

SYSTEM ARCHITECTURE FOR INTEGRATING GIS AND PHOTOGRAMMETRIC DATA ACQUISITION

P. A. Woodsford

Laser-Scan, Cavendish House, Cambridge Business Park, Cambridge, CB4 0WZ and Dept. of Geomatic Engineering, University College, Gower Street, London, WC1E 6BT, UK – email: consultancy.woodsford@ntlworld.com

ICWG II/IV

KEY WORDS: Photogrammetry, GIS, Databases, Integration, Interoperability, Updating, Accuracy, Three-dimensional

ABSTRACT:

The paper surveys the reasons behind the trend towards the increasing integration of photogrammetric data acquisition and GIS databases, describing both the drivers and the resulting benefits. The shift from primary data acquisition to the maintenance and/or enhancement of increasingly rich datasets is discussed, and the assertion is made that this should increasingly be approached as a conflation issue. Different levels of system architecture are enumerated and characterised, including file-based information exchange, loose- and close-coupling (with differing balance of functionality between client and server-side) and multi-tier architectures. The role of the database in supporting transaction management is discussed, as is the importance of interoperability between all the maintenance, quality assurance, analysis and delivery processes either within the firewall or across a distributed web-based system. A number of key use cases over and above the standard update or revision task are examined, including positional accuracy improvement, the maintenance of public persistent feature identifiers and the maintenance of topological integrity. The level of support available for 3D information in geospatial databases is reviewed together with the issues of migration from 2D to 3D. The level of standards support, or lack thereof, is also discussed. The paper concludes with some thoughts on the role of the photogrammetrist, image analyst or geomatician of the future.

RÉSUMÉ

Ce papier examine les raisons derrière la tendance vers l'intégration croissante de l'acquisition de données et des bases de données photogrammétriques de SIG, décrivant les forces d'entraînement et les avantages résultants. Le décalage de l'acquisition de données primaire à l'entretien et/ou du perfectionnement des ensembles de données de plus en plus riches est discuté, et l'affirmation est faite que ceci devrait être approché comme issue de conflation. Différents niveaux d'architecture de système sont énumérés et caractérisés, y compris l'échange de l'information dossier-basé, lâchement et étroitement lié (avec l'équilibre différent de la fonctionnalité entre le client et le serveur-côté) et les architectures à plusieurs niveaux. Le rôle de la base de données dans la gestion de support de transaction est discuté, de même que l'importance de l'interopérabilité entre tous les entretien, garantie de la qualité, analyse et processus de la livraison dans le firewall ou à travers un système web réparti. Un certain nombre de cas principaux d'utilisation au delà de la mise à jour ou de la tâche standard de révision sont examinés, y compris l'amélioration de position d'exactitude, l'entretien des marques persistantes publiques de dispositif et l'entretien de l'intégrité topologique. Le niveau de l'appui disponible pour l'information 3D dans les bases de données géospatiales est passé en revue ainsi que les questions de la migration de la 2D à 3D. Le niveau de l'appui de normes, ou le manque d'appui, est également discuté. Le papier conclut avec quelques pensées sur le rôle du photogrammetrist, l'analyste d'image ou géomatician du futur.

KURZFASSUNG

Der Artikel befasst sich mit den Gründen der Tendenz zu einer zunehmenden Integration von Photogrammetrischer Datenerfassung und GIS Datenbanken. Dabei werden sowohl die Motivation als auch der resultierende Nutzen beschrieben. Die Verschiebung von primärer Datenerfassung zur Wartung und/oder Erweiterung zunehmend größerer Datensätze wird diskutiert, und es wird die Behauptung aufgestellt, dass dies im zunehmenden Maße als Konflationsproblem behandelt werden sollte. Es werden verschiedene Level einer Systemarchitektur aufgezählt und charakterisiert, einschließlich Dateibasierter Informationsaustausch, lose und feste Kopplung (mit verschiedenen Gewichtungen der Funktionalität zwischen Client und Server) und mehrschichtigen Architekturen. Die Rolle der Datenbank bei der Unterstützung des Transaktions-Managements wird diskutiert, sowie über die Bedeutung der Interoperabilität zwischen Wartungs-, Qualitätssicherungs-, Analyse- und Auslieferungsprozessen entweder innerhalb eines Firewalls oder über eine verteiltes Web-basiertes System. Es werden einige Schlüssel-Anwendungsfälle und die üblichen Aktualisierungs- und Revisionsvorgänge überprüft, einschließlich Verbesserung der Positionsgenauigkeit, Wartung der öffentlichen persistenten Feature-Identifizierer und die Wartung der topologischen Integrität. Der Stand der verfügbaren Unterstützung für 3D-Information in Datenbanken mit Rauminformation wird zusammen mit den Möglichkeiten zur Migration von 2D nach 3D untersucht. Der Stand der Unterstützung von Standards oder deren Fehlen wird ebenfalls diskutiert. Der Artikel schließt mit einigen Gedanken über die Rolle des Photogrammeters, Bildanalysten oder Geoinformatikers in der Zukunft.

1. INTRODUCTION

1.1 Rationale of Paper

The start point for this invited paper can be summarised by the following extract from the ISPRS Annual Report 2003 (ISPRS 2003) covering the work of Intercommission WG II/IV: Automated Geo-Spatial Data Production And Updating From Imagery:

“Concerning updating of GIS, besides the data capture itself the management of the updating information in the database is a relevant topic itself. Automatic update including topological changes still is nearly unsolved. In some cases from the operational point of view the acquisition of the complete data set still seems to be easier than to incorporate acquired changes into an existing data set. Digital photogrammetric workstations more and more approach a GIS leading to integrated solutions which cover the complete process from data capture to data management, analysis, visualisation and dissemination. At present they are incorporating database and visualization functionalities, partly in 3D. In general the cooperation and exchange between GIS and Photogrammetry still is rather low, especially from the commercial point of view. There still is a lack in standardized exchange between the respective systems but the companies seem to have recognized the lack in integrated solutions for the end-user.”

The importance of the shift in requirement, from the acquisition of complete data sets to the incorporation of acquired changes into an existing data set was highlighted for the author by a remark at the OEEPE/ISPRS Workshop “From 2D to 3D – Establishment and Maintenance of National Core Geospatial Databases”, Hannover, Germany, October 2001. In describing the history of the Topographic Database of Catalonia, 1:5,000 (Pla et al, 2001) Josep Lluís Colomer said that in 1996 they had done something ‘they would never be allowed to do again – they had abandoned the previous data set and started again’. This represents the new reality for most providers of topographic data. The investment in the current data holding, and its increasing richness in attribution and structure as well as topographic detail, means that there is in practice no alternative but to update (or revise) it, to enrich it and on very rare occasions to re-engineer it. The only exception to this trend is the case of ‘mission-specific’ datasets, which may be of sufficient importance to bear the cost of one-off creation from source, but even in this case the trend is towards intensification of sustained and maintained ‘framework’ data.

1.2 Structure of Paper

Section 2 surveys the reasons behind the trend towards the increasing integration of photogrammetric data acquisition and GIS databases, describing both the drivers and the resulting benefits. The shift from primary data acquisition to the maintenance and/or enhancement of increasingly rich datasets is discussed, and the assertion is made that this should increasingly be approached as a conflation issue.

Section 3 enumerates and characterises different levels of system architecture, including file-based information exchange, loose- and close-coupling (with differing balance of functionality between client and server-side) and multi-tier architectures. The role of the database in supporting

transaction management is discussed, as is the importance of interoperability between all the maintenance, quality assurance, analysis and delivery processes either within the firewall or across a distributed web-based system.

Section 4 covers a number of key use cases over and above the standard update or revision task, including positional accuracy improvement, the maintenance of public persistent feature identifiers and the maintenance of topological integrity.

Section 5 discusses the level of standards support, or lack thereof for the integration of photogrammetry and GIS databases. The level of support available for 3D information in geospatial databases is reviewed together with the issues of migration from 2D to 3D.

The paper concludes with a summary and some reflections on the future role of the practitioners – the photogrammetrists, image analysts and geomaticians of the future.

2. DRIVERS AND REQUIREMENTS

2.1 The Drive to Database-Centric Operations

The fundamental requirement for the database management system (DBMS) is of course that of being a secure and widely accessible repository for the geospatial data gathered at such considerable cost. At the same time as the case for the integration of photogrammetry with GIS is increasing, the GIS vendors are making use of mainstream DBMS technology for its security, availability, scalability, archiving, transaction management, query support and enterprise-wide characteristics. Oracle Spatial has emerged as the DBMS of choice for the GIS vendors and hence for integration with photogrammetry, albeit often via an intermediate GIS layer or component.

In the GI enterprise, the DBMS needs to support all the processes involved in the care and nurture of the data and in its delivery to customers. In addition to photogrammetric workstations it has to support desktop applications, field update operations, quality assurance processes and delivery mechanisms. There is some evidence of a trend towards separate, but linked, maintenance and delivery databases, because of the markedly different functionality and performance characteristics of these two regimes (Murray, 2003). This separation can serve to simplify the required level of integration.

2.2 Joined-Up Data

Another driver towards the database-centric approach is the need for joined-up data. At its simplest level, this is expressed as the desire to escape from the arbitrary constraints of ‘map sheets’ and the associated breaking up of features in an artificial manner. Even if the system still operates on a sheet or unit basis, the management of the resultant edge-matching task is much more feasible in a database environment. Some mapping agencies (for example Ordnance Survey Great Britain) have migrated to a continuous national coverage, with much greater ability to deliver customised selections of features to users and the prospect of greater flexibility in internal maintenance. Only

in the very recent past has DBMS technology been available to support such extended continuous spatial coverages. The need for consistency across joined-up data arises also when topographic framework data is brought together with similar data from other adjacent jurisdictions or with other layers of business data.

2.3 Conflation and use of best Available Sources.

Imagery is becoming increasingly important as the primary source for topographic framework data. According to Heipke, it will account for around 50% of such data in most NMA's over the next few years (Heipke, 2004). Other sources are also important including field survey and other data. An enterprise-wide database-centric approach is needed to support exploitation of this rich set of sources. Integration of photogrammetry with the database is part of this.

A key element for improved efficiency in the future is conflation – the combination of two datasets to produce a merged dataset with the best elements of the inputs. Automation of conflation depends on the articulation of the rules for determining the required best elements, and their implementation in a rules-based processing environment. Such rules will likely be imprecise and data will never quite fit, so fail-safe recourse to human interpretation is needed. In future this kind of approach will increase the degree of automation in data update and enhancement using imagery as well as other sources.

3. LEVELS OF INTEGRATION AND OF INTEROPERABILITY

3.1 File-based Data Exchange

Most implementations combining photogrammetry with GIS databases in use at present use file based-data exchange with the database. Data is extracted, updated and returned, either on a whole replacement basis or in some cases as files of deleted and created features. The data model is essentially a simple feature model. A variety of proprietary CAD or GIS formats are used (DGN, DXF, Shape) and often there are problems arising from loss of information (eg multiple attributes). Some organisations have developed lossless exchange formats. Validation of the modified data takes place at the end of the session(s), on return to the database. Failures of validation result in repeated revision cycles, often over long timescales if the validation processes are prolonged. Nevertheless this architecture is simple to realise and widely adopted. It represents the initial level of integration and can of course support remote operation.

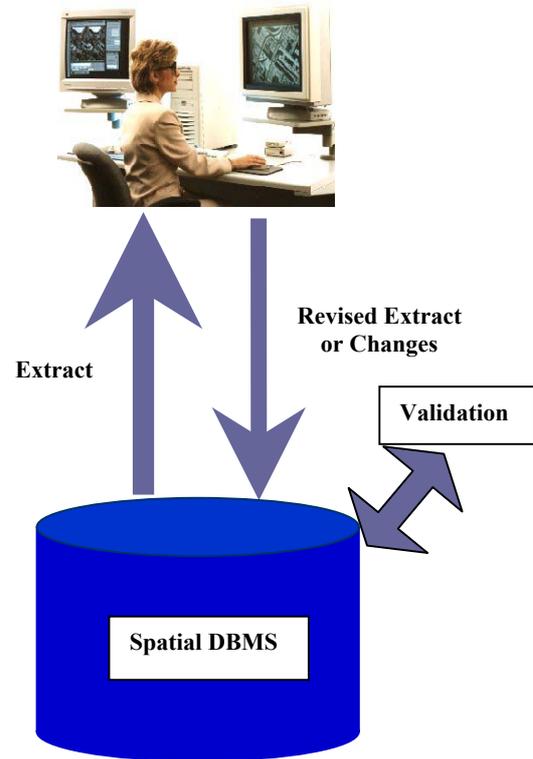


Fig. 1 File-based Data Exchange. Several cycles over extended timescales.

3.2 Direct Link to GIS Database

The next level of integration is achieved by a direct link between the photogrammetric system and the GIS database, by programmatic connection using the API's of the two systems. This avoids problems of information loss. Instead of a single commit to the database, there is a series of commits on a per completed operation basis. There is some gain in efficiency due to the closer coupling and a reduction in the 'floppynet' effect. Most mainstream photogrammetry systems can now operate in this mode with a GIS database such as ESRI's ArcSDE or via a GIS layer to Oracle Spatial.

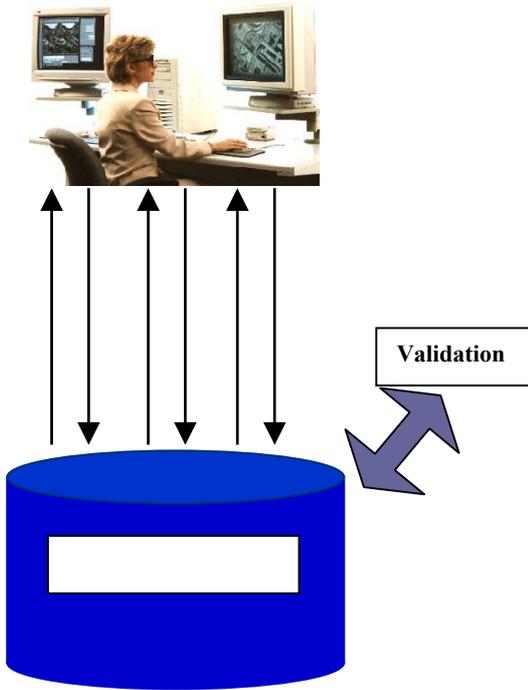


Fig. 2 Direct Link to GIS database.

3.3 Direct Link with Active Validation

Gains in efficiency are considerably greater if the GIS layer and/or database provides a high degree of active (or on-line) validation. Active topology maintenance is a key example, expanded in more detail in section 4.4 below. More generally active validation using rules-based processing is needed to efficiently support the more complex data models coming into use. A strong candidate for supporting rules-based processing is object-oriented technology. A pioneering example of the use of this integrated with photogrammetry is provided by the Laser-Scan LAMPS2/SOCET Set integration (Edwards et al, 2000; Hayles, 2001).

At the present time such architecture involves two data management environments, with the object-oriented technology providing the active validation layer and RDBMS technology being used as the storage repository. The close coupling between the photogrammetry system and the active validation layer results in large efficiency gains and single cycle operation.

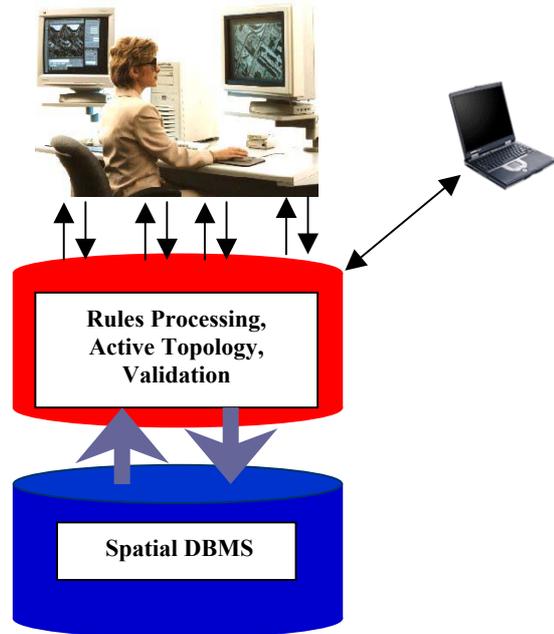


Fig. 3 Direct Link with Active Validation and Underlying Database. Single cycle operation.

The link to the underlying RDBMS still uses file-based data exchange, but only of completely validated data. More and more of the active validation functionality is becoming available in middleware or in the database and in future we can expect to see further integration with a multi-tier architecture. This will be coupled with use of better industry standard rules-based processing components and with a significantly greater role for thin clients.

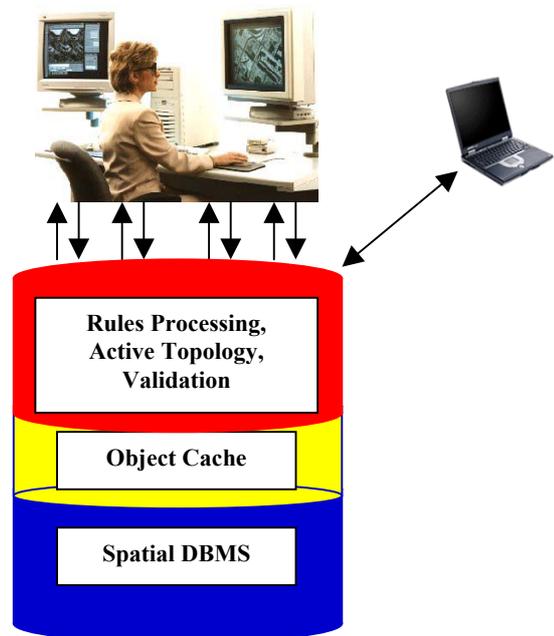


Fig. 4 Direct Link with Multi-tier Architecture.

3.4 Interoperability and Choice of Application Interface

The use of a central database across a geospatial information production enterprise should lead to interoperability 'within the firewall' of all client applications needed for the care and

nurture and delivery of the information – capture, maintenance, validation, analysis, query, elaboration, rendering and delivery. The recent initiative (Intergraph et al, 2004) to ensure client-neutral interoperability with Oracle Spatial is currently restricted to 2D data but may well be extended to 3D so as to support photogrammetric clients.

Another current trend, set to be realised in 2004, is for photogrammetric capabilities to be made available in the guise of familiar GIS editing tools operating in a 3D/Stereo environment. Examples that are anticipated include Z/I Imaging photogrammetry used via the Intergraph Geomedia interface and BAE Systems SOCET Set photogrammetry used with ESRI's ArcGIS. It will be interesting to see which style of user interface proves the more popular.

These levels of integration and interoperability presently depend on active collaboration between the different technology suppliers. Particularly for 3D data, the broader vision of interoperability, with photogrammetric capabilities deployable in a Web Services architecture and 'plug and play' integration are still some considerable distance away in the future. Some of the standards issues involved are discussed in section 5 below.

4. SOME IMPORTANT USE CASES

4.1 Update (or Revision).

Market requirements for better currency of framework data are increasing, especially from new business areas like mobile navigation, urban planning and business applications. There is in consequence more emphasis on update and or flexible patterns of update. Updating (used here as a synonym for revision) is the task of comparing the present state of the database with a more recently generated source or dataset, detecting and capturing changes and reflecting these in the database. By means of updating the database is regularly adapted to reflect changes in the real world. As such, it increasingly represents the 'bread-and-butter' tasking of photogrammetric workstations. Some industry observers (Keating, T., private communication) have gone so far as to state that 'The need to populate and maintain GIS databases has driven a re-growth in the photogrammetry community'.

The trend towards richer data models in these GIS databases is a powerful motivation towards closer integration of all update tasks (including photogrammetry) with the database. The richer the model, the more checking is needed, and the greater the cost of remedying undetected errors, particularly as contamination of the data can spread enormously. Indeed such contamination may not be recoverable except by very expensive manual intervention and re-doing the whole process. Automation of the checking processes is both necessary and to a large extent achievable, but there is a big incentive to centralise these processes at the database level. Two particular cases are examined in sections 4.3 and 4.4 below.

Other important aspects of Update include the maintenance of metadata, to reflect the current status of update, and, in many instances, the preservation of the history of previous states of the data. Both of these aspects are suitably handled at the database level.

4.2 Refinement and Positional Accuracy Improvement.

Refinement is the process of increasing the quality or content of existing data in terms of its geometric accuracy, its topological structure or its thematic content (by the addition of further attribution).

The advent of widely available high accuracy GPS positioning has highlighted inaccuracies in absolute coordinate positioning in many core or framework topographic datasets. These have historically maintained high levels of relative positional accuracy – the much lower level of absolute accuracy having been less material. With the wider use of such data in association with contemporary GPS systems, this particular 'nettle' is now having to be grasped by national mapping agencies and their customers. The situation is much more widespread than might be generally known (EuroSDR, 2004) and rectifying it will be the cause of considerable investment over the next few years.

Suitably controlled imagery (either used in photogrammetric workstations, esp. when 3D data is involved, or as orthoimagery if 2D data is involved) is a primary source for positional accuracy improvement (PAI). Almost inevitably, in PAI programmes, changes in the data due to real world change (update) and changes due to accuracy improvement are generated together. It is important for users of the data to be able to distinguish between these, in so far as this is possible. It is easy for the data supplier to be unmindful of the problems posed for users in modifying the positional content of their own data, consequential on the refinement of the accuracy of the topographic or framework data. Unfortunately the remaining errors in absolute position are typically unsystematic since systematic errors will have been dealt with already – see Fig. 5 for example.

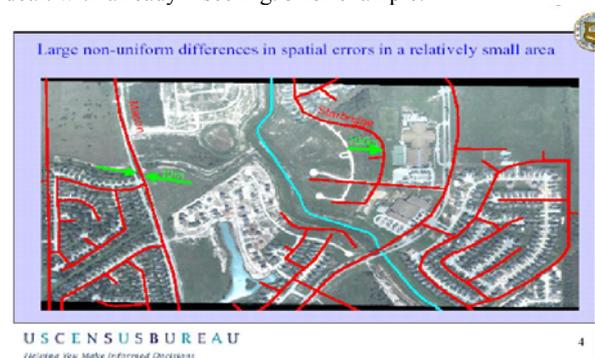


Fig. 5. Non-uniform differences in spatial errors in US Census TIGER data.

The provision of sufficient information (eg fields of shift vectors) to allow users in conjunction with their software systems to adjust their own data is a key aspect of the task, and poses an additional requirement on any systems (including photogrammetric workstations) used for PAI to record and generate such information.

The PAI task for topographic or framework data may well be of such a magnitude as to call for some degree of automation, although to date it has been addressed in a more or less piecemeal short-term manner. Given the present state of the art in feature extraction, it would seem that the constrained problem – 'There is, very likely, a feature with this geometry somewhere near this location in the image. Is it still there? If so, what are its coordinates? If not, refer to operator' –

should be capable of a degree of solution sufficient to cut timescales and costs.

4.3 Support for Persistent Public Feature Identifiers.

A major aspect of the new generation of data models being adopted by national mapping agencies and other data suppliers is the shift to meaningful features (often termed 'objects', at this level of discourse the terms are effectively synonymous) with unique and persistent identifiers. This model provides strong support for a change-only or incremental update service to users. In the geospatial community this technique was first effectively adopted in the S57 standard of the International Hydrographic Organisation (IHO). It is now well established for topographic framework data, and is used for example by Ordnance Survey Great Britain in providing an incremental update service for its national MasterMap and Integrated Transportation Network datasets. Progress in this area can be monitored via the regular Joint ICA/ISPRS/EuroGeographics Workshops on Incremental Updating and Versioning of Spatial Databases, one of which precedes this Congress (ICA, 2003).

Identifiers also play a key role as the 'hooks' by which user data can be related to framework data. Because they are relied upon by user applications, they force a more rigorous approach to the semantics of update. Lifecycle rules have to be defined across the tasks of creation, deletion, splitting and merging of objects and the modification of their geometric, thematic, topological and temporal descriptions. These have to cover identifiers, and the circumstances in which identity is retained or lost.

All update processes need to be aware of these rules and to enforce them if update is to be efficient. They also need to respect existing identifiers and be capable of issuing new ones, either by access to a central registry or by a surrogate mechanism, leaving the final assignment to the commit stage. In all events the trend towards this form of data model strengthens the case for tighter integration between the database and the photogrammetric systems.

It is worth noting that if framework data had provided identifiers as hooks from the outset, the consequential problems for user data from PAI would have been avoided.

4.4 Support for Topology

The desirability of active topology maintenance in update processes has been well documented and the lack of it is one of the major contributors to the extended timescales and repeated round trips of the early generation of systems (Edwards, 2000). For 2D data, support for topology is becoming increasingly available. The recently released Oracle 10g provides support for the storage of topology. Laser-Scan's Radius Topology complements this with active server-side topology maintenance (Laser-Scan, 2004). ESRI's ArcGIS (ESRI, 20023) provides topology support on the client side. There is still debate, as yet unresolved, as to which approach provides the better overall efficiency.

The essential point for the discussion of integration is that the technology for active topology management in 2D is well-established and available. A level of integration which exploits this will deliver substantial productivity gains, as slivers and overshoots become things of the past. A key

overall architecture decision to be made is whether active topology maintenance needs to be available only through the GIS interface, or to any application that uses the standard database interface.

Photogrammetry is of course intrinsically concerned with 3D data. Support for topology in 3D is a much more open issue and is left to the next section.

4.5 Further Benefits

The use cases described in this section, which all represent major aspects of the overall task, all argue towards a greater degree of integration between photogrammetry and the database. Further benefits are potentially achievable with better data management and closer alignment with the IT mainstream. Photogrammetry system vendors and GIS vendors seek to off-load transaction management, history management and archiving to the DBMS. Furthermore, it is arguable (Garland, 2004) that the greatest savings may arise from the adoption of workflow management technology, which becomes accessible with improved data management environments.

5. STANDARDS AND 3D SUPPORT

5.1 Standards

There is a common feeling that the integration of photogrammetry and GIS databases is not yet at a sufficiently mature stage to be the subject of Standardisation. In particular the lack of adopted and implemented standards for 3D data is an obstacle. GML2 from the Open GIS Consortium has proven its worth for transferring and serving 2D data and for supporting incremental update. It is being widely adopted. Implementations of GML3, which supports 3D data and topology, are at a very early stage. Some photogrammetry vendors have reported early experiments with GML (Olhof et al, 2004).

Nevertheless, from the wider perspective, the shift to database-centric environments stands to benefit from the progress towards interoperability and better access to information through the application of OpenGIS and ISO specifications.

As an aside, the progress towards the Sensor Web (OGC, 2004) as reported in a special session of this Congress will have significant effects on front-end data and imagery gathering processes.

5.2 Stages towards 3D support

Databases which support spatial data typically include support for 3D data (although with significant restrictions in areas such as indexing and query). There is no difficulty in storing 3D information, although many national mapping agencies still do not retain z-values even when they are captured. This is changing as market demand for 3D vector data in addition to DEM increases. In practice 2.5D vector data (z as a single valued function of (x,y)) can be created from 2D data and sufficiently high quality DEM data.

Issues arise in the handling of multiple z-values and topology in 3D. The implementation of full 3D topology, whilst defined in the ISO standards, is a long way off. Many observers, the author included, would advocate seeking this

ISPRS, 2004. Annual Report 2003. Vol. 9, No 1 March 2004. <http://www.isprs.org/publications/highlights/highlights0901/annualreport03.pdf>. (accessed 27 April 2004).

Laser-Scan, 2004. Radius Topology at http://www.laser-scan.com/technologies/radius/radius_topology/index.htm. (accessed 27 April 2004).

OGC 2004. Sensor Web Enablement and OpenGIS SensorWeb at <http://www.opengis.org/functional/?page=swe>. (accessed 27 April 2004).

Pla, M., Sánchez S., and Colomer J-L., 2001. Updating topographic databases: the example of the Topographic Database of Catalonia at scale 1:5000. OEEPE/ISPRS Workshop "From 2D to 3D – Establishment and Maintenance of National Core Geospatial Databases", Hannover, Germany, October 2001. http://www.ipi.uni-hannover.de/isprs-wg2-4/oeepe01_pdfs/colomer.pdf (accessed 27 April 2004).

Acknowledgements

The author is pleased to acknowledge the inputs into the thinking behind the paper of several players from the commercial and academic sectors including Terry Keating and Paul Garland of Z/I Imaging, John Allan of BAE Systems, Eberhard Gülch of Stuttgart University of Applied Sciences, Paul Watson of Laser-Scan and David Maguire of ESRI.

Developments of commercial products is rapid and the interested reader should track them via the websites of the companies involved.