Integrating GIS technology with RiverWare (Phase 2)

Functional Requirements

Author: Patrick Lynn

Graphical information systems (GISs) support storage, analysis, and display of spatial data. The document "Integrating GIS technology with RiverWare" presents requirements for the phased integration of GIS technology within RiverWare, including a good deal of background information as well as a detailed description of the requirements and design for the first phase of integration. Phase 1 introduced a new workspace view, the Geospatial View, which includes support for georeferenced objects and a background image registered to the view's coordinate system. This phase was completed in Spring 2011 and was first released in RiverWare 6.0.

In this document we develop functional requirements for the second phase of GIS integration. The general goal of this phase is to provide support for additional GIS data formats, and initially we focus on the ESRI World File format for georeferencing raster data. The ability to read and write the World Files will simplify the registration process for the Geospatial View's background image -- when the user provides a background map image for the Geospatial View, RiverWare will now be able to simultaneously acquire georeferencing information for that image. This information will locate the map in the Geospatial View's projected coordinate system, enabling RiverWare to display it in proper registry with the workspace objects without requiring the user to interactively register it. In addition, the ability to write image georeferencing information in the World File format will facilitate sharing of maps between RiverWare models and between RiverWare and external GIS applications.

While the limited time and resources available for this development limit the scope of this work to the World File format, RiverWare could make use of numerous other data formats employed by GIS applications. The last section of this document discusses how RiverWare might make use of additional spatial data formats and estimates the level of effort required to achieve that functionality.

1 Motivation

1.1 Map images and coordinate systems

Map images represent projections of geospatial data onto a two dimensional coordinate system. To display an image in a coordinate system, such as that of the Geospatial View, requires projection information that is not always available as part of the image. Consider the following three categories of image (raster) data formats, each with progressively more projection metadata:

- Plain image: a plain image can be conceptualized as an array of colors, one for each pixel in the image. Formats vary along many dimensions, e.g., whether or not support is provided for compression, transparency, or multiple layers. RiverWare currently supports the reading of many standard image formats, including: PNG, JPEG, BMP, and SVG.
- Registered image: a plain image plus additional information describing the position of this image in a coordinate system. The additional information could be as simple as a pair of coordinates defining the bounding box of the image in the coordinate system, or it could be more general, such as an equation defining the translation from pixel coordinates to map coordinates.
- Georeferenced image: an image, a description of the coordinate system associated with the image, as well as registration information for the image. For geographic and projected coordinate systems the metadata include a spec-

ification of the datum, and for projected coordinate systems the metadata also include a specification of the projection.

Note that some image formats, such as TIFF and JPEG2000, support the embedding of image metadata in a header section of the image file. For example, the GeoTIFF image format uses this feature to support fully georeferenced images. On the other hand, image metadata can also be stored in a separate file, in which case the two files are typically associated by name. One such format is the ESRI world file format which consists of 1) a plain text file containing parameters for a simple image-to-coordinate system transformation, and 2) rules for naming the world file based on the name of the image file. For example, for the image file MyMap.jpg, the accompanying world file would be MyMap.jgw or MyMap.jpgw.

1.2 Limitations of the Geospatial View

For the Geospatial view background image, RiverWare 6.0 supports only plain images. After providing a plain image file, the user is required to interactively register that image in the Geopspatial view's coordinate system by either: 1) directly entering the coordinates of the lower left and upper right corners of the image, or 2) providing coordinates for two reference points in the image (via the Image Location Dialog). The image registration information is then saved with the canvas configuration information in the model file.

The Geospatial View's support for coordinate systems and the background image has the following limitations:

- Registering an image can be burdensome to the user. For example, it involves typing, and in addition the user might not know the required information (the bounding box of the image or the coordinates in the image's projection of two reference points).
- Each time an image is set as the background image it must be registered, even if the user had previously registered the same image. This is because RiverWare maintains registration information for the current image only
- A single image used by multiple models must be registered for each model. This is because RiverWare saves registration information with the model file, as part of the Geospatial view's configuration.
- Users might not be able to easily provide SimObj coordinates in the Geospatial view's coordinate system. This might be because they only know them in another system or they do not know what the coordinate system of the Geospatial view is (i.e., they don't know the projection of the image).
- All coordinates must be in the same coordinate system, and it is not possible to convert coordinates from one system to another. This means that the user can only enter and view object coordinates in the Geospatial view's coordinate system, when they might prefer geospatial coordinates, for example.

To relax all of these limitations would require the ability on the part of RiverWare to represent coordinate system specifications and to use this knowledge to perform translations from one system to another. While supporting this functionality is part of the long-term vision for RiverWare, it is beyond the scope of the current work.

2 Requirements

The goal of the current task is to is to simplify the image registration process, addressing the first three limitations discussed in the previous section. To meet this goal, we propose to:

- Add support for the reading of at least one common registered image format. This will allow users that have access to registered images (in a supported format) to skip the image registration step.
- Add support for the writing of an image in at least one common registered image format. This will allow users to avoid having to perform the registration process more than once for a particular image.

2.1 Choosing a format for image registration (JPEG/PNG + ESRI world file)

We would like to support a registered image format to which our users are likely to have access, that is, we would like it to be the case that those users who have access to registered map images, have access to them in a format supported by RiverWare. As a starting point for narrowing down the list of possibilities, we considered the formats supported by ESRI's ArcMap 10.0 for map export with registration information. They are: PDF, JPEG, PNG, TIFF, GIF. Of these we eliminated from consideration PDF and TIFF because RiverWare does not currently support reading of those formats, and GIF because it only supports 8 bits/pixel, leaving JPEG and PNG for consideration. Since both formats use the world file format to provide the registration information, provide support for both formats requires minimally more effort than just supporting one format, so we elect to support both JPEG and PNG.

The world file format was introduced over 10 years ago (with ArcInfo 7.0), is simple to read and write, and appears to be commonly used within the broader GIS community. A world file contains 6 floating point parameters, A-F, which define an affine transformation in the form:

$$x_1 = Ax + By + C$$
$$y_1 = Dx + Ey + F$$

where

x_1 = calculated x-coordinate of the pixel on the map
y_1 = calculated y-coordinate of the pixel on the map
x = column number of a pixel in the image
y = row number of a pixel in the image
A = x-scale; dimension of a pixel in map units in x direction
B, D = rotation terms
C, F = translation terms; x, y map coordinates of the center of the upper left pixel
E = negative of y-scale; dimension of a pixel in map units in y direction

RiverWare 6.0's Geospatial View assumes rectangular pixels with no rotation, thus not all map images that can be described with a world file can be displayed as background images. Initially RiverWare will check that B and D are zero and that A and E are equal, and issue an error if one of these conditions is violated. Eventually, we hope to extend support to all world file transformations.

For many GIS applications, creating an image with a world file is relatively straightforward. To do this in ArcMap 10.0, select File->Export Map...", then check the "Write World File" checkbox option within the subsequent dialog. On the other hand, when this is not possible, the world file format is simple enough that it can be created by hand with a text editor.

2.2 Detailed requirements

Within the Geospatial Canvas Configuration dialog, when the user loads a background image file (either by entering a path, selecting a file in the file chooser, or clicking the Reload button), RiverWare will determine if registration information is available for the image by looking for a world file image with the appropriate name. If a valid world file is discovered, the user will be asked if they would like to set the Background Image Rectangle Location from this information. If RiverWare does not find registration information for the image or if it does but the user does not want to use it, then the behavior remains unchanged (RiverWare attempts to give the image a reasonable default rectangle location). In any case, the user can proceed to change the registration values directly if they wish.

A button labeled “Export World File” will be added to the bottom of the Background Image Rectangle Location panel of the Geospatial Canvas Configuration Dialog. Clicking this button will bring up a file chooser, allowing the user to select the output file name, but pre-selected to a file name that adheres to the world file naming conventions.

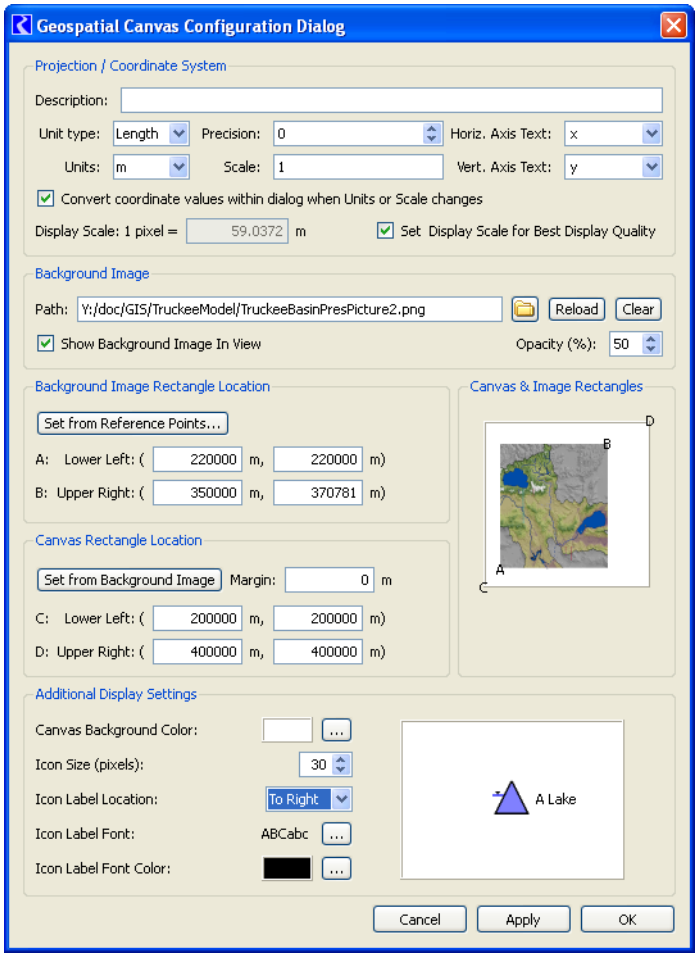

Figure 1. RiverWare 6.0 Geospatial Canvas Configuration Dialog.

3 Supporting additional spatial data formats

In this section we discuss how future RiverWare development might make use of additional GIS data formats, and we estimate the level of effort required to extend RiverWare in this way. The Geospatial View uses different types of spatial data for the purposes of image registration and object location. We consider support for additional data formats for these two purposes separately.

3.1 Supporting additional data formats for georeferenced images

Displaying an image in the Geospatial View requires raster data which has been rendered with a particular color table and which is located in the projection's coordinate system. RiverWare supports the reading of images in the PNG and JPG formats; while neither of these image formats supports the inclusion of metadata within the image file, the images can be georeferenced (i.e., located in the Geospatial View) by providing an accompanying file in the ESRI World file format.

Rendering spatial data so that it can be displayed is, in general, a complex process involving different techniques for different types of data and different rendering goals. For example, one of many options for displaying single-band raster data is to draw contour lines. Once the decision has been made to use a contour line plot, additional decisions must be made about the contour interval as well as the line color, width and style.

Most modern GIS applications support the rendering of many different types of data and typically provide a rich set of display options. Instead of trying to recreate this complex functionality, RiverWare assumes that users have access to spatial data which has already been rendered, i.e., maps.

Therefore, when considering additional formats for Geospatial View images we focus our attention on raster data formats that include both color and georeferencing information. For example, we do not consider the Shapefile format because it is designed for representing vector data. The following are some of the formats which can support georeferenced images:

- GeoTiff: this is an extension of the TIFF image format which adds georefencing information. A medium-sized open source library exists for reading and writing this format.
- MrSID: this is a proprietary format designed explicitly for georefenced raster graphics. A library for reading files in this format is available from LizardTech.
- JPEG 2000: is an image compression standard which support the embedding of metadata, allowing the inclusion of georeferencing information for map images.

3.2 Supporting additional data formats for georeferencing objects

The object location data is vector data in the form of spatially located points with attributes giving the name and type of the object at that location. RiverWare supports the reading and writing of these data in the ESRI Shapefile format, allowing RiverWare to display icons at the associated object's location in the Geospatial View.

3.3 Plan and time estimate

To avoid maintaining separate code and data abstractions for each format type, we would like to extend RiverWare's spatial data support by linking in a more general library. The prime candidate is the GDAL/OGR open source library which supports the reading of the formats mentioned above as well as more than a hundred additional raster and vector formats. This is a large library and so should only be linked in when it is being used (probably a plug-in architecture will be appropriate).

In addition to allowing RiverWare to provide additional flexibility with respect to input of images and object locations, access to a library such as GDAL/OGR, would be useful for extending RiverWare functionality in other ways as well. For example, support for the KML format would be useful for linking RiverWare to web applications displaying spatial data. As another example, when projection information is available, a library such as this would support translation between different coordinate systems, such as geospatial (lat/long) to map coordinates.

We estimate that incorporating a general GIS library into RiverWare would take approximately 3 weeks of effort, with the following subtasks requiring about 4 days each:

- Further research on library alternatives, considering: cost, availability for target architectures, functionality, licensing, suitability for high-level software architecture (i.e., as plug-in). Installation and testing of library.
- Inclusion in configuration management process (build & release processes, licensing).
- Design and development of conditional link mechanism (presumably a QPlugin).
- Integration with existing RiverWare code.

Note that this work would only be undertaken for the Windows platform, but should ideally be supported for both 32- and 64-bit operating systems and executables.