The race between data production and processing capacity has been going on for many decades, with data production usually on the winning team. This is also true for airborne and space-borne imagery, as the amount of images captured by satellite sensors and aerial cameras is growing not only steadily but also rapidly. How can the abundance of pixels be processed into photogrammetric products quickly and effectively? The answer lies in parallel computing. Today, computer clusters enable fast and affordable processing of photogrammetric tasks. Read on to learn how parallelism speeds up the creation of seamless orthomosaics.

By Andrey Yu. Sechin, Racurs, Russia

Traditionally, software has been written for serial computation. The algorithm is put into operation as a series of instructions which are executed on a central processing unit (CPU), one instruction at a time in succession. Parallel computing – a dominant research area in computer architecture aimed at speeding up computation – is mainly implemented through multi-core processors. The use of multiple CPUs enables many calculations to be conducted simultaneously. As a result, complex computational tasks are broken down into smaller components which can be processed at the same time. Each CPU executes its part of the process simultaneously with and independently of the others. The results are combined afterwards. Photogrammetric processing of massive volumes of images may also benefit from parallel computing. To illustrate the massiveness of the amount of data produced by spatial and airborne sensors, European Pleiades-1D and Pleiades-1B satellites have the capacity to acquire 2,000,000 km² per day, while the VisionMap A3 Edge aerial camera captures 5,000 km²/hour of imagery with a GSD of 20cm.

Seamless Mosaics

The creation of seamless orthomosaics consists of several steps, including (Figure 1):

1. project creation and reading the images from storage devices
2. ortho creation of each image using the corresponding digital terrain model (DTM)
3. determination of seamlines
4. image statistic gathering for brightness adjustment
Photogrammetric algorithms can be effectively run on computer clusters with 100 to 200 CPUs. A fast data retrieval and storage device by splitting huge amounts of data into smaller parts and pushing these parts to different HDDs simultaneously. How the data is distributed across the devices – usually hard disk drives (HDDs) – depends on (1) the level of redundancy required to avoid loss of data due to sector read errors or disk failures (reliability); (2) speed of reading and writing (performance); and (3) available capacity. An extension number from 0 onwards indicates the balance between these goals. Two principles underlie a RAID: stripping and mirroring. Stripping allows writing in parallel to a storage device by splitting huge amounts of data into smaller parts and pushing these parts to different HDDs simultaneously. Mirroring is storing the same data on several HDDs to improve reliability of data storage and also to speed up reading, as the data can be read from different HDDs simultaneously. Typical RAIDs have 300MB/s sequential access throughput and do not constrain CPU usage.

Random Access

The data needs to be transported from the RAIDs to the CPUs and back through a local area network (LAN). Therefore, the reading and writing times also depend on the LAN throughput. The IO time is roughly linearly related to the array speed, i.e. a 300MB/s array is around three times faster than a 100MB/s array based on the same LAN speed. The authors’ experiments show, however, that a 600MB/s array connected to 4 GB/s LAN is only two times faster than a 300MB/s array connected to a 4 GB/s LAN. This is also true for a 100 MB/s array connected to a 2 GB/s LAN. This configuration is four times faster than a 20MB/s to 30MB/s array connected to a 1 GB/s LAN, although one would expect performance to be eight times better. Hence, the array and LAN speed alone do not tell the whole story. The reason is that some photogrammetric algorithms require random data access for which modern HDD-based RAID storage systems will give low throughput. Furthermore, the experiments show that using solid state drives (SDDs) instead of HDDs improves the reading performance even when only stripping is used.

Concluding Remarks

Photogrammetric algorithms can be effectively run on computer clusters with 100 to 200 CPUs. A fast data retrieval and storage...
system and high LAN throughput ensure the highest productivity.

**Biography of the Author**

Andrey Sechin graduated from Moscow Institute of Physics and Technology, Russia, in 1980 and obtained a PhD degree in mathematics in 1984. He is co-founder of Racurs, a photogrammetric company based in Moscow, where he has been scientific director since 1994. Before founding Racurs he was with Troitsk Institute for Innovation & Fusion Research.

Email: sechin@racurs.ru

**Figure captions**

*Figure 1*, Flowdiagram of the orthomosaic creation process.

*Figure 2*, Block of around 200 GeoEye images on a background representing the DTM and detail (bottom).

*Figure 3*, Speed-up using multiple CPUs: $T_1$ and $T_p$ are computing time on one and $p$ CPUs, respectively; $a$ is a nonparallel ratio.

*Figure 4*, GeoEye orthomosaic – over 600 sheets have been produced at scale 1:10,000.