Concepts and Applications of Digital Orthophotos

By Robert A. Fowler, OLS, CST, CET

The following paper was presented at the AOLS Geodetic Picnic on September 28, 1996.

The title doesn't sound very exciting, but it is, because where I work, we think this technology is going to totally "change the way you look at mapping." So much so we have adopted that slogan in our advertising.

Digital orthophotography is a technology that has been around for a few years now (at Intermap we've been doing it for more than six years) but has not had the reception in Canada that it has in other parts of North America and around the world. And this is rather strange, given that Canada has always considered itself at the forefront of Geographic Information Systems and digital map data development. As a personal observation, I wonder if this is actually one of the reasons - because so much energy has been spent on structuring data in conventional vector systems, that there is some resistance to a new technology which, in many cases, throws out so much of the work done in this regard. But I will return to that a little later.

As the title says, we will look at the fundamental background to digital orthophoto production and then review applications for which it can be used.

So what exactly is a digital orthophoto?

In short, it is a raster image file in a format that can be used by computer systems with the capability of handling images. A raster is simply a grid of information, in this instance a grey scale. So a raster image is made up of a number of rows and columns of pixels which have an intensity value assigned to them. The usual values vary from 0 which is black through 256 shades of grey to 255 which is white.

The images do not have to be simply black and white, they can also be in colour. However, that means that each colour image is made up of a composite of three raster files with intensity values for each of the primary projection colours: red, blue and green.

The digital orthophoto raster image constitutes a fully rectified base map with x and y (or east and north) coordinates for each pixel. This mean that if we measure distances, areas or angles on the image they will be correct. In addition, there is a digital terrain model which comes with the territory.

The next problem is to rectify and georeference the image. This is where the mathematics become a little tricky.

The next question is how is a digital orthophoto made?

It is made from a combination of a scanned aerial photograph, which is then processed using a number of other bits of information.

Taking each in order. The aerial photograph is simply a photograph taken using one of the standard metric aerial cameras. That is to say a camera which has been calibrated for mapping measurement purposes. These are almost always 9" x 9" format negatives and 152mm focal length cameras, with extremely fine resolving power of the lenses. All of the new cameras have motion compensation devices which compensate for image blurring along the line of flight at low levels. This does have some significant quality ramifications for orthophoto production at large scales.

The aerial photo is scanned using a precision scanner. A number of different types of precise scanners are around. The reason I say precise is because, there are numerous document scanners (including fax machines) which are scanners but offer a very coarse resolution. To provide the eye with the illusion it is still seeing a real photograph rather than a coarse series of dots or more correctly pixels, the image or aerial photograph must be broken up into extremely small squares (picture elements - hence pixels) each of identical size and shape and at a specific point in the raster grid. A precise scanner does this.

The next problem is to rectify and georeference the image. This is where the mathematics become a little tricky.

An aerial photograph contains a number of distortions due to the lens, the cameras attitude and the shape of the earth. Briefly, because of the shape of the lens, there is radial distortion on every photograph. This is simply a function of the length of distance light travels from the lens to the film. The effect of this is a change in scale the further away from the centre of the photograph a point in the image is located.

Second, the rotation of the camera (aircraft) in relation to the earth creates additional distortions. Finally the shape of the earth adds additional distortions.

Just looking at a single photograph, of course, it is impossible to ascertain which distortions are where and how much they are affecting the accuracy of the position of any detail.

However, by adding other information we can systematically eliminate each of these distortions. The first requirement is to obtain the camera calibration report. Distortions due to lens aberrations and flatness of the camera platen are removed as these tested knowns are factored into the equation.

The second requirement is to produce an aerial triangulation of the block, (assuming more than one photographic model is involved). The aerial triangulation process extends the ground control so that any model in a block can be set up at any time. The aerial triangulation data is effectively the control elements for the orientation of each photograph in space in relation to the earth. In other words, it provides us with a specific level and scale for each photo.

Will the system choose the grey scale of the scanned pixel that covers the largest part of the pixel it must fit in - or some combination of pixels?

Finally, to remove the distortions due to changes in the terrain, we need to add a digital terrain model. A DTM is a collection of points to indicate the changes of ground elevation, which is usually collected using standard photogrammetric mapping techniques. DTMs are usually collected across the photogrammetric model by the photogrammetrist reading profiles spaced at specific intervals apart. (There are other ways of providing a DTM, but for the purposes of this paper we will ignore them.)

In addition to the DTM grid, the operator collects break lines and spot elevations. A break line is collected to show a change in the nature of the land. For example, the break line could be a stream. Streams are standard break line features as they occur at the lowest point in a valley - indicating that the land slopes down one way, reaches the bottom, then slopes upwards. Spot elevations generally indicate the top of a hill or the bottom of a depression, or simply help fill in an area which is relatively flat.

Once the digital terrain model is collected, it is processed to a very dense grid. The digital orthophoto algorithms then determine the closest elevation for a given pixel and apply that correction.

However, the actual rectification process proceeds in the reverse of what you may think. The design of any digital orthophoto product starts with the final output. Once that is established, the design of the project process is decided. For example, once we know the size of the output pixel (i.e. what each pixel is to represent) and the size of the tile or map sheet, then everything follows.

For example, if we know our pixel is to represent 0.2 metres and the map sheet will cover an area 500 metres square, then our raster for that tile/file/sheet will be 2500 pixels by 2500 pixels. In other words we have to fill a block of empty pixels each of that size, 0.2 metres, with grey scale information. The corners of our tile are given geographic coordinates and the system knows to systematically look for the grey scale value to fit each pixel in turn.

One of the problems is that the pixel to be filled, in all likelihood, will not line up exactly with an existing pixel from the scanned image. Depending on the complexity of the software, a choice then has to be made. Will the system choose the grey scale of the scanned pixel that covers the largest part of the pixel it must fit in - or some combination of pixels?

If it the former, it is known as resampling by nearest neighbour. If the software is slightly more complex, it would choose to do this by a process known as bi-linear resampling.

A bi-linear system picks the pixel that most covers the area of its new location, plus the value of the pixel each side in both directions (i.e. north, south, east , west) and then makes some sort of average based on this cross of five pixels.

Cubic convolution is the most complex algorithm, whereby the system picks the pixel most covering the new location and then produces an averaging based on percent coverages from the block of nine pixels on the aerial photo closest to immediately surrounding the pixel on the final raster it must fill. Cubic convolution provides the optimum result, although the difference between bi-linear and cubic convolution is not as significant as the difference to either of these and nearest neighbour, which can produce much more blockiness in the image. Once a raster is filled, it is an orthorectified image.

In most cases, the area to be covered, however, requires more than one photograph. The problem then is to mosaic two or more photos together.

So, if the images are to be mosaicked, the mosaicking software joins the images in a way that is least offensive. You could simply butt join one image to the next, and in theory they should join precisely. However, the lean of above ground features such as buildings on the edge of one photo will not match the lean from the adjacent photo, causing an image discontinuity which is very offensive to most people. (Usually the more you look at it the more unbearable the join is.) While not all ortho systems can mosaic effectively, the best solution is for an operator to interactively decide a seam line along a feature or through an area where it is least likely to be offensive.

Finally the image is converted to its destination file type. A large number of image file formats are around. There are also unfortunately several versions of the ubiquitous TIF files. Most GIS systems have a delivery file format, and most of these are based on one of the big three: ARC/INFO, Intergraph or AutoCAD. (An ARC/INFO file is a .BIL file, and Intergraph file is a .COT and an AutoCAD file is an .IGS file) AutoCAD requires an ad-on package called CAD Overlay. There are three versions of this and the lowest is only for documents. So if you are contemplating using CAD Overlay you will need, at minimum, the GS version. Compatibility among raster formats is gradually settling down and, in the future, it is hoped there will be more standardization in this regard.

Delivery these days is most often on CD-ROM. For while other delivery media are available (and we use them for specific clients) CD-ROM seems to be the most convenient and cost effective at the present time.

So that's a brief look at how digital orthophotos are made. The next topic is why would you want to use one.

The answers to this, of course, depend on what you hope to achieve. However, the fundamental reason is orthos make it generally easier to appreciate what is on the ground - simply because most people can relate to an aerial photograph.

A digital orthophoto is a fully corrected image, which means you can use it for a map. In other words, anything you can use a map for, you can use a digital orthophoto.

A digital orthophoto has a couple of fundamental considerations which are important. The first is, depending on the area, it is often less expensive than a conventional vector map of the same area. Next, it is flexible, because most maps are interpretations of what someone else thinks you need to know. The digital orthophoto has everything on it that was on the original air photo. Third, it is very easily shared. Now, you may argue so is a line map. However, often the public works line map showing roads and sewers is not suitable for the parks department which wants to see trees and flower beds. Fourth, it will allow incremental addition or integration of vector data. For example, if you need maps in a rush, but don't have the time to wait for vector data collection, which let's say in a major city project could take months or years, a digital ortho project will provide all of the base information very quickly, and you can collect the vectors as and when needed. Fifth, a very interesting exercise, it can be used to verify the accuracy and completeness of vector data sets!

Last, it is relatively inexpensive to maintain and update. For example, if there is a portion of the digital orthophoto coverage which has changed, the updating is a matter of obtaining a new photo, scanning it at the same resolution, controlling the scan with a new aerial triangulation, and then if there have been no major changes in terrain slope, simply using the old DTM to correct the image. The new portion can then be "fenced" and dropped on top the old data which is then blown away.

Now, someone is going to say what about the attribute information I had recorded for that area?

Well the first thing is to realize that a raster has no topology. The question then is, "why do we need topology?"

And there are good reasons for topology, however, it should be recognized a lot of topology is there to simplify vector data use. If you don't have vector data, you don't need as much topology.

However, you do have to recognize that there will be applications which call for topology and if that's the case, then a digital orthophoto is not the full answer. However, I might suggest, in the general scheme of things these are not many. What there are often requirements for are road linkages and attributes. With respect to the former, road centrelines can be digitized photogrammetrically or on the screen. At Intermap, we recognize that in most cases (unless the client specifies otherwise) road centrelines are collected as a break line in the DTM collection. So these already exist as part of the DTM. We code them separately, so they can be extracted from the DTM and used by the client in other ways. In the latter case, attributes can be assigned by creating a linkage point or centroid. Now some software packages, such as ARC/INFO will create centroids automatically, once all of the polygons are closed. However the database usually has to be populated by some sort input. Linkages have to be made through the centroid to the database. So if you are typing in a whole bunch of information on 32 Elm Street, it takes only a second or two to put the computer cursor somewhere near the centre of house number 32 on the digital orthophoto, and create the point which allows the linkage to be made to the tabluar database in which you are typing the attribute information. As the point data is usually created on a separate level or coverage (depending on what software you are using), when you subsequently replace the image during an update, the centroid stays there because it is on another level (not in the raster file).

Applications? These are as varied as map users. We have created digital orthophotos for people managing landfill sites, as base maps for municipalities and for state and national governments.

A couple of cost indications may surprise you.

For example, to produce digital orthophoto quarter quad sheets for the US Geological Survey, the standard small scale map sold by the US Government, the cost of producing the orthophoto is US$300. Compare that to say $4,000-$7,000 for a 1:10,000 scale OBM sheet. Now the $300 does not include aerial triangulation and DTM collection, but its still a significant difference over the time and cost to produce a vector map.

Looking at a large scale application, in the city of Boston, 1:1,200 scale digital orthophotos (in colour) are being produced for about US$475. A similar map sheet produced as a vector map would cost $2,500 (plus the photo control and aerial triangulation, of course).

Applications? These are as varied as map users.

Where the digital orthophoto is less competitive is in rural areas, where the cost of producing a digital orthophoto remains the same, but the cost of vector mapping decreases because there is less detail to map.

The other disadvantage until fairly recently was the production of hard copies. To obtain hard copies, most engineering companies or municipal departments did not have quarter of a million dollar plotters to reproduce raster images. However, over the last couple of years, a number of manufacturers have introduced ink jet plotters which can plot raster files (in black and white and colour) and produce a pretty good plot at very reasonable outlays. The sample we have here was made by Hewlett Packard Design Jet 755 plotter (not brand new) which we purchased for $12,000. (Not a huge outlay when you consider the costs of computer equipment, software and data services to set up a municipal GIS. And the plotter can be used for other types of plot.)

If you have a project which you think may benefit from this technology, call someone with the service and they will answer your questions. And if you haven't looked at what's happening in the digital world recently, its probably time you changed the way you look at mapping too!

Robert Fowler is manager of marketing for Intermap Technologies (see ad this issue), as well as North American correspondent for Surveying World (U.K), and editor of Northpoint Magazine.