A solution for topographical map updating

by Samantha Goodchild-Brown and Zoltán Szecsei, Geograph

Production challenges and solutions used in updating data for topographical mapping.

Cape Town-based spatial information data capture company, Geograph, recently won tenders from the Department of Land Affairs for topographical map updating. This is work not commonly undertaken in the private sector and therefore generates much interest particularly due to the variety of features being captured.

The objective

- A list of 192 features need to be identified from orthorectified imagery for topographical mapping and GIS database capture – not all of these features occur on each image or map.
- To capture all the information topographically and topologically correct, in as automated a fashion as possible.
- To achieve client specifications with minimum costs, i.e. minimum skills and experience levels to get the job done, in minimum time.

Software

Our GIS software is Linux based and the Ubuntu variant of Linux was chosen due to its well-maintained releases. Linux is a true multi-user, multi-tasking operating system and this enabled us to configure our PCs with three screens and three keyboard/mouse combinations, allowing three operators to simultaneously work off one PC box with no noticeable effect on computer response speeds.

The proprietary GIS package used is fully customisable and well over 10 000 lines of custom code were written in-house, just for the main feature capture script. We specifically chose to make the operator menu available on the keyboard, and not to have drop-down menus and icons for choosing the various functions, features or routines.

The reason for this is that we believe that the need to move the cursor and refocus the eye to find all of these drop-down menus will slow down the operator. With our system, the operator does not take the mouse off the feature they are working with, and only briefly glances to the keyboard from time to time, to check on the position of his or her fingers. We have used this methodology for over 15 years and time and again it has proven extremely productive.

The software used was not without its weak spots. Like all software, it is not about the weak spots, but about knowing them and having suitable methods of dealing with them. This software has a known performance weakness with polygons; therefore the decision was made to only capture data using lines and points, and to automatically build the polygons at time of export.

Data management

All data is stored on a central server, enabling easier management and backup processes.

It was decided to make all 192 features available to the operators as feature codes (numerical rather than alphabetic). Remembering all the numbers might seem a bit daunting, but we found the operators adapted very quickly. There is a group of about ten more common features which were easy to memorise, and it became a competition in the office to see who could remember the more obscure and rare feature types e.g. siphon, shipwreck, ground sign, cave and zoo.

The topology issues needed to be automated as much as possible in order to minimise human error. To this end our main feature-capture script references a table we compiled. This table specifies the actions and relationships between features e.g. 'snap', 'move away', 'ignore' or 'log for operator review'. It is automatically activated when nodes, vertices or edges are placed by the operator within a specific distance of other nodes, vertices or edges.

This snap-routine is also re-run later as a post-process, and a log is made of all combinations where human intervention is required for resolution. Where any combination is not found in the table, it is also spatially logged (as 'undefined', for addition to the snap-table).

Operators with suitable experience and capability then go to each record and accept or modify accordingly. This should never be done with only the captured line work visible, but rather with the imagery visible so that a more comprehensive quality check can be done in these areas.

The GIS database we create is seamless and continuous, but as our source data, capture units and delivery units are map sheet based, edge matching needs to be considered. This sheet-based imposition also creates the need for polygons to be "closed" per sheet and not per feature. For example if a "sandy area" overlaps a sheet corner (and hence overlaps four map sheets), the feature needs to be cut into four polygons, the cut-lines being the map sheet boundaries.

The main aim of the export routine is to reformat our data structures to that of our clients' needs. We do however also use this export routine to build and check the above complex polygon and topological rules as specified by our clients. Creating our final polygon and other data structures at export time minimises the possibility of an operator introducing topology errors, and maximises the topological correctness of our deliverables.

Fig. 1: Raster image of Lainsburg with colour photography.

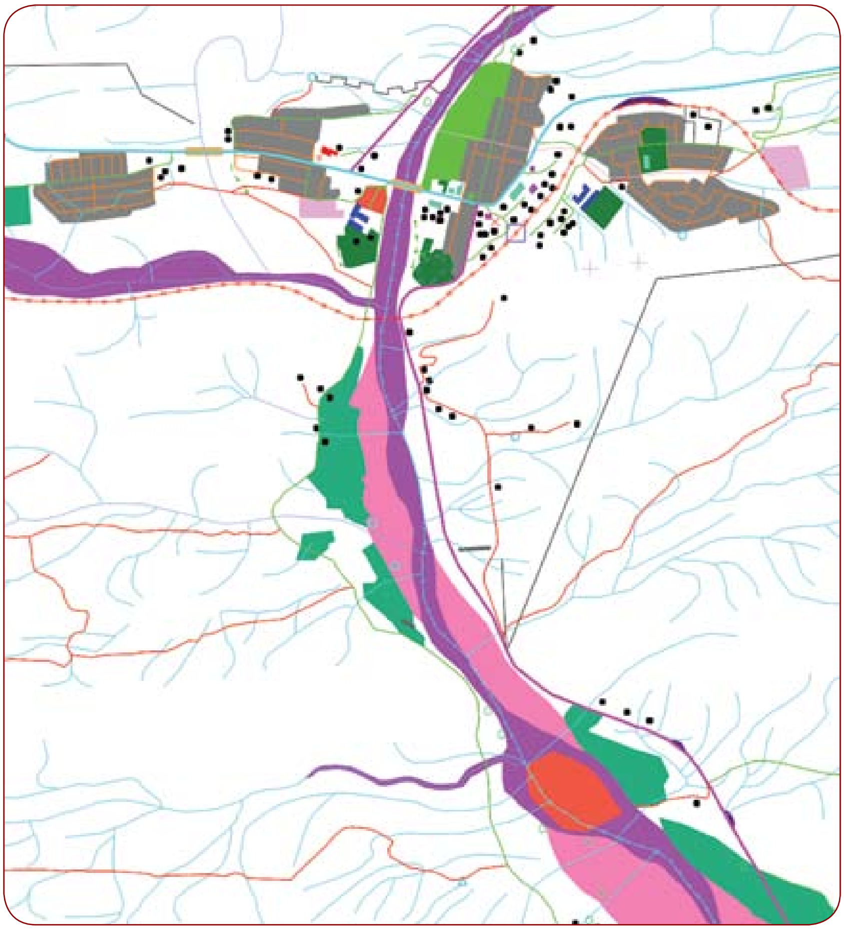

Fig. 2: Vector image of Lainsburg – a work in progress.

Project management strategies

As experience shows, software is rarely the ultimate answer to anything but is rather complimentary to other systems. The software solutions above could never work without a suitably designed 'human production system'.

We recognise that:

- Cost constraint and the specialised nature of the work precluded the use of well-experienced people for the bulk capture phase
- This sort of capture work requires a long training period (photo interpretation skills)
- Not everyone is capable
- Different people need different training methods
- Learning speeds vary greatly

Our solution thus far, for the processes and management of these issues is outlined below.

Break the production line up so that no one person does everything. This means that people can be trained in stages and this minimises the amount of time trainers are occupied before you can get some useful production out of the trainees.

Select a subset of features that can be done at a simpler level of understanding, where feature relationships are minimal and only limited image interpretation is required. This subset then becomes the group of features that inexperienced operators begin with. This enables them to learn the main package functionality and limits the rest to sizeable bites.

Once they have shown that they have achieved full competency at this level, they can be moved into the group that does the rest of the feature capture (general capture).

After a suitable degree of experience has been gained in general capture, and productivity shows they are comfortable in this function, they can be trained to the function of topology checking. This group would also be concerned with exporting of the data and resolving any issues that may arise. The export process is a batch job, typically run overnight, that re-does full topology and data consistency checks. Where open polygons, kickbacks in vectors, and other anomalies are found, they are logged to a file. Operators will use this log file to visit and repair all issues prior to the creation of our final deliverables.

The capture of rivers is handled differently to that of other features, as special skills are required. The existing rivers and contours dataset is provided to us by our client, for update (rather than re-capture) purposes. River update is done in conjunction with contours. We have found that many people think this is easy, resulting in underestimation of the work, which results in failures where people miss too much, or create more errors in the dataset. e.g. drawing in rivers over watersheds.

Edits in the contour dataset are only done where the contours do not fit the image. For minor problems, like contours turning on a river, or crossing a dam high-water mark, we will draw in a best fit for map logic purposes only. Where major contour problems are found, an investigation needs to be done to decide on a method to proceed (see Fig. 3). The method chosen will depend on the severity of the problem being faced, and might be anything from shifting the entire map sheet by a constant, to sourcing a digital elevation model (DEM) to generate more accurate contours.

We have found that breaking the work into smaller bites like this means that each person's work can be completed faster than if they were doing all the features on one map sheet themselves. This has a psychological benefit in that they can complete a task within a few hours or a few days, instead of being busy with one thing for a longer period. The achievement of these short term goals has the effect of keeping up morale and ensuring greater overall productivity.

Quality assurance

Most crucial to this system is the quality control that happens at each stage. At the initial stage this process is to check the individual operator's abilities and understanding of the specifications. This first quality control (QC) step is done digitally. This ensures that each operator's learning curve and motivation is as steep as possible, making them productive in the workplace as soon as possible. This also allows us to ascertain if the person is suited to this type of work or not, sooner rather than later. A clear sign they are failing is when they are consistently failing to follow specifications after repeated report backs of errors. Also, when they cannot maintain the required level of work as soon as the trainer moves away.

Feedback is given to the individual operators at all times. This is done via team leaders who get reports on individual's errors from the QC team, which they then feed back as part of the training. It is also done with one-on-one sessions directly between the operators and the QC person.

Once the general features have been captured, a plot is made and compared to the hardcopy photography supplied by the client. This might sound stupid to many people who feel that the digital environment should be paperless, however, due to limitations caused by the digital environment such as screen resolution, we have found this to be quite effective at spotting blunders of misclassification, specifications and omissions. In certain areas, such as urban areas, this is not an advantage and digital checks must be done. The volume of data to analyse, and the space for writing comments in, makes this imperative.

At no stage can it be assumed that an operator is experienced enough to not need quality control. When moving to a new part of the country, it is vital to approach the job as if operators have no experience.

As operators tend to focus their attention on the centre of the dataset they are required to capture, they sometimes inadvertently neglect the edges. For this reason, visual edge matching is manually performed as part of the quality control. Here all the data as well as the imagery is viewed together per map sheet. In this way, one not only edge matches but also

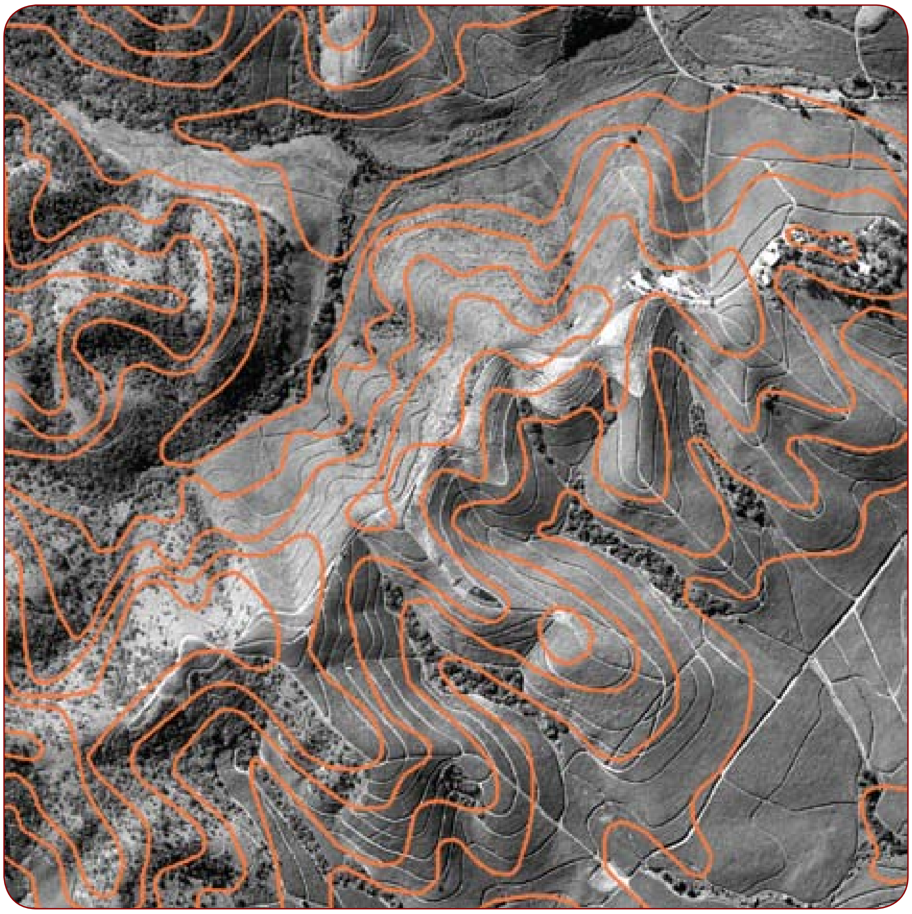

Fig. 3: An example of contours out of specifications and requiring investigation.

Fig. 4: Panchromatic imagery with a variety of feature classes.

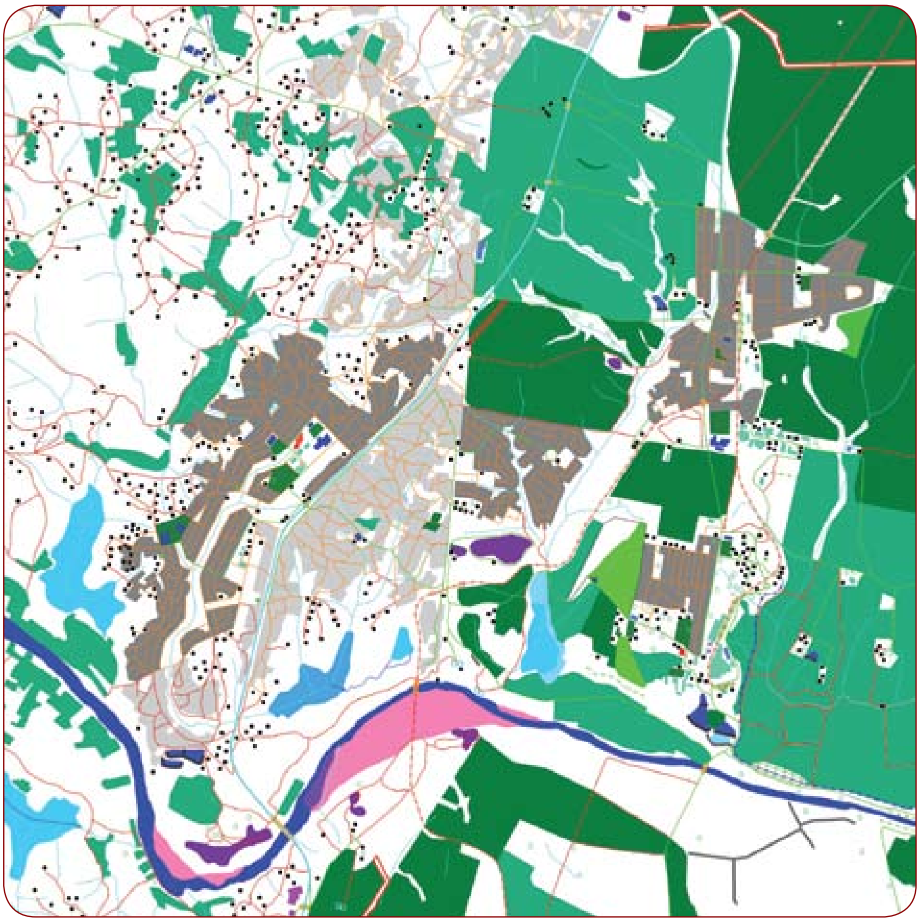

Fig. 5: The resultant vector dataset – contours are not included.

identifies more omissions and ensures more consistency across the dataset. Only a select few operators can do this work so that no new errors are introduced.

Moderating is done in conjunction with edge matching. Here one will review all the sheets together and compare the data coverage between the sheets. We ensure that where a change of density of a feature type exists, it does not occur on a sheet boundary but rather on a natural or political boundary. We would also check for human blunders like accidental misclassifications e.g. one sheet having a different class of railway line to that of the adjoining sheet.

The formal QC procedures do not operate independently. By noting the nature of the problems found and feeding them back into the production line, we have also been able to improve our electronic processes. The less you find at the moderating stage, the more you know you have succeeded in the previous management strategies.

The future

As we do more work of this nature in the future, the reality of receiving only digital imagery will necessitate a change to the current strategy. Hard-copy photos will no longer be feasible, meaning that some of our quality control strategies will need to be redeveloped. Using a stereoscope for 3D problem solving will also no longer be an option and thus a digital solution must be chosen.

While everyone would agree that there are many ways to carry out this task, everyone would also agree that image interpretation is not without serious challenges. Our solution is one way of safely taking inexperienced people and making them productive within an extremely short space of time. What is vital, is to have a strong core team that can work together to create all the software and hardware, QA and training capacities that are required for high volume map compilation.

Contact Samantha Goodchild-Brown, Geograph, Tel 021 788-4897, samanthag@geograph.co.za