here the poor performance of the 100Base-T cluster where the efficacy peaks at two processors and correspondingly the efficiency soon drops below 0.5 as you add more processors. This indicates that a cluster built on the 100Base-T network is not efficiently scalable beyond two processors. In contrast the 1000Base-T network provides a much better basis on which to base a cluster. The peak efficacy was not reached with the 8 processors that were available to us indicating that it would still be greatly beneficial to add processors. The speedup achieved with 8 processors was almost a factor of 5 (see figure 4b).

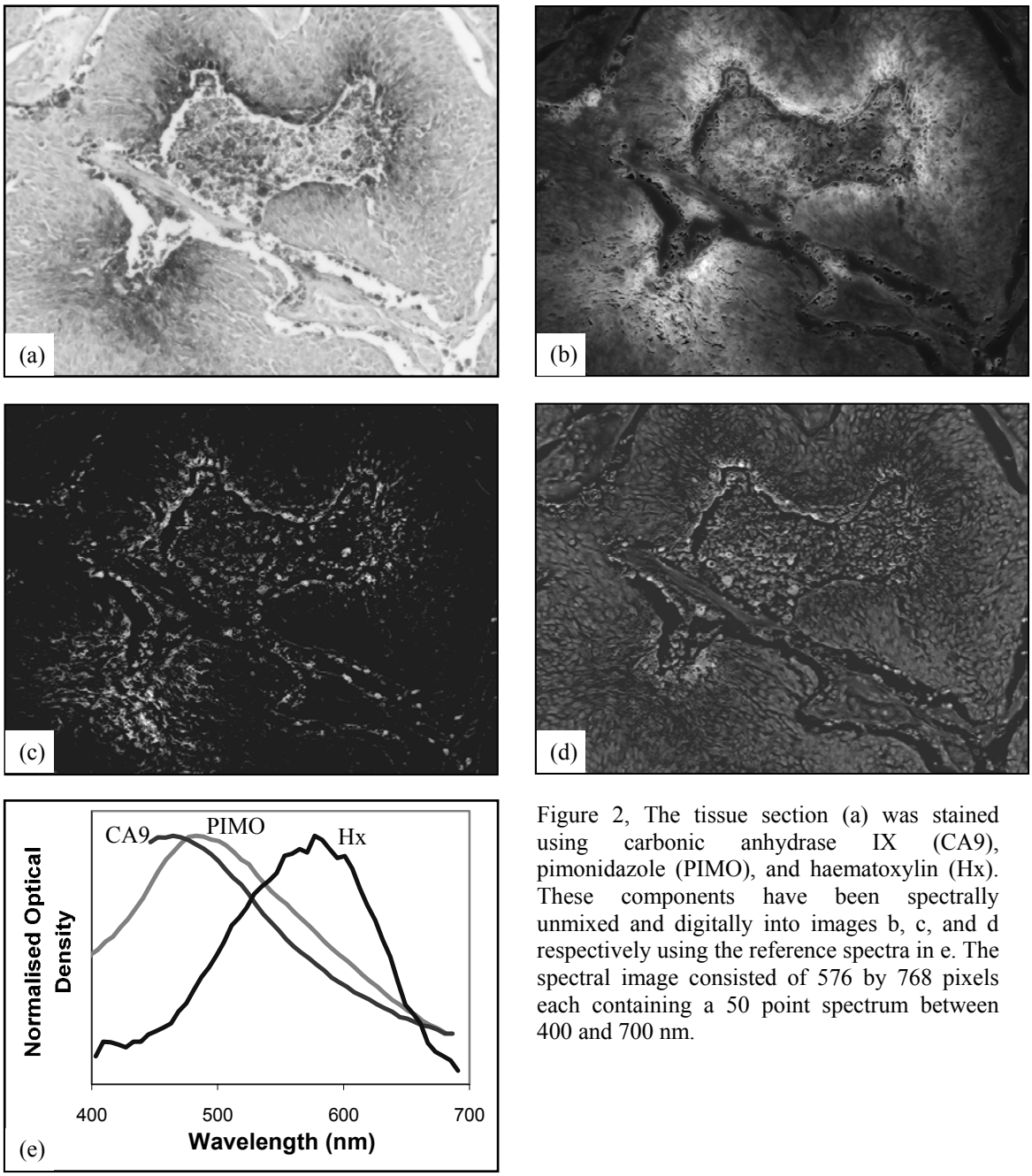

Figure 2, The tissue section (a) was stained using carbonic anhydrase IX (CA9), pimonidazole (PIMO), and haematoxylin (Hx). These components have been spectrally unmixed and digitally into images b, c, and d respectively using the reference spectra in e. The spectral image consisted of 576 by 768 pixels each containing a 50 point spectrum between 400 and 700 nm.

In figures 3 and 4 the points indicating the performance of a cluster with zero processors (n=0) are also shown. These points indicate an initial drop in performance when clusters of numbers of processors are formed. The addition of network transfer time adds a time overhead that is not present when n=0 and is the source of this drop in performance. The number of processors had to be increased for the increase in processing speed to outweigh the network transfer time and provide an overall benefit.

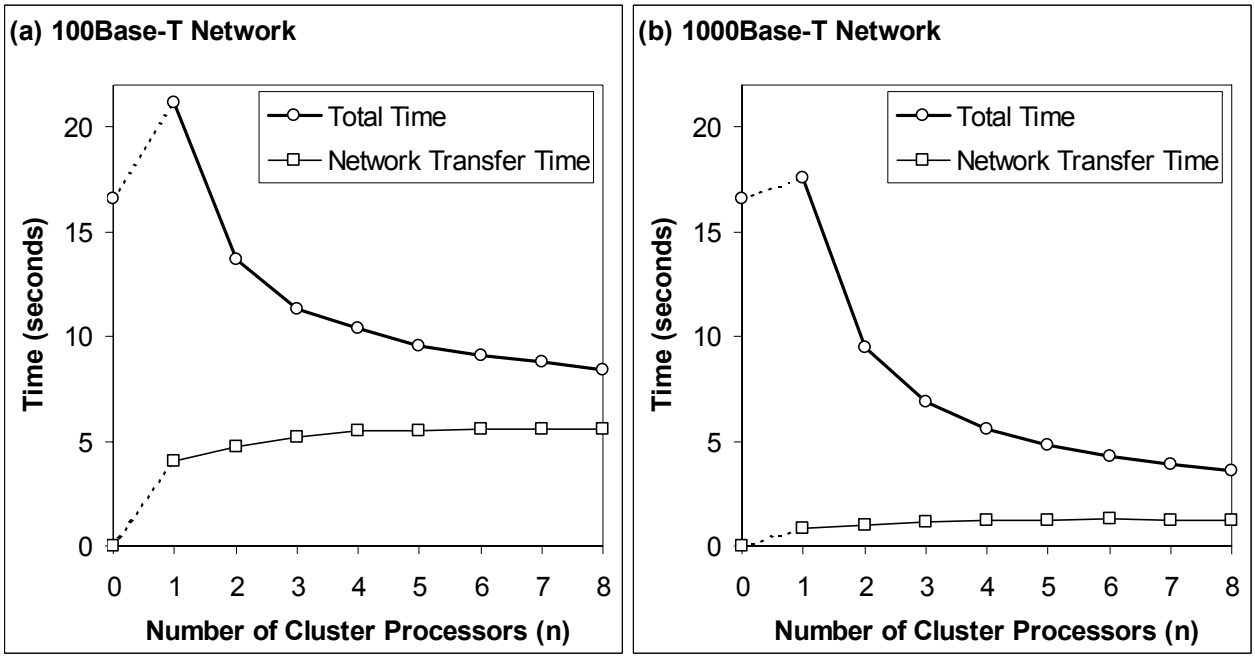

Figure 3, Testing the computing cluster with an increasing number of cluster processors.

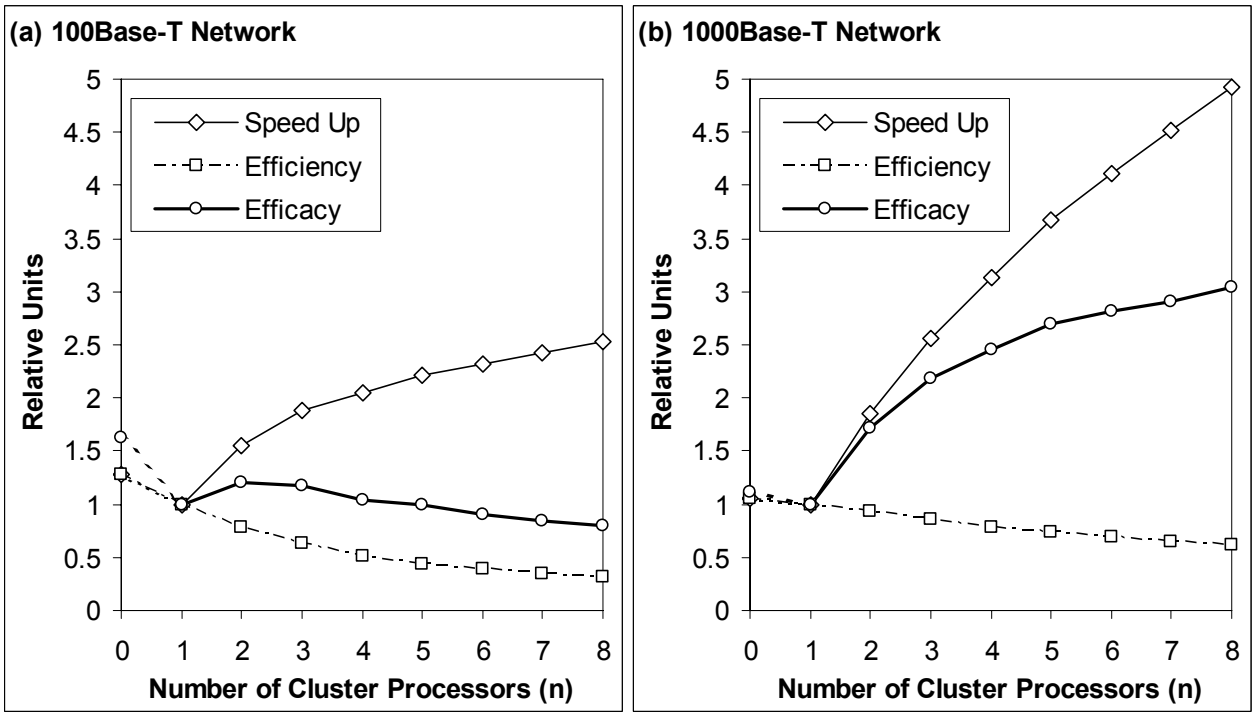

Figure 4, The speedup, efficiency and efficacy of different sized clusters. Using the 1000Base-T network the efficacy is still increasing at 8 processors, with the efficiency still above 0.5 (b).

To explore the network transfer performance, data transfer rates were measured directly using just two slave computers and a 41.3 Mbyte dataset. Transfers were performed between the RAM of the two computers and so were not dependant on the computers hard disk drive read/write and transfer performance. Example transfer rates through 100Base-T, 1000Base-T, and a IEEE 1394 network configurations are shown in table 2. The IEEE 1394 network was formed using

the standard "Internet Protocol (IP) over IEEE 1394" provided by the WindowsXP operating system. The theoretical maximum data rates for the different networks is also shown. The IEEE 1394 network was tested as a possible, cost-effective, alternative to 1000Base-T, requiring only an OHCI IEEE 1394 interface which is standard with most new PC's. In this case it was desirable that the dataset is too large to be stored in any cache of the AMD chipset which may give false results.

Table 2, Comparison network transfer times for memory to memory transfer of a single 41.3 Mbyte dataset. The average and standard deviation of 5 measurements are shown.

	100Base-T Network	***1000Base-T Network***	***IEEE 1394 Network***
Transfer Time	3.88 ± 0.03 s	0.69 ± 0.03 s	1.88 ± 0.03 s
Data Rate	85.2 ± 0.7 Mbit/s	481 ± 21 Mbit/s	176 ± 3 Mbit/s
Theoretical Maximum	100 Mbit/s	1000 Mbit/s	400 Mbit/s

4. CONCLUSIONS

A cost effective spectral imager, as previously described[6], can be used to un-mix multiple stains or markers from tissue sections or cell samples. This linear un-mixing of multiple stains has application for aiding quantitative analysis of fluorescent and absorptive samples whilst retaining feature context through both imaging and by allowing the use of counter-stains that provide visibility of the rest of the sample. We have previously shown how the un-mixing of distinct colours can be achieved using the tri-band (red-green-blue) colour information of an ordinary colour camera, but the use of a spectral imager offering a 15-nm optical wavelength resolution can un-mix stains that spectrally overlap to a great extent.

Quantitative biological assays usually involve many samples to achieve an acceptable statistical accuracy and so the stain un-mixing process has to be performed many times. Indeed, a system that offers almost 'real-time' un-mixing of the sample stains would be highly desirable. Unfortunately, the linear un-mixing process on a desktop computer can be time consuming when performed on spectral images with many megabytes of data. This problem has been the subject of this coding project to employ cluster computing, in a transparent and easy to use manner, in order to reduce the processing time. To that end a computing cluster based on the WindowsXP operating system and code written in C has been developed and tested. The Message Passing Interface (MPI) versions 1 and 2 have been used to implement the transfer of data and results between master and slave computers. Two Ethernet networks were evaluated as part of the cluster hardware. An Ethernet 100Base-T network was form using the motherboard-integrated network adapters of the computers and the performance was compared to that achieved when adding 1000Base-T network adapters to the computers 64-bit PCI bus. The performance of the 100Base-T network was poor in this application where large datasets have to be transferred. The 1000Base-T network proved very scalable up to our maximum number of processors (eight) with a speedup of 5 times and a cluster efficacy of 3.0 with al eight processors running. The main aim of this work was high-power computing to return a single result to the user quickly such that real-time un-mixing may be built upon it. Code optimisations have reduced the processing time to 17 seconds from previously reported times (over 130 seconds on a 1 GHz PC[6]) on a single desktop computer. Introducing cluster computing has reduced processing times again to around 3.5 seconds. This is now of the order of the image acquisition time and "real-time" un-mixing could be achieved by processing one image as the next is acquired.

An IP over IEEE 1394 network was also tested but despite a theoretical maximum data rate of 400 Mbit/s, only 176 Mbit/s was achieved compared to the 481 Mbit/s of the 1000Base-T network. At the time of writing, 1000Base-T networking is now becoming common in new pc's and the component prices are likely to fall and so seems to be the network of choice for cost effective cluster solutions.

We have shown that using standard, low-cost, computer components an effective computing cluster can be built for "high power" applications. Code has been written for a standard and widely used operating system and will be portable to newer and faster computers and network protocols as they are developed.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge the technical support of D. Stewart, R. Newman, J. Prentice and R. Locke for computing, electronic, mechanical workshop and interface coding assistance respectively. We must also give thanks to Critical Software SA, Portugal, for supplying WMPI II at a greatly reduced cost. This work was funded by Cancer Research UK, programme grant C133/A1812 – SP2195-01/02, and by the UK Research Councils Basic Technology Programme.

REFERENCES

[1] Ornberg, R. L., Woerner, B. M. and Edwards, D. A. "Analysis of Stained Objects in Histological Sections by Spectral Imaging and Differential Absorption." *The Journal of Histochemistry & Cytochemistry* **47**, 1307–1313 (1999).

[2] Zhou, R., Hammond, E. H. and Parker, D. L. "A multiple wavelength algorithm in color image analysis and its applications in stain decomposition in microscopy images." *Medical Physics* **23**, 1977-86 (1996).

[3] Lavi, M., et al. "A new compact design interferometer based spectral imaging system for bio-medical applications." *Proc. SPIE* **3261**, 313-321 (1998).

[4] Barshack, I., Kopolovic, J., Malik, Z. and Rothmann, C. "Spectral morphometric characterization of breast carcinoma cells." *British Journal Of Cancer* **79**, 1613-1619 (1999).

[5] Farkas, D. L. in *Methods in Cellular Imaging* (ed. Periasamy, A.) 345-361 (Oxford University Press, 2001).

[6] Barber, P. R., Vojnovic, B., Atkin, G., Daley, F. M., Everett, S. A., Wilson, G. D. and Gilbey, J. D. "Applications of cost-effective spectral imaging microscopy in cancer research." *J. Phys. D,* **36**, 1729 (2003).

[7] Carnell, D. M., Smith, R. E., Daley, F. M., Barber, P. R., Hoskin, P. J., Wilson, G. W., Murray, G. I. and Everett, S. A. "Target validation of cytochrome p450 cyp1b1 in prostate carcinoma and protein expression in associated hyperplastic and pre-malignant tissue." *Int. J. Radiat. Oncol. Biol. Phys.,* **58**, 500-509 (2004).

[8] Greer, M. L., Daley, F. M., Barber, P., Murray, G. I., Patterson, L. H. and Everett, S. A. "Cytochrome P450CYP1B1 versus GLUT 1 protein expression during the development of head and neck squamous cell carcinomas (HNSCCs)." *British Journal of Cancer* **88**, S33-S33 (2003).

[9] http://www.beowulf.org

[10] Daggett, T. and Greenshields, I. R. "A cluster computer system for the analysis and classification of massively large biomedical image data." *Computers in Biology and Medicine* **28**, 47-60 (1998).

[11] http://www.cs.wisc.edu/condor/

[12] Farkas, D. L., Du, C., Fisher, W., Lau, C., Niu, W., Wachman, E. S. and Levenson, R. M. "Non-invasive image acquisition and advanced processing in optical bioimaging." *Computerized Medical Imaging And Graphics* **22**, 89-102 (1998).

[13] Lawson, C. L. and Hanson, R. J. *Solving Least Squares Problems* (Prentice-Hall, Englewood Cliffs, N.J., 1974, Revised edition, SIAM Society for Industrial & Applied Mathematics, 1995, 1974).

[14] AMD-760TM MPX Chipset Overview, Publication number 24494, Dec. 2001, http://www.amd.com/us-en/assets/content_type/white_papers_and_tech_docs/24494.pdf.

[15] Commercial Building Telecommunications Wiring Standard, EIA/TIA 568.

[16] http://www-unix.mcs.anl.gov/mpi/

[17] Argonne National Laboratory, University of Chicago and Mississippi State University. http://www-unix.mcs.anl.gov/mpi/mpich/

[18] Goller, A., Glendinning, I., Bachmann, D. and Kalliany, R. in *Digital Image Analysis: Selected Techniques and Applications* (eds. Kropatsch, W. G. and Bischof, H.) 135-153 (Springer-Verlag New York, 2001).

[19] Goshal, D., Serazzi, G. and Tripathi, S. K. "The processor working set and its use in scheduling multiprocessor systems." *IEEE Trans. Software Engineering* **17**, 443-453 (1991).

[20] Hockel, M. and Vaupel, P. "Biological consequences of tumor hypoxia." Seminars in Oncology 28, 36-41 (2001).

[21] Vaupel, P., Kelleher, D. K. and Hockel, M. "Oxygen status of malignant tumors: pathogenesis of hypoxia and significance for tumor therapy." *Seminars in Oncology* **28**, 29-35 (2001).

[22] Wykoff, C. C., Beasley, N. J., Watson, P. H., Turner, K. J., Pastorek, J., Sibtain, A., Wilson, G. D., Turley, H., Talks, K. L., Maxwell, P. H., Pugh, C. W., Ratcliffe, P. J. and Harris, A. L. "Hypoxia-inducible expression of tumor-associated carbonic anhydrases." *Cancer Research* **60**, 7075-83 (2000).

[23] Arteel, G. E., Thurman, R. G. and Raleigh, J. A. "Reductive metabolism of the hypoxia marker pimonidazole is regulated by oxygen tension independent of the pyridine nucleotide redox state." *European Journal of Biochemistry* **253**, 743-50 (1998).