BRINGING A VECTOR/IMAGE CONFLATION TOOL TO THE COMMERCIAL MARKET

Lou Martucci, Emeritus Scientist
Battelle Pacific Northwest Division
Richland, Washington 99352
lou.martucci@pnl.gov

Boris Kovalerchuk, Professor
Central Washington University and
BKF Systems
Seattle, Washington 98178
bkf1@BKFSystems.com

ABSTRACT

This paper addresses the conflation problem of integrating/aligning/fusing vector and image data in geospatial products, with special focus on the aspect of bringing a solution to the commercial market. Users of geospatial data in government, military, industry, research, and other sectors have need for accurate displays of information such as roads and other terrain information in areas of interest and operations. Our general approach to vector/raster conflation examines the problem in three activity areas: preprocessing, conflation processing, and postprocessing. We use two well-developed and complementary methodologies with the goal to integrate them into a unified framework for an optimized conflation solution. This research is conducted within an Army Small Business Innovation (SBIR) project with the critically important aspect of pursuing a technology transfer and commercialization strategy that would result in a likely pathway for transition into an operational capability. We describe fundamental principles and generalized roles of participants in the commercialization process. Further, we introduce the concept of putting technically sound products to beneficial use through the steps of (i) defining the specific use scenarios and the respective operational/business environment of that use, and (ii) performing product marketing in accordance with use scenarios and the stimulation of related environments. Several sample scenarios are presented, along with operating/business environments, to demonstrate the concept. The approach assesses the technological readiness of the user for a vector/raster product with a view towards application of a more penetrating market analysis that attempts to pinpoint the technology transition opportunities in a complex and ever expanding geospatial data arena.

INTRODUCTION

Users of geospatial data in government, military, industry, research, and other sectors have need for accurate displays of roads and other terrain information in areas where there are ongoing operations or locations of interest. In general, two types of geospatial data are available—maps and satellite images—but the persistent problem has been that the information contained in them over common areas does not always match. Because the data sets come from different sources and their spatial accuracy varies, there are difficulties in integrating imagery (raster data) with maps (vector data). For example, the image of a road may not match the existing map of that same road. In such cases, the geospatial analyst typically uses editing software and follows a complicated and lengthy process to adjust the roads on the maps to match with the satellite images. This problem and process are generally referred to as *conflation*, and finding a solution that is significantly more automated than the employment of costly and scarce human resources has become a challenging technical issue for the geospatial community. A typical example of the conflation problem is presented in Figure 1.

The integration/fusion/alignment of vector and raster geospatial data is a difficult and time-consuming process. Data to be integrated may have inaccurate and contradictory geo-references or may not have them at all. Contributing to the challenges of robust integration are issues of different and unknown rotations, disproportional scales, uncontrolled noise, and other factors. Vector and raster data conflation problems are discussed in the literature in publications such as (Cobb et al., 1998), (Jensen, Saalfeld et al., 2000), (Chen, Knoblock, et al, 2004), and (Kovalerchuk, Schwing, 2005; Kovalerchuk, 2007). The vector/raster conflation (VRC) problem is closely related to the problems of imagery registration and feature extraction; e.g., (Brown, 1992), (Shah, Kumar, 2003), (Seedahmed, Martucci, 2002), and (Doucette et al., 2004).

Figure 1. Example of road mismatch between satellite imagery and map data (Google map).

The authors of this paper are collaborating with a team of researchers to investigate conflation methodologies and develop an appropriate solution (Doucette, Kovalerchuk, Brigantic, Seedahmed, Graff, 2007). The general technical approach in our vector to raster conflation (VRC) technology addresses the problem in three categories of activity. First, *preprocessing* of data identifies and assesses *metadata and knowledge* that are available but may not be contained explicitly in the image or vector data. Such information, if available, can guide processing steps that actually conflate/align data. Second, *processing* steps are designed to integrate two well-developed methodologies for conflation: algebraic structural algorithms (ASA) and similarity transformation of local features (STLF). Third, *postprocessing* steps are designed to approve/disapprove and fix the conflation result. Figure 2 depicts these processes.

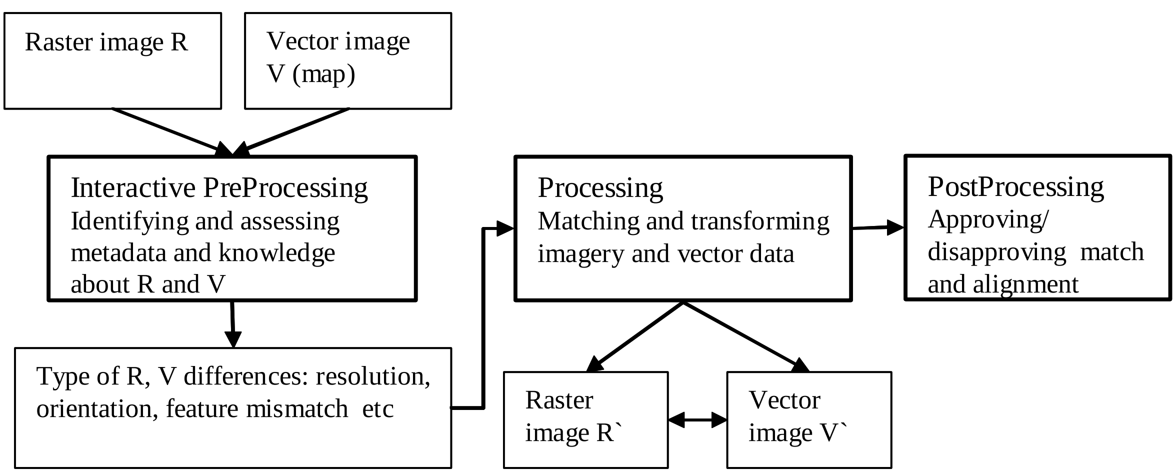

Figure 2. General framework of technical approach.

Both methodologies match and transform imagery and vector data. Preliminary results have explored feature extraction from semi-automated and manual algorithms, and the matching of features extracted from imagery with roads presented in the vector source data using ASA and STLF algorithms. Where the road data contained only short fragments of actual roads, we have found that the use of ASA and gap analyses and filling algorithms were able to discover and structurally match extracted lines. We have also seen that STLF was able to successfully conflate data with gaps. Our continuing research work will focus on the complementary nature of the two conflation methodologies. Our goal is to precisely identify the relative strengths and weaknesses of these approaches and to integrate the findings into a unified methodology toward developing an optimized conflation solution.

A critically important aspect of the current phase is the *technology transition and commercialization* strategy that describes the vision of the research and the most likely pathway for technology transition of the SBIR from research to an operational capability that satisfies one or more of the Army's operational or technical requirements.

The 2006 National Defense Authorization Act encouraged commercialization of SBIR technologies in order to accelerate the transition of SBIR technologies, products, and services, including the acquisition process. The Army's overall objective is to increase SBIR technology transition and commercialization success, thereby accelerating the fielding of capabilities to soldiers and to benefit the nation through stimulated technological innovation, improved manufacturing capability, and increased competition, productivity, and economic growth (United States Army, 2008). Furthermore, the Army and other Federal agencies acknowledge the civilian marketplace with related business opportunities and recognize the potential significant value gained for the United States economy through commercialized dual-use products.

The approach taken in this paper is *not* to describe the technical details of the conflation research and development but rather to examine basic commercialization principles for such geospatial products and to present an analytical marketing approach using case scenarios and the operational/business environment from a user perspective.

TECHNOLOGY TRANSITION AND COMMERCIALIZATION APPROACH

Overview of Technology Transition Principles

This section draws heavily on the concepts and procedures for technology transfer documented by (Sandelin, 2008). Although Sandelin's manual was created primarily to assist university technology transfer offices in commercialization efforts of faculty, staff, and students, it is also a valuable resource for small businesses and non-profit organizations engaged in technology transition, transfer, or commercialization activities.

Underlying Legislation. The most well-known and basic law related to the technology transition/transfer process is given in the Bayh/Dole Act of 1980 (Public Law 96-517), which provides rules for ownership and handling of inventions created with Federal Government support. Among other things, it gives U.S. universities, small businesses and non-profit organizations intellectual property control of their inventions that result from federal government-funded research and allowed such entities to pursue ownership of an invention before the government. The policy and objective of Bayh/Dole are given below (U.S. Code Collection, 2008):

> *"It is the policy and objective of the Congress to use the patent system to promote the utilization of inventions arising from federally supported research or development; to encourage maximum participation of small business firms in federally supported research and development efforts; to promote collaboration between commercial concerns and nonprofit organizations, including universities; to ensure that inventions made by nonprofit organizations and small business firms are used in a manner to promote free competition and enterprise without unduly encumbering future research and discovery; to promote the commercialization and public availability of inventions made in the United States by United States industry and labor; to ensure that the Government obtains sufficient rights in federally supported inventions to meet the needs of the Government and protect the public against nonuse or unreasonable use of inventions; and to minimize the costs of administering policies in this area."*

Key Participants in Pursuit of a Commercialization Objective. The *inventor(s)* of the intellectual property to be licensed usually initiates the process by disclosing the invention, identifies people in industry who should be interested in the invention, participates in obtaining patent protection, and responds to technical questions about the invention. In the above discussed SBIR research and development of a conflation solution, the potential inventors would be the identified members of the technical team and any other participants who contribute to the intellectual property.

The *representative of the technology transfer and intellectual property function*, who is situated within the respective organizations of the inventor(s), evaluates the proposed intellectual property/invention and considers it for patenting or other legal protection. The representative also has a role in determining the potential value of the invention and typically participates in the negotiation of licensing agreements and in subsequent monitoring of the product development, including adjustment of license agreements with licensee parties as necessary. In our SBIR project, this role would be identified within BKF Systems, NG4 GeoSpatial Solutions, Battelle PNWD, and ITT Corporation as warranted by intellectual property/invention claims of the researchers.

A fundamental tenet of patent law is that a government provides a legal grant to inventors in order to encourage innovation, promote technology development, and to foster economic growth. The basic concept is *not* to give the patent owner exclusive right to exploit the patent but only *to exclude others* from practicing the invention so long as the patent is in effect. Another basic concept is to understand that a patent is effective in the country of issue and to appreciate that questions about patentability and patent validity can be interpreted differently in different countries. It is important to rely upon patent professionals in pursuing intellectual property concepts and rights. A suggested reference on this subject is (Knight, 2001).

The third key participant in the commercialization process is the *representative of the potential licensee* who determines the value and importance of the intellectual property/invention to his organization. In our SBIR project, we are considering partnering with a licensee already successfully engaged in the marketplace with available products or tools that can be extended by incorporating our conflation solution; we may also consider a standalone product and market it in some consortium arrangement among our respective organizations with potential venture capitalists. This critical decision is influenced by the operating concept and ongoing development of the envisioned conflation system. Nonetheless important in our endeavor are the early contacts and discussions with potential licensees and other interested parties as we move forward in Phase II of the SBIR. The roles of participants and the commercialization process are generalized in Figure 3. It is understood that there is some overlap and collaboration among the participants in this process.

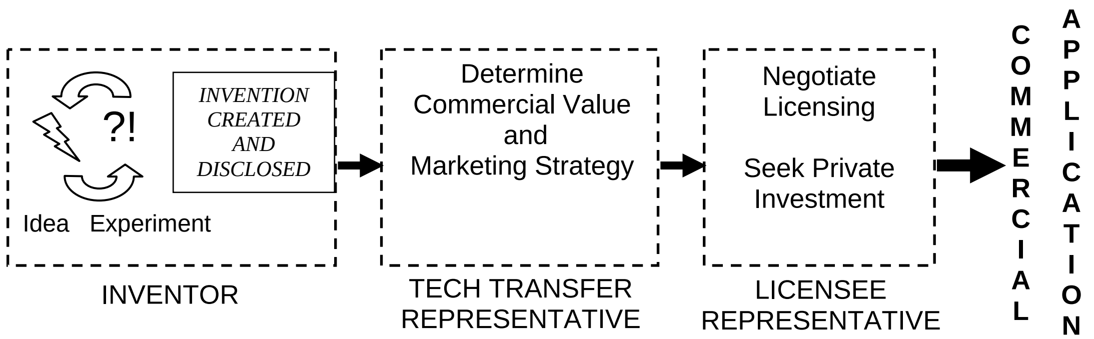

Figure 3. Generalized Roles and Process.

Understanding the Need. Tantamount to success in this technology transition and commercialization activity are identification and validation of the need for the envisioned product by potential users. This topic was touched on in the Introduction but is now included as an essential element of this discussion on the technology transfer/transition approach. The need argument enters into the roles and context of each participant discussed above. For example, the Army has preference for a conflation solution that is compatible with commercial off-the-shelf (COTS) software such as ArcGIS, ERDAS IMAGINE, INTERGRAPH, FEATURE ANALYST, or others. The need expression is common to many users or producers of geospatial intelligence (GEOINT) information. For example, the National Geospatial-Intelligence Agency (NGA) St. Louis is a GEOINT provider and has critical need to maintain current databases in order to support their customer's mission. In order to update and add new features to existing databases various new sources are exploited, such as commercial satellites, national satellites, aero/ground mobile platforms, and field-provided data. The growth of these data is exponential. The few available COTS tools typically conflate partial solutions and business rules are required for each combination of data set pairs in order to use a COTS tool.

We further argue that present methods of conflation require extensive human intervention and labor, which is an expensive effort and drain on human resources that could be more productively spent in cognitive, decision-making tasks. Hence, we have established a technical goal to solve the conflation problem with a solution that is optimally automated (i.e., minimizing human involvement), robust and reliable, and provides uncertainty metrics. Throughout the technical research and the commercialization process, keeping in mind the need and its evolving impact on the geospatial users community will drive the interests and vision of the key participants over the life cycle of the SBIR project.

Concept of Use Scenarios and Environment

The technology transition and commercialization process for geospatial products is in essence a process of putting technically sound products to beneficial use. At first, this requires clarification of four critical aspects:

(1) how the product can be used; that is the *use scenario*, and

(2) the operational and business *environment* of that use.

(3) *finding* and/or *stimulating* the needed operational and business environment for the product to be used.

(4) *marketing* of the product in accordance with (1) **-** (3).

If the needed operational and business environment already exists, it is obvious then that marketing can be quite successful. Otherwise, the focus should be on stimulating the needed environment. Typically this type of activity falls into a *crack*, because it is neither totally within traditional technical development of the product nor fully encompassed by typical marketing activity. This is probably the major reason why many technically successful SBIR projects never reached market or did not become profitable—because it seems that no one is working on stimulating the market. The second difficulty is that even when inventors and small businesses recognize this problem and try to solve it, they have little capacity for stimulation of the needed environment where government and private organizations operate. Therefore, the wider involvement of the Government and professional societies such as ASPRS is needed. On the positive side of this situation, the required environment is essentially common to many geospatial products, not just a vector to raster conflation tool, and therefore many offerings will benefit from stimulating the encompassing business environment.

Below we clarify what the use scenarios are for a vector to raster conflation (VRC) product and what the required environment is.

Scenario Example. Assume that an analyst has a duty to update the road network for a specific geographic area using satellite imagery at an organization such as The National Geospatial-Intelligence Agency (NGA), Army Topographic Engineering Center (TEC), U.S. Geologic Survey (USGS), or the government of a State, County, or City. The scenario includes capabilities for the steps of:

(a) buying/producing imagery and vector data,

(b) conflating the data,

(c) storing results for later use by different users at these organizations and their clients, and

(d) distributing results.

Environment. While steps (a) **-** (d) are common for all listed organizations, their respective environment readiness is quite different. Most of these organizations receive or buy imagery from third parties. Some of these organizations self-produce vector data, and some of them receive and/or buy vector data from other sources. Respective storing and distributing capabilities and their restrictions also are quite different. We may grade capabilities on the scale from 0 (do not exist) to 5 (full capabilities). Say organization A has these capabilities: a = 5, b = 3, c = 4, d = 5. As we see, probably only training of the users is the obstacle for the use of VRC in this organization. It is most likely that large federal organizations belong to this category represented by organization A, but small Counties may belong to a category B represented with capabilities of: a = 1, b = 1, c = 2, d = 2. If the needed/desirable environment is a = b = c = d = 4, then it is obvious that stimulating this environment for an organization in category B is much harder than in category A.

Technological Readiness Assessment

The marketing analysis based on described principles is much deeper than traditional marketing research that sends inquiries or questionnaires to a potential user for determining interest in buying a product and for evaluating the size of the market using the few responses typically received.

The proposed approach analyzes the technological readiness of the user for a VRC product. Indicators (a) **-** (d) are more or less illustrative and can be elaborated further. For example, there are numerous sources (both commercial and government) that provide imagery (raster) and/or map (vector) data for a price, such as: Digital Globe, MDA Federal, Positive Systems, Sanborn, Infotech, USGS, National Aeronautics and Space Administration (NASA), Google Maps, Microsoft TerraServer, MapQuest, GlobeXplorer, and others not listed here. Many of these sources provide geospatial data as well as technical services particularly for the integration of imagery and map information to create raster-vector products that potentially could be enhanced with a VRC conflation tool.

Another aspect of this category of providers is the growing trend for Web-based technical services and commitment to open standards such as those implemented by the Open Geospatial Consortium (OGC). In view of this geospatial data/services provider sector, our analysis might be adjusted to reflect capabilities of: (a) collection and storage of imagery and vector data, (b) abilities to conflate the data within Web-based or otherwise offered technical services, (c) retention of conflation results for resell to other customers as standalone or augmented products, and (d) innovative and progressive applications of conflation to growing needs of customers. Again, the main point here is the use of a *more penetrating market analysis approach* that better pinpoints the commercialization opportunities in a complex and ever expanding geospatial data arena.

SAMPLE SCENARIOS

Here we provide some representative scenarios to further illustrate the proposed marketing approach. In scenarios S1 to S4 below we assume a County or other local government entity has probable need for VRC but has capabilities of a = 1, b = 2, c = 2, and d = 2. The County has several geospatial objectives where a vector to raster conflation solution can be productively employed as the availability of geospatial data increases. It is realized that the data sets come from a variety of sources and have different spatial accuracies; hence, the data do not match very well, causing problems such as misalignment of vectors with the respective imagery in areas of interest. The current approach to conflating these is employing intensive human labor; i.e., analysts have to manually move the vectors, or adjust the images (rubbersheet) to match the vectors, or possibly redigitize the vectors from existing data sets and transfer the associated attributes. The types of problem sets facing this County have a wide range but could encompass the following scenarios:

S1: The integration of aerial or satellite imagery covering the County's internal road network with existing map data of this transportation system. This problem possibly includes raster to raster registration to fuse local aerial data with broad area coverage of satellite imagery.

S2: The special problem of matching end points of road segments that cross the border between this County and adjoining counties to enable consistent and accurate road connectivity of mutual interest. The combination of vector to vector and VRC conflations might be needed in this situation.

S3: The issue of significant seismic hazard zones where earthquakes tend to occur and induce ground settlement or liquefaction. Specific image/vector information is needed for detailed planning of commercial and/or residential developments.

S4: The objective of strengthening State, County, and local preparedness for management of emergencies caused by natural phenomena or man-made events. Robust integrated geospatial information assets are essential for adequate planning, mitigation, response, and recovery actions. Combined vector to vector, raster to raster, and vector to raster conflations might be needed in this multi-faceted project.

We also introduce two additional possible scenarios for users in other than government sectors:

S5: Detection and clearance of mines and unexploded ordnance, left over from military operations in worldwide locations, is a humanitarian concern for many nations and communities. Mine clearance entails surveys, mapping and minefield marking, as well as the actual clearance of mines from the ground. Geospatial data from various sources are needed to produce multi-layered geographic information systems. Acquisition of aerial and satellite data support this endeavor and need to be appropriately integrated and matched with vector data in various regions where existing map products vary greatly.

S6: Scientific researchers in the earth sciences, such as geology, might have need for very precise and rapidly performed conflation of raster and vector data covering a geographical area of interest wherein a special study is underway. For example, a geostatistical methodology for integrating elevation estimates derived from digital elevation models (DEM) and elevation measurements of higher accuracy. It would be very important that existing maps of sparse ground-based elevation measurements are matched and aligned precisely with the more abundant points on DEMs derived from acquired aerial/satellite raster data.

ENVIRONMENTS AND APPROACHES

With respect to the County scenarios above, our proposed marketing approach would assess the County's technological readiness in meeting its geospatial problem sets where a VRC would be beneficial. To help improve the operating environment surrounding each scenario, we outline the following four respective approaches (numerically keyed to the above scenarios):

E1: In view of the County's low a,b,c,d capabilities, in order to convince the County to buy or use our VRC tool a vendor could offer free training and free demonstration of road data conflation to offset the probable limited financial resources of the County.

E2: A vendor could propose to work with the County to seek joint funding with neighboring counties on the border-crossing road segment matching problem.

E3: A vendor could propose to work with the County, other State and local offices, as well as private land development enterprises, to address image/vector information that supports planning of development projects in areas of seismic threat.

E4: A vendor could propose to work with the County to help procure grants and other types of funding from offices of emergency management and from the Department of Homeland Security to produce a case study on the use of a vector/raster conflation tool to enhance geospatial capabilities leading to better emergency preparedness.

Regarding the other two Scenarios, a vendor possibly could propose the following:

E5: With respect to the mines Scenario, a VRC vendor would identify several of the firms and institutions engaged in actual mine clearing activities, as well as determine the relevant components of the United Nations and all of those member nations that support this humanitarian endeavor. It is likely that most of the engaged demining organizations concentrate on use of equipment and techniques in the detection and mitigation process and rely on conventional or available sources for the geospatial data acquisition and integration. In view of the critical need to precisely locate and delineate the mine fields, our approach would assess the technological readiness of these firms to perform raster/vector conflation and also demonstrate how our VRC tool would possibly achieve better accuracy and higher throughput of survey information before actual mine clearance can begin. A VRC vendor would work with those benefiting firms directly or in collaboration with them to acquire sponsorship from available funding sources.

E6: With respect to the scientific research scenario, we would identify ongoing or planned geoscience research projects such as the given example and the funding sources of such activities (e.g., DARPA, NASA, NSF, National/Federal Laboratories, academia, etc.). It is likely that some contacts would be derived from collegial and professional networks of the VRC vendor. It is likely that researchers in geoscience areas will focus on their specific disciplines and take advantage of available sources for relevant geospatial data acquisition and integration. In order to perform experiments and deep analyses it is safely assumed that researchers will want to eliminate or minimize the contribution of errors and "noise" within those analytic processes of interest. Again the VRC vendor approach would offer use of the VRC tool to effectively and accurately conflate geospatial data and thereby reduce error/noise contributions from that particular source. The VRC vendor could offer the technical services in this activity or provide technical training to the researcher. Moreover, the VRC vendor could offer to work with the researcher in a collaboration supported by his/her funding source, or could jointly work with the researcher in finding a sponsor of the project.

APPROACHES TO IMPROVE ENVIRONMENTS

In the previous section, we discussed some possible scenarios and associated operating/business environments surrounding those scenarios. It can be seen clearly that the given environment of a potential user might encompass the several and various operational scenarios in which the user intends to operate or perform. Thus the lack of a specific capability, such as inability to conflate geospatial data, will probably impact all use scenarios where conflation has a significant role. In other cases, where the user basically acts in a singular type of scenario, such as a researcher in a given domain of expertise, the weakness in conflating geospatial data might not be a factor in all situations. This distinction will be important to the VRC vendor engagement with potential clients/users of geospatial data and the respective assessment of their technological readiness for VCR. Further, we recognize that our approaches to improve given operating/business environments have to consider the user's business and technical objectives, his/her clients' interests and requirements, the constraints of financial resources, the available expertise/knowledge within the user organization, and other factors that will influence user interest in our VCR tool. From the above brief outline of several possible scenarios/environment, it is evident that the approaches will vary and that a "one size fits all" approach will *not* be beneficial to the effective technology transition and commercialization goal.

CONCLUSION

This paper addressed aspects of the technology transition and commercialization process for a vector/image conflation tool. An overview of technology transition principles was presented in the context of current efforts in the performance of an ongoing SBIR project. A fresh perspective was given on product use scenarios and an analytic assessment of the associated operating/business environment of users. In pursuing Phase II of our project, we expect to invoke the appropriate principles of technology transfer and to carefully consider business opportunities that mutually satisfy our objectives and the needs of potential customers for our VRC product. The outlined technology transition and commercialization approach can be beneficial and effective for the wider geospatial community due to commonality of the environments for different geospatial products.

ACKNOWLEDGEMENT

We are pleased to acknowledge the support and interest of the Army Topographic Engineering Center (TEC), Alexandria VA, in our ongoing SBIR activity, and in particular we acknowledge the efforts of Brian Graff, at Army TEC.

REFERENCES

Brown, L., 1992. A survey of image registration techniques, *ACM Computing Surveys*, vol. 24 (4), pp. 325-376.

Chen, C., Knoblock, C., *et al.*, 2004. Automatically and accurately conflating orthoimagery and street maps, In: *Proceedings of the 12th ACM International Symposium on Advances in Geographic Information Systems* (ACM-GIS'04).

Cobb, M. Chung, Foley, M., *et al*., 1998. A rule-based approach for the conflation of attributed vector data, *GeoInformatica*, 2/1, pp. 7-36.

Doucette, P., Agouris, P., Stefanidis, A., 2004. Automated road extraction from high resolution multispectral imagery, *Photogrammetric Engineering & Remote Sensing*, 70(12), pp. 1405-1416.

Jensen, J., Saalfeld, et al., (accessed June 2001). Spatial data acquisition and integration, 2000 NSF Workshop GIS and Geospatial Activities.

Knight, H. Jackson, 2001. *Patent Strategy for Researchers and Research Managers*, John Wiley & Sons, LTD, England. 2001.

Kovalerchuk, B., Schwing, J., (Eds), 2005. *Visual and Spatial Analysis: Advances on Data Mining, Reasoning and Problem Solving*, Springer.

Kovalerchuk, B., April 9-12, 2007. Affine invariant and robust image registration/conflation algorithm, In: *SPIE Defense and Security Symposium Proceedings*, Orlando, FL, Vol., 6568, 6568-05.

Doucette P., Kovalerchuk B., Brigantic R., Seedahmed G., Graff, B., Oct., 2007. A methodology for automated vector-to-image registration, In *IEEE Applied Imagery Pattern Recognition (AIPR) Workshop*, Wash. D.C.

Sandelin, J., (accessed January 2008). *An Operations Manual for a Technology Transfer Organization Based on the Stanford Model*. http://www.innovationmatters.com/publications/TTO.php

Seedahmed G., Martucci, L., (Sept. 9-13, 2002). Automated image registration using geometrically invariant parameter space clustering (GIPSC), *Archives of the International Society of Photogrammetry and Remote Sensing*, Symposium 2002, Graz, Austria.

Shaw, M., Kumar, R. (Eds.), 2003. *Video Registration*, Kluwer.

United States Army SBIR/STTR Programs, (accessed January 2008). *Instructions for Submitting Phase II Proposals*. http://www.armysbir.com/sbir/07p2_instruct.htm

U.S. Code Collection, Cornell University Law School, (accessed January 2008). Text of the Bayh-Dole Act (35 USC 200), policy and objective. http://www4.law.cornell.edu/uscode/35/200.html