FAST PARALLEL IMAGE MATCHING ALGORITHM ON CLUSTER

Zhenge.Qiu [a, *], ZengBo.Qian [a] , Zhihui Gong [a], Qing Xu [a]
[a] Zhengzhou Institute of Surveying and Mapping, 66# Longhai Middle Road Zhengzhou China,
Qiuzhenge@sina.com

Commission IV

KEY WORDS: Parallel Algorithm, Cluster, DSM (Distributed Shared Memory), Image Matching, DEM (Digital Elevation Model), No-uniform Image sampling

ABSTRACT:

This paper introduces our experiences with developing fast Parallel stereo matching algorithm on Cluster. The key technique of generating DEMs or DSMs' from remote-sensing images is stereo matching, and it's one of the most time consuming algorithms, so that many efforts have been carried out to develop fast algorithms. Many difficulty problems continue to challenge researchers in this field; they are occlusion, large parallax range, and radiant distortion. To solve these problems need more complicated image-matching algorithm that is more time consuming. Recent advance in high-speed networks, rapid improvement in microprocessor design, and availability of highly performing clustering software implementations enables cost-effective high-performance parallel computing on clusters very attractive. The cluster is a very useful platform on which to develop fast parallel image matching algorithm dealing with the difficult problems above-mentioned. A novel operator of image matching based on no-uniform image resample is used in parallel image matching algorithm, this operator reduces the influence of geometric distortion on those matching operators which are in effect under equal parallax assumption, for example, correlation operator, and expands the scope of them to be in effect.

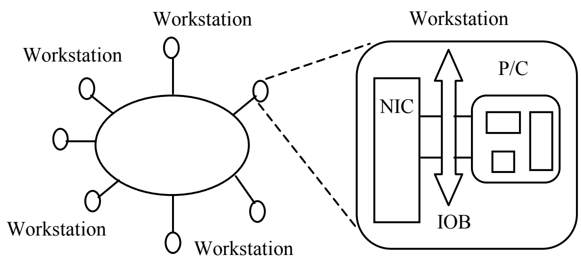
1. INTRODUCTION

The key technique of generating DEMs or DSMs' from remote-sensing images is stereo matching, and it's one of the most time consuming algorithms, so that many efforts have been carried out to develop fast algorithms of image matching. Many difficulty problems continue to challenge researchers in computer vision and digital photogrammetry and other scientific field. The main problems are occlusion, large parallax range, and radiant distortion. To solve these problems need more complicated image-matching algorithm that is more time consuming.

The current trend in high performance computing is clustering and distributed computing. In clusters, powerful low cost workstations and/or PCs are linked through fast communication interfaces to achieve high performance parallel computing as depicted in figure1. Recent increases in communication speeds, microprocessor clocks, protocol efficiencies coupled with the availability of high performance public domain software including operating system, compiler tools, and message passing libraries, make cluster based computing appealing in terms of both high-performance computing and cost-effectiveness. Parallel computing on clustered systems is a viable and attractive proposition due to the high communication speeds of modern networks.

MPI (Message Passing Interface) approach is considered one of the most mature methods currently used in parallel programming mainly due to the relative simplicity of using the method by writing a set of library functions or an API (Application Program Interface) callable from C, C++ or Fortran programs.

DSM (Distributed Shared Memory) is another mature methods currently used in parallel programming. DSM systems provide the illusion of shared memory on top of standard message passing hardware at very low implementation cost, but provide acceptable performance for much class of applications. DSM provides a good compromise between programmability of shared memory multiprocessors and hardware simplicity of message passing multicomputers. Many software DSM systems, such as Ivy [8], Midway [l], Munin [2], TreadMarks [7], Cashmere [10], and JIAJIA[3] have been implemented on the top of message passing hardware.



P/C (Microprocessor and Cache) NIC (Network Interface Circuitry)
IOB (I/O Bus)

Figure1. Cluster of PC

This paper introduces our experiences with developing fast Parallel stereo matching algorithm on the home-based software DSM JIAJIA [3] and discusses techniques of parallelizing a sequential image matching program to run on software DSM.

Satisfactory speedups are achieved for all of these applications on a cluster of eight Pentium II PCs connected by a 100Mbps switched Ethernet. The rest of this paper is organized as follows. Section 2 discusses the image-matching algorithm based on no-uniform sampling, JIAJIA Software DSM as a useful parallel software-developing tool is introduced in Section 3. Section 4 discusses methods for parallelizing sequential image matching programs. The conclusion of this paper is drawn in Section 5.

2. IMAGE MATCHING ALGORITHM BASED ON NON-UNIFORM SAMPLING

The geometric distortion is the main problem of image matching; the no uniform sampling can effectively reduce the influence of geometric distortion, which, in essence, is the relief displacement of pairs of stereoscopic pictures. Figure 2 illustrates the case of vertical aerial photographs over flat terrain, in which all points in their planimetrically correct positions, every corresponding image points has the same parallax, the shapes of object imaged on the two photographs are identical and the correlation coefficient which is typically used in area matching will work perfectly with large image correlation windows. As shown in figure 2, $G(X,Y,Z)$ depicts the reflectance map of terrain and $g_1(x_l, y_l)$, $g_2(x_r, y_r)$ are the corresponding images on the pairs of stereoscopic pictures and their relationship can be defined by a constant d.

$$g_1(x_l, y_l) = g_2(x_l + d, y_l) \tag{1}$$

In this case, in the matching windows every corresponding points has a unique parallax, large window can be used to avoid ambiguous and at the right matching position the correlation will reach maximum.

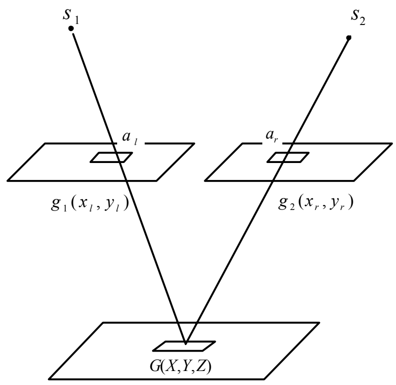


Figure 2. Vertical aerial photographs over flat terrain

Figure3 illustrates the case of vertical aerial photographs over no-flat terrain. For it is not enough, only base on the image of points a_l, a_r, to determine if they are corresponding points or not, so the neighbour image should be used, for example b_l, b_r which are the images of object point B. However, we have a trouble when we use the images of b_l, b_r, for the parallax of a_l, a_r and parallax of b_l, b_r are not identical, in other word, there is a geometric distortion between the image on left photo and that on right photo.

The geometric distortion between the image on left photo and that on right photo, in essence, is the problem of relief displacement due to height difference. As shown in figure3, object point B and point B_0 has the same planimetrically positions and different elevation, the difference between the displacement of b_l (on left image) and that of b_r (on right image) is geometric distortion, where b_l, b_r are imaged from B; b_{l_0}, b_{r_0} are imaged from B_0. Let δh_l be the relief displacement between b_l, b_{l_0} on left image, it can be expressed as

$$\delta h_l = \frac{\Delta h \cdot rb_l}{H} \tag{2}$$

Where rb_l is the radial image distance of b_l, Δh is the height of B above the reference object plane, H is the flying height.

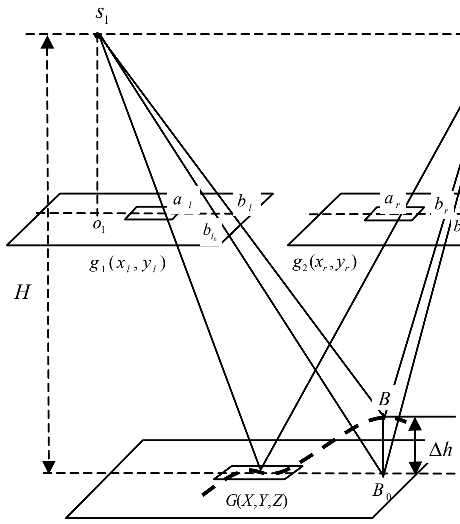


Figure 3. Vertical aerial photographs over no-flat terrain

On the right photo, b_r imaged from object point B also has a relief displacement δh_r between b_r, b_{r_0} which can be described as

$$\delta h_r = \frac{\Delta h \cdot rb_r}{H} \tag{3}$$

Where rb_r is the radial image distance of b_r 、 Δh is the height of B above the reference object plane, H is the flying height.

The geometric distortion between the image on left photo and that on right photo can be depicted as:

$$\Delta_{lr} = | \delta h_l - \delta h_r | \tag{4}$$

Equation 4 shows that if $|\delta h_l|$,$|\delta h_r|$ can be reduced, Δ_{lr} can be reduced dramatically.

From equation 2 and equation 3 we can obviously find such relationship that relief displacement is in direct ratio to Δh, rb_l or rb_r, and in inverse ratio to H. It is reasonable to suppose that terrain relief has a continuous change. For studying the change trend of relief displacement, it is also reasonable to suppose that possible maximum value of Δh is in direct ratio to distance of object B from A, this relationship can be depicted by figure 4. Figure 4(a) describes the distribution of relief displacement on left photo, and shows the case that a_l superposes fiducial center o_l, these circles represent the image points having same distance from fiducial center o_l or a_l, under the hypotheses above mentioned, these points also have the same possible maximum value of Δh, namely

$rb_l = c_1, \Delta h = c_2$, and the possible maximum value of Δh is in direct ratio to rb_l . According to equation 2 and equation 3, we can bring out that at each circle, every image points have a relief displacement of same possible maximum value that is in direct ratio to rb_l .

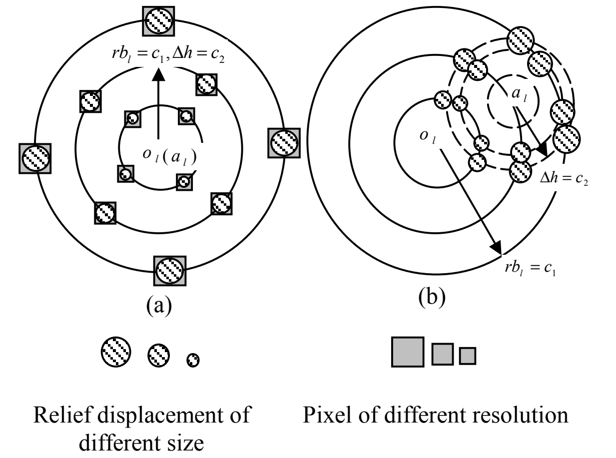


Figure 4. The distribution of relief displacement

The common case is showed by figure 4(b), the focusing point a_l does not superposes fiducial center o_l , the points at circle of real line all have same distance from fiducial center o_l , namely $rb_l = c_1$, the points at circle of broken line all have same distance from focusing point a_l , according to previous assumption it can be accepted that the their possible maximum value of Δh is also same , namely $\Delta h = c_2$, thus at the cross points of circle of real line and circle of broken line, all the value $rb_l \cdot \Delta h$ are same. Although in this case the distribution of relief displacement does not have symmetry characteristic as the case shown in figure 4(a), they are similar and the distribution of the relief displacement at circle of broken line also has the property of being in direct ratio to rb_l (namely the distance from focusing point a_l), the dissimilarity is that at the point near fiducial center o_l the value is smaller, on the contrary, at the point far away from fiducial center o_l the value is larger.

If we use no-uniform sampling as depicted in figure 4(a), namely in this kind of sampling method, the pixels at different circle has different size, and the size is in direct ratio to semi diameter of the circle (in the case depicted in figure 4(b) the size is in direct ratio to semi diameter of the circle of broken line). In discrete representation, if this kind of pixels of different size is used as metric unit, the numerical value of relief displacement will observably reduce. For example at certain circle, the value of relief displacement: $\delta h_l = 0.1mm$ and the pixel size of no uniform sampling here: $\Delta pixel = 0.2mm$, so the integer value of relief displacement is zero (0 $\Delta pixel$). Another advantage of no-uniform sampling is that the resolution at inner circle is higher than that at outer circle, in this way these image information near the matching point a_l (focusing point) is enlarged and those image information away from the matching point a_l is compressed. So if using same size of match window more information can be used in matching, more detail information in central part of match window is used, more information in low resolution in marginal part of match window is also used and the relief displacement of them is markedly reduced. This analysis of the function of no uniform sampling is also exact to right photo of stereo pair.

In this way every corresponding points' parallax in match window will be approximately equal at exact matched position, so this method can reduces the influence of geometric distortion on those matching operators which are in effect under equal parallax assumption, for example correlation operator, and expands the scope of them to be in effect.

The image matching strategy depicted in figure 5 used here are:

(1) Re-sampling the image at the focusing point on left and right photos ,and get two new images;

(2) Calculate the correlation coefficient of the two new images;

(3) Determine if they are matched images.

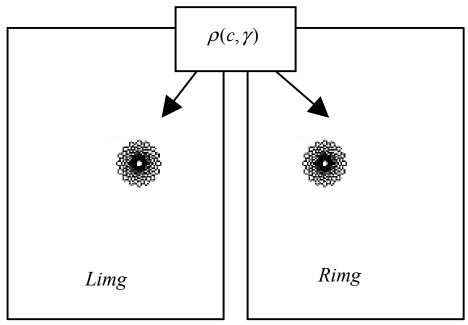


Figure 5. Image matching strategy

The correlation coefficient of discrete digital images can be calculated as follow:

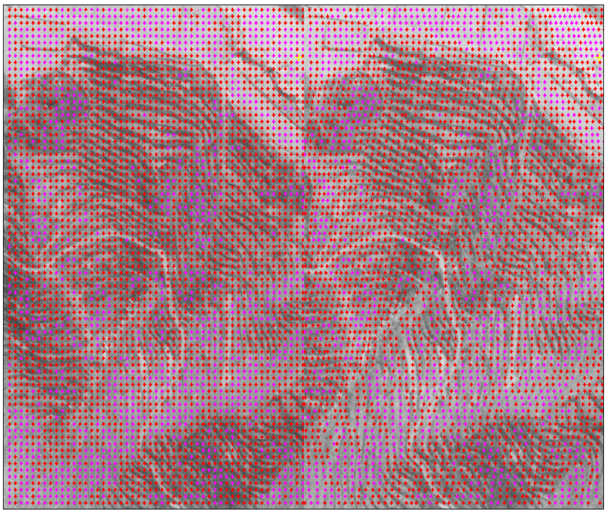
$$\rho(c,r) = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}(g_{i,j} - \overline{g})(g'_{i+r,j+c} - \overline{g}'_{r,c})}{\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}(g_{i,j} - \overline{g})^2 \sum_{i=1}^{m}\sum_{j=1}^{n}(g'_{i+r,j+c} - \overline{g}'_{r,c})^2}} \quad (5)$$

Where:

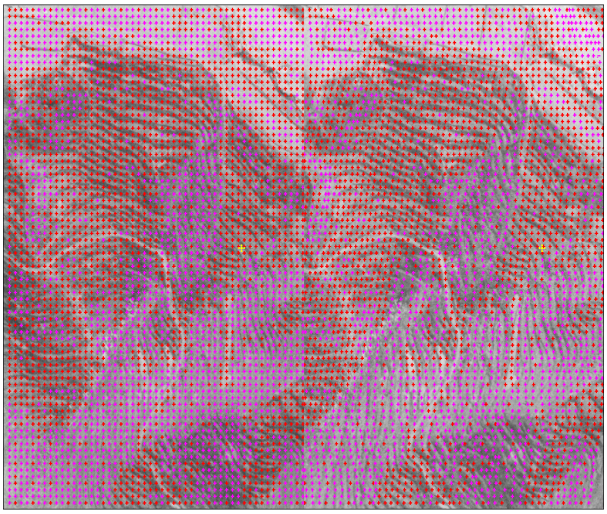
$$\left.\begin{aligned} g &= \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} g_{i,j} \\ \overline{g}'_{r,c} &= \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} g'_{i+r,j+c} \end{aligned}\right\} \quad (6)$$

This novel operator of image matching imitates the distribution of cone photoreceptors in the retina of man. In fovea there is high density of cone photoreceptors and high resolution of scenery, whereas in circumference there is low density of cone photoreceptors and low resolution of scenery. This mechanism not only can centralize the resource of computation on the centre part of match-window and use more wide scope image information, but also reduces the influence of geometric distortion on image match, which is the difference between relief displacements of two stereoscopic images in essence.

The aerial images of some area of china (1:20000) are used to test this novel image matching algorithm, the right matching ratio of this algorithm is about 70%-95%, about 15% higher than that of algorithm only using correlation coefficient, one of the example is showed in figure 6.



(a) Image matching algorithm based on no uniform sampling



(b) Image matching algorithm only using correlation coefficient

Figure 6. Image matching results

3. THE JIAJIA SOFTWARE DSM

With recent progress in software DSM and connection networks, software DSM can achieve an acceptable performance. On the one hand, optimization methods such as multiple-writer protocol and lazy implementation of release consistency can significantly improve the performance of software DSM. On the other hand, recent progress on networks reduces the sensitivity of parallel application performance to the communication amounts. Previous study shows that on a large variety of programs, the performance of well-optimized software DSM system is comparable to that of a message passing system. The fast parallel image-matching algorithm introduced in this paper is developed on the home-based software DSM JIAJIA.

JIAJIA is a home-based software DSM which supports scope consistency [6]. It has two distinguishing features compared to other recent software DSM systems. First, it combines physical memories of multiple computers to form a larger shared space. Second, it combines the lock-based cache coherence protocol and many optimization methods.

3.1 Programming Interface

The API of JIAJIA is similar to that of other software DSM systems. It provides following basic routines supporting SPMD (Single program multiple datum) shared memory parallel programming to the applications:

(1) jia_init(argc,argv) and jia_exit(): Initialize and shut down JIAJIA.

(2) jia_alloc(size): Allocate size bytes of shared memory.

(3) jia_lock(lockid) and jia_unlock(lockid): Acquire and release a lock specified by lockid.

(4) jia_barrier(): Performs a global barrier.

Besides, JIAJIA offers some subsidiary calls such as jia_Configure() to set optional optimization methods, jia_setcv(), jia_resetcv(), jia_waitcv() to provide the conditional variable synchronization method.

JIAJIA provides two global variables: jiapid to specify the host identification number and jiahosts to specify the total number of hosts of a parallel program. JIAJIA looks for a configuration file called .jiahosts in the directory where the application runs. This file contains a list of hosts to run the applications, one per line. The first line of .jiahosts should be the master on which the program is started.

JIAJIA allows the programmer to control the initial distribution of homes of shared locations. The basic shared memory allocation function in JIAJIA is jia_alloc3(size,block-size,starthost) which allocates size bytes cyclically across all hosts, each time blocksize bytes. The starthost parameter specifies the host from which the allocation starts. The simple call jia_alloc(size) equals jia_alloc3(size,Pageszie,0).

JIAJIA also provides some MPI-like message passing calls: jia_send(),jia_recv(),jia_bcast (),and jia_reduce().

3.2 Coherence Protocol

Figure 7 shows memory organization of JIAJIA. As shown, each shared pages has a home node and home of shared pages are distributed across all nodes. References to home pages hit locality, references to no-home pages cause these pages to be fetched from their home and cached cache locally. A cached page may be in one of three states: Invalid (INV), Read-Only (RO), and Read-Write (RW). When the number of locally cached pages is larger than the maximum number allowed, some aged cache pages must be replaced to its own home to make room for the new page. This allows JIAJIA to support shared memory that is large than physical memory of one machine.

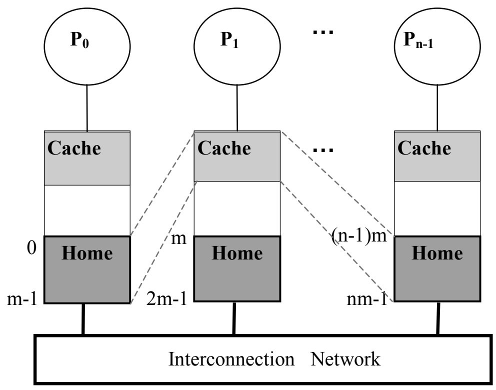


Figure 7. Memory organization of JIAJIA

JIAJIA implements the scope memory consistency model. Multiple writer technique is employed to reduce false sharing. In JIAJIA, the coherence of cached pages is maintained through requiring the lock-releasing processor to send to the lock write notices about modified pages in the associated critical section, the lock-acquiring processor to invalidate cached pages that are notified as obsolete by the associated write-notices in the lock. This protocol maintains coherence through write notices kept on the lock and consequently eliminates the requirement of directory. Some optimization methods of the protocol include single-writer detection [3], incarnation number technique [3], write vector technique [5], home migration [4], lazy home page write detections, and SMP optimization. These optimizations can reduce page faults, message amounts, and diffs dramatically and consequently improve performance significantly.

4. FAST PARALLEL IMAGE MATCHING ALGOR - ITHM ON THE SOFTERWARE DSM

Parallel program segments can be categorized into five patterns; single-process sequential, mutual-exclusive sequential, data-parallel, task-parallel, and common-parallel. The data-parallel pattern is used in parallelizing the sequential matching program with the API of JIAJIA. In our experiment, a cluster of eight Pentium II PCs connected by a 100Mbps switched Ethernet is used. The stereo images are divided into eight parts as showed in figure 8, each PC carries out the matching task of one parts of stereo image. For example, P_0 carries out the matching task of （Limg0,Rimg0,Φ），theΦ is the matching operator；P_1 carries out the matching task of （Limg1,Rimg1, Φ）；… P7 carries out the matching task of（Limg7,Rimg7, Φ）。P_0 ,P_1…….P_7 compute parallelly.

Data locality is crucial for performance in DSM machines, due to the difference in access times between local and remote memories. In image matching algorithm, high data locality can be reached by proper data dividing and proper task mapping to each computer. We put each part of stereo image data in the computer that used them, so that every computer can get the needed image data locally, in this way, massive communication operations are avoided. Even though the optimizing method above mentioned is used, at boundary, matching operator based on window have to fetch datum from the memory of other computers (remote memory). As shown in figure 8, in P_0, when the matching operator deals with the points at boundary, it needs the data in area FL0 and FR0, and the data in area FL0 and FR0 are in computer P_1 (FL1 and FR1), so that P_0 has to fetch the data in area FL0 and FR0 from P_1 . As shown in figure 7, for reducing this kind of remote fetching operations, in each computer partial memory is used as cache to keep the data fetched from remote memory. If the cache is big enough the total line of datum in P_1 that are needed by P0 can be fetched into the cache of P_0. In this way, at boundary, image matching operator only need one remote fetching operation, after this remote fetching, it can find the datum of area FL0 and FR0 in its own cache.

The main steps of the program are:

(1) Edit .jaconf, .jiahosts files to give the initial parameters.
(2) Call Jia_init(argc,argv) to create multithreads.
(3) Call Jia_alloc() to allocate memory for each thread.
(4) loop:
    Computer P_0~ P_n reads part of image separately,
    Call jia_barrier() to Synchronize.
(5) loop:
    Computer P_0~ P_n creates pyramid of part image separately,
    Call jia_barrier() to Synchronize.
(6) loop:
    Computer P_0~ P_n matches image separately,
    Call jia_barrier() to Synchronize.
(7) Output results.

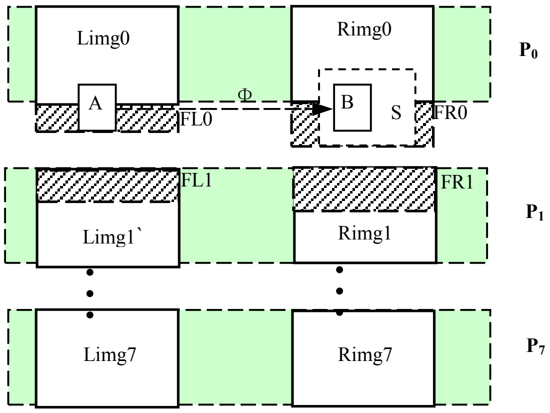


Figure 8. Division of image matching task

Satisfactory speedups are achieved for this application on a cluster of eight Pentium II PCs connected by a 100Mbps switched Ethernet. As showed in table 1,when two PCs are used, the speedup ratio is 1.8; when four PCs are used, the speedup ratio is 3.7; when eight PCs are used, the speedup ratio is 7.5.The speedup ratio is near the ideal linearity speedup ratio. The speed of finding corresponding points reaches 3200 pair/second, when eight PCs are used.

Table 1 depicts the speedup ratio of parallel image matching algorithm.

Number of computers	Two	Four	Seven	Eight
Speedup ratio	1.8	3.7	6.4	7.5

Table 1. Speedup ratio of parallel image matching algorithm

5. CONCLUSIONS

In this paper, we have presented a novel operator of image matching that imitates the distributing of cone photoreceptors in retina of man and its fast parallel algorithm based on the softer ware DSM.

The advantage of suggested image matching algorithm base no-uniform sampling is that this mechanism not only can centralize the resource of computation on the center part of match-window and use more wide scope image information, but also reduces the influence of geometric distortion on image match, which is the difference between relief distortions of two stereoscopic images in essence.

From the experimental results, we have observed that it is important to pay great attention to the development of cluster system. A well-designed cluster system can provide highly available access to applications, system resources and data, together with other features such as scalability and single system image. There is a great demand of large memory size and more high calculation speed in the field of photogrametry and remote sensing, cluster system can meet this demand at low cost, especially in processing and analysing remote sensing image of great size and establishment of image data bank.

6. REFERENCES

[1] Bershad,B., 1993f. The Midway Distributed Shared Memory System. *In Proc. of the 38th IEEE Int'l CompCon Conf.*, pp. 528-537.

[2]Carter, J., 1991o. Implementation and Performance of Munin. *In Proc. Of the 13th Sym. on Operating Systems Principles,* pp.152-164.

[3] Hu, W., 1999a. Reducing System Overhead in Home-Based Software DSMs. *In Proc. Of the 13th Int'l Parallel Processing Symp.,* pp. 167-173.

[4] Hu, W., 1999j. Home Migration in Home-Based Software DSMs. *In Proc. of the 1st Workshop on Software Distributed Shared Memory,* pp. 21-26.

[5]Hu,W., 1999j. Reducing Message Overheads in Home-Based Software DSMs. *In Proc. of the 1st Workshop on Software Distributed Shared Memory,* pp. 7-11.

[6] Iftode, L., 1998a. *Home-based Shared Virtual Memory*. Ph. D. Thesis, Princeton University.

[7]Poggio,T., 1985s. Computational vision and regularization theory. *Nature,*317(26),p.324-319.

[8] Qiu,Z., 2001o. *The Theory/Algorithms and Realization of Stereo Vision*. Ph.D.Thesis, Zhengzhou Institute of Surveying and Mapping.