CREATING AN IMAGE DATASET TO MEET YOUR CLASSIFICATION NEEDS: A PROOF-OF-CONCEPT STUDY

James D. Hurd, Research Associate
Daniel L. Civco, Director and Professor
Center for Land use Education and Research
Department of Natural Resources Management and Engineering
College of Agriculture & Natural Resources
University of Connecticut
Box U-4087, Room 308, 1376 Storrs Road
Storrs, CT 06269-4087
james.hurd_jr@uconn.edu
daniel.civco@uconn.edu

ABSTRACT

The amount of spatial information, particularly remotely sensed image data, has increased dramatically during the past 10 years. As researchers, we have gone from appreciating the spectral resolution of sensors such as Landsat TM and ETM, ASTER, and IRS and their capabilities to distinguish among various earth ground features to appreciating the spatial resolution of sensors such as satellite based IKONOS and QuickBird, and a multitude of aerial based sensors such as Leica Geosystems' ADS40 and Z/I Imaging's DMC with their ability to better define the spatial extent of earth ground features. While each of these image types are valuable in there own right, the type of imagery used is ultimately dependant on the objective of the application.

Over the past several years, researchers at the University of Connecticut's Center for Landuse Education and Research (CLEAR) have used Landsat TM and ETM imagery to map land cover and land cover change for five dates over a 21 year period for the state of Connecticut. While this type of information has proven very valuable to land use decision makers at a regional scale, the spatial and thematic detail has been limiting at the more localized level. Now that powerful computing processing technology and disk storage space has become more affordable, and with the abundance of available image types, researcher's at CLEAR have begun to look at the creation of a higher spatial, more thematically rich, land cover map for Connecticut.

This paper will overview a proof-of-concept study that examined the development of a suitable image dataset based on the data fusion of a 2006 Landsat ETM image with 2006 NAIP color digital aerial imagery and will address potential issues and techniques for classification of this fused image dataset to derive an improved land cover map over CLEAR's Landsat based land cover for a portion of the Connecticut coastal area.

INTRODUCTION

Image fusion, resolution merge, or pan-sharpening, as used in the art and science of remote sensing, are typically used interchangeably to refer to the technique of combining the spatially detailed structure of a high-resolution panchromatic image with the spectral information of a lower resolution multispectral image to produce a high-resolution multispectral image (Klonus and Ehlers, 2007). The purpose is to enhance the imagery with the intent of obtaining information of greater quality; with the exact definition of 'greater quality' ultimately depending on the intended application (Wald, 1999).

There exist a number of methods for image fusion. Common techniques include:

- **The Modified IHS (intensity, hue, saturation) Transformation** – This process involves the conversion of three input multispectral bands from red-green-blue (RGB) space to intensity-hue-saturation (IHS) space. The higher resolution panchromatic band is substituted for the intensity channel and the IHS transform is converted back into RGB space maintaining the panchromatic bands spatial structure (Vrabel, 1996; Zhou *et al.*, 1998; Siddiqui, 2003).
- **Principal Components Merge –** Principal components analysis is commonly used for image enhancement and compression. The original spectral data are mathematically transformed into new, uncorrelated images called components (Chavez and Kwarteng, 1989; Vrabel, 1996). As a data fusing method principal components are calculated from the multispectral image. The higher resolution panchromatic band is

stretched to have similar mean and variance to the first principal component and then substituted for the first principal component followed by an inverse principal components transformation to derive the fused image (Zhou *et al.*, 1998).

- **Brovey Transform –** This process uses a ratio algorithm to combine the multispectral and panchromatic images. Essentially the technique normalizes each of the multispectral bands used and multiplies these with the higher resolution panchromatic band. When layer stacked, the result is a synthesized higher resolution multispectral image (Pohl and Van Genderen., 1998).
- **Wavelet Transformation –** This process first histogram matches the higher resolution panchromatic band to each of the possible three input multispectral bands. These modified panchromatic bands are converted into three wavelet coefficients which identify the high frequency spatial structure in three directions (vertical, horizontal and diagonal). A wavelet decomposition is constructed using the multispectral image (3 bands) and high frequency panchromatic information to produce four components. An inverse wavelet transform is then performed introducing the spatial detail into each multispectral band to produce a final 3 band higher resolution multispectral image (Sanjeevi *et al.*, 2001).
- **High Pass Filter (HPF) Resolution Merge –** This process is an arithmetic technique that applies a spatial enhancement filter to the high-resolution image before being merged with the multispectral data set on a pixel-by-pixel basis. The HPF fusion combines both spatial and spectral information using the band-addition approach (Chavez, 1991).

Each of these techniques has their benefits and detractions depending on the expected use of the fused image. For this project, the Ehlers Fusion technique was used. No attempt was made to examine or compare the aforementioned fusion techniques.

Ehlers Fusion Technique

The Ehlers fusion technique was developed as a fusion method that would better preserve the spectral characteristics of the multispectral data set while improving the spatial quality of the fused image (Ehlers, 2004). The driving principal behind preserving the spectral characteristics is that the high resolution image has to spatially enhance the multispectral image, that is enhance high frequency changes such as edges and high frequency gray level changes in an image, without adding new gray level information to its spectral components, particularly in areas of homogenous land cover features (Ehlers, 2007). To accomplish this, two processes need to occur. First, the spatial and color information must be separated. Second, the spatial information needs to be altered in a way that allows adaptive enhancement of the images.

For optimal color separation, The Ehlers Fusion technique uses the IHS transform applied to the multispectral image. Since the IHS transform is limited to three input bands (RGB), the process is extended to include more than three bands by using multiple IHS transforms until the total number of bands is exhausted. An adaptive filter design used to enhance the spatial information is derived using a Fourier transform. With the fast Fourier transform (FFT), the intensity channel from the IHS transform is filtered using a low pass (LP) filter and the panchromatic image (P) is filtered with an opposing high pass (HP) filter. These images are converted back into the spatial domain using an inverse fast Fourier transform (FFT^{-1}) process and added together to generate a fused intensity channel with low frequency information from the coarse spatial resolution image (via the Intensity channel) and high frequency information from the high resolution panchromatic image. An inverse IHS transformation (HIS^{-1}) is performed to produce the final fused image that contains the spatial information supplied by the panchromatic image with the spectral resolution of the multispectral image. Again, since the IHS transform is limited to using three bands, the process is extended by using multiple IHS transforms until the total number of bands is derived. An overview flowchart of the Ehlers Fusion process is presented in Figure 1. The Ehlers Fusion technique was applied using the interface found in ERDAS IMAGINE 9.2

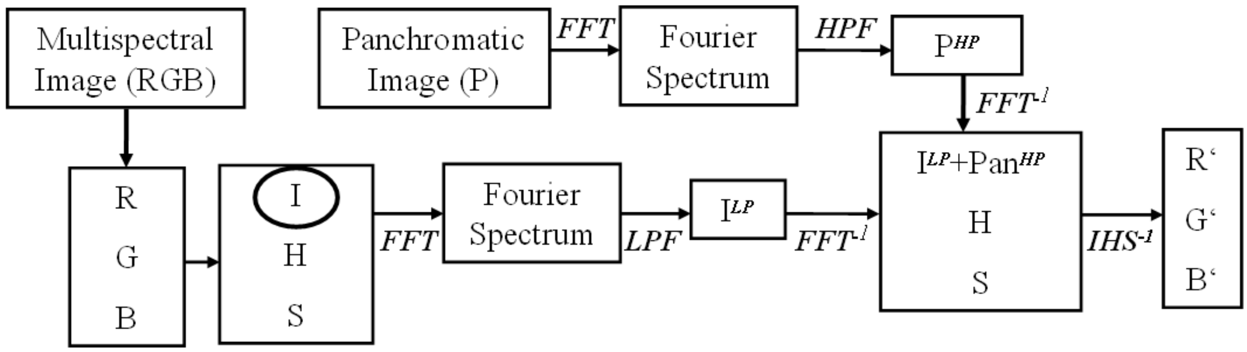


Figure 1. Basic overview of the Ehlers Fusion process (adapted from Ehlers *et al.*, 2006).

METHODS

The goal of this proof-of-concept-study was simple; create an image of sufficient spatial and spectral quality to identify the primary land cover features of interest that would be found in a heterogeneous landscape composed of various anthropogenic and natural objects using readily available image data.

Test Area and Data

To test the feasibility of fusing two data sets collected by different sensors, a small study area was selected along the central coast of Connecticut. This site covers a nine square kilometer area and is composed of various land cover features such as naturally occurring forests, tidal wetlands, beaches and water in addition to human altered environments such as residential and commercial structures and grasses, parking lots, roads, railroads, and agricultural fields. It is at this level of thematic and spatial detail that is the intended goal of the classification of the fused image dataset.

The data sets used for the fusion process include (1) a Landsat ETM satellite image from July 17, 2006 which served as the lower resolution data set and (2) imagery collected for the National Agriculture Imagery Program (NAIP) in late August 2006 which served as the higher resolution data set. For best results, datasets collected on or near the same date are preferable. This eliminates problems associated with phenological differences among images due to seasonal variations or changes in land cover that may occur within the region of interest over time. For the Landsat data, only the six 30-meter spatial resolution reflective bands plus the 15-meter spatial resolution panchromatic band were utilized. The NAIP imagery contained three bands covering the visible portion of the electromagnetic spectrum (blue, green, red) at a 1-meter spatial resolution.

(a) Landsat ETM Multispectral **(b) Landsat ETM Panchromatic** **(c) NAIP True Color**

Figure 2. Comparison of (a) the original 30-meter Landsat multispectral image data, (b) the original 15-meter Landsat panchromatic image data, and (c) the 1-meter NAIP imagery covering the study area.

Fusion Procedure

The first step of the process was to re-project the original image data into the Connecticut Stateplane projection, NAD83 feet datum. During the re-projection process, the image data were also re-sampled, using the nearest-neighbor method, to slightly adjust pixel sizes so that smaller pixels would be completely encompassed by larger pixels during the fusion process. The Landsat multispectral image was re-projected and re-sampled to 80 feet whereas the Landsat panchromatic image was re-projected and re-sampled to 40 feet. The NAIP imagery was first "degraded" from 1-meter to 3-meter spatial resolution using the Image Degradation tool in ERDAS IMAGINE 9.2. This process reduces the resolution of an image by an integer factor in the X and Y directions and averages all of the original "small" pixels that make up the new "big" pixels. In the case of this project, an integer factor of three was used for the X and Y. The degraded NAIP image was then re-projected and re-sampled to 10-feet.

One of the major considerations when fusing image data, particularly from different sensors, is the geometric registration of one scene to the other (Chavez, 1986). Any mis-registration among images will result in a reduced spatially detailed fused image. Since the three image datasets used in this project were acquired at nearly the same date, the image data sets were merely overlaid in turn and toggled to visually discern if any variations occurred along roads, and shorelines. In addition, vector roads and hydrography GIS layers were draped over the images to identify any spatial problems. These visual examinations did not detect any noticeable mis-registration problems among the image data sets.

Since Landsat ETM data was used with both multispectral and panchromatic data available, the second step of the process was to fuse these data, a 1:2 image fusion ratio, using the Ehlers Fusion technique to generate a 40 foot resolution multispectral image. The intent was to derive a multispectral image with increased spatial information that could be further enhanced using the NAIP imagery with a smaller image fusion ratio of 1:4 as opposed to a ratio of 1:8 if only the Landsat multispectral data was used. The Ehlers Fusion technique as implemented in ERDAS IMAGINE 9.2 provides several processing options (Figure 3). A nearest neighbor re-sampling technique was used to resample the multispectral image to the panchromatic image. All six reflective bands were used in the fusion process which results in the use of multiple IHS transforms, and an additional IHS algorithm is applied to resample the multispectral image prior to the fusion. The output fused image is an unsigned 8-bit image. The filter design was set to automatically adjust to fit the imagery, using an urban or mixed image content (due to the heterogeneity of the study area) and a normal setting for the color/resolution tradeoff resulting in a balance between spatial and spectral preservation. No advanced features, which allow for the customization of the fast Fourier transformation, were used.

The result of the Landsat fusion compared to the original Landsat multispectral image are provided in Figure 4 (a and b). Although an improvement in the spatial information is apparent in Figure 4 (a and b), it was desired to determine if it was possible to further enhance the spatial information based on just the Landsat imagery. A crisping filter was identified as a potential procedure to perform this function. The crisping filter as provided in ERDAS IMAGINE 9.2, sharpens the overall scene luminance without distorting the variance content within the image. The process performs principal components analysis, applies a sharpening (convolution summary) filter to the first PC, then performs an inverse principal components analysis to produce a sharpened image. The result of this process was deemed successful and beneficial in further enhancing the Landsat fused image and is shown in Figure 4c.

The third step was to fuse the NAIP imagery with the fused and sharpened Landsat imagery. Since the NAIP imagery is a three band true color dataset, it was necessary to generate a panchromatic image from the three bands. A simple process of averaging the three bands was used to generate the panchromatic NAIP imagery. Again, the Ehlers Fusion technique was used to fuse the NAIP panchromatic and Landsat image to derive a 10-foot spatial resolution dataset. A similar process was followed with the exception that only bands 4, 5, 6 were utilized from the Landsat image. The result is a three band, 10-foot resolution image of bands 4, 5, 6 which were layer stacked with the true color bands of the NAIP imagery to derive a final six band 10-foot spatial resolution image. A comparison of the original 80-foot resolution Landsat image, 40-foot fused Landsat image, and 10-foot NAIP/Landsat fused image displayed in various band combinations is provided in Figure 5, with a larger view of the NAIP/Landsat fused image provided in Figure 6.

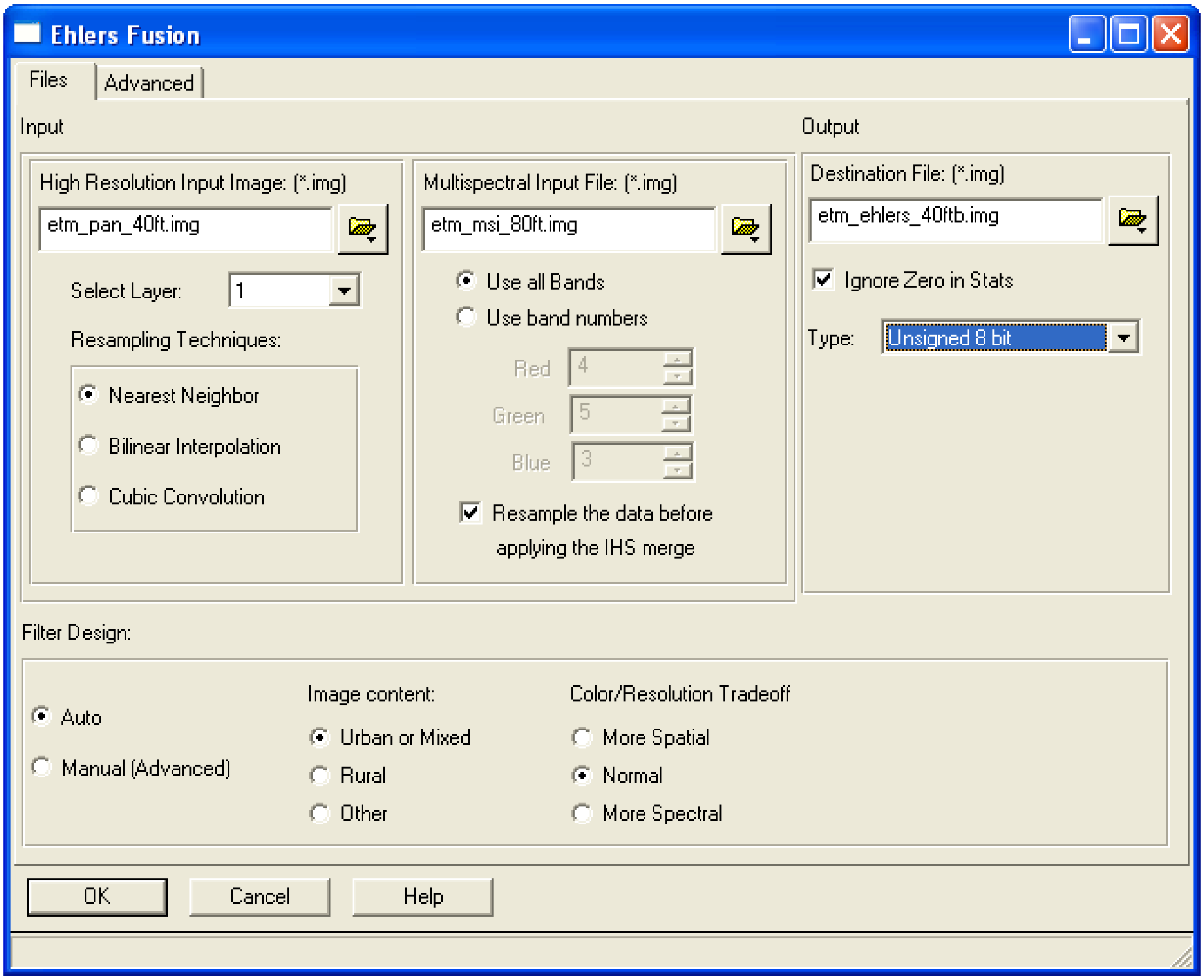


Figure 3. Example of the Ehlers Fusion technique interface in ERDAS IMAGINE 9.2.

(a) Landsat ETM Multispectral
80-foot spatial resolution

(b) Landsat ETM Fused
40-foot spatial resolution

(c) Landsat ETM Fused
following crisp filtering
40-foot spatial resolution

Figure 4. Comparison of (a) the resampled 80-foot Landsat multispectral image data, (b) the fused 40-foot Landsat multispectral image data, and (c) the fused and sharpened 40-foot Landsat multispectral image data covering the study area.

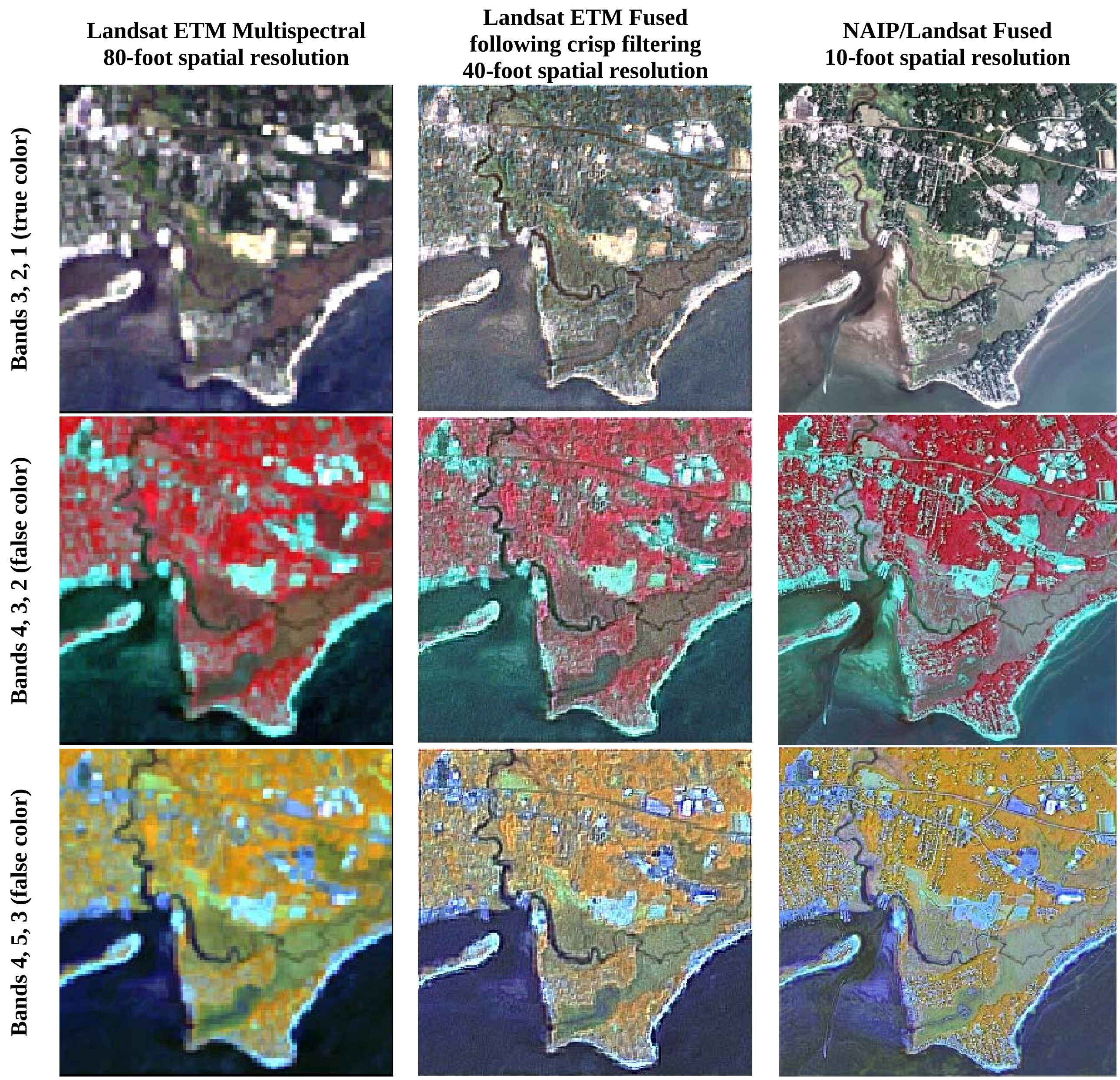


Figure 5. Comparison of original re-sampled Landsat multispectral image (left column) with fused Landsat multispectral and Landsat panchromatic images (center column) and fused Landsat and NAIP images (right column) displayed in various band combinations.

Final Fused Dataset (bands 4, 3, 2)

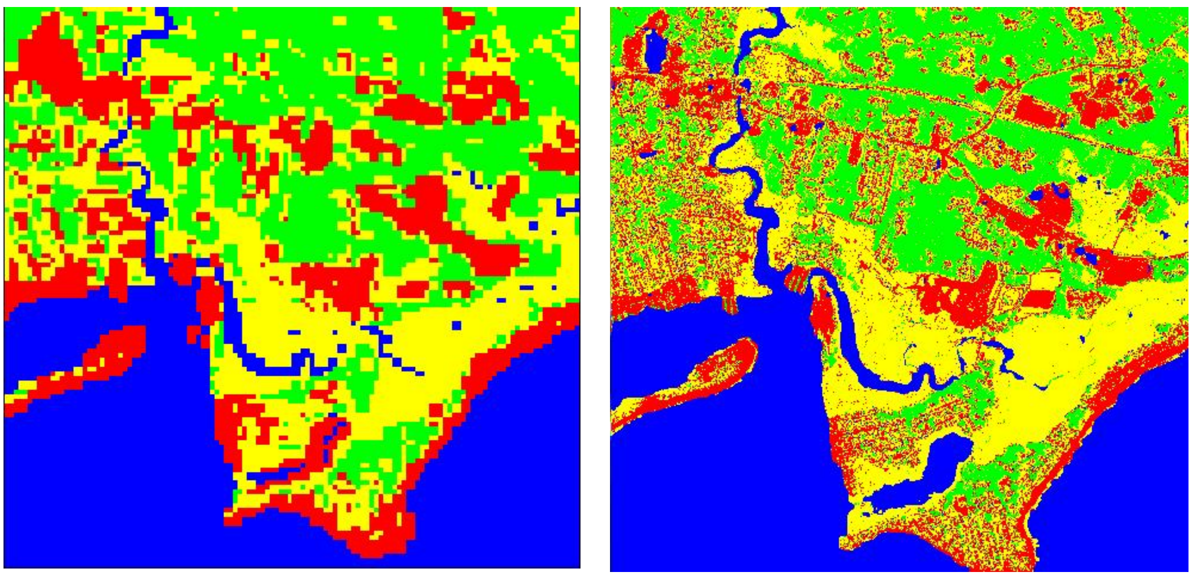
Final Fused Dataset (bands 4, 5, 3)

Figure 6. Results from the fusion of Landsat ETM and NAIP imagery using the Ehlers Fusion Technique.

Classification

As with image fusion, there exist a number of techniques that can be used to classify image data to derive thematic land cover maps. These include conventional per-pixel classifiers such as unsupervised ISODATA, or supervised approaches such as minimum distance to mean, contextual logical channel, or maximum likelihood (Kanniah *et al.*, 2007). Other classification techniques fall under the category of sub-pixel classification or spectral un-mixing including deterministic approaches, fuzzy set theory based approaches, neural network based approaches and linear mixture modeling approach (Emami, 2002). There are also various machine learning techniques such as support vector machines, artificial neural networks and decision trees (Kotsiantis, 2007). Lastly is image segmentation and object-based classification.

At the time of this writing, various classification approaches were still be examined, focused on per-pixel classification approaches and image segmentation and object-based classification. Figure 7 provides a comparison of the application of ISODATA to the 80-foot spatial resolution Landsat multispectral image (Figure 7a) and the 10-foot spatial resolution NAIP/Landsat fused image (Figure 7b). While the thematic information was kept to a minimum (only identifying four general land cover types), the difference in the level of spatial detail is apparent. This basic analysis provides evidence that the derived fused image has the potential to provide an improved land cover product.



(a) ISODATA Unsupervised Classification Landsat 80-foot Multispectral Image

(b) ISODATA Unsupervised Classification Fused 10-foot Multispectral Image

Figure 7. Comparison of ISODATA classification on (a) 80-foot spatial resolution Landsat multispectral image and (b) 10-foot spatial NAIP/Landsat fused multispectral image resulting in four general land cover categories (built-up and barren (red), forest (green), grass and wetland (yellow), and water (blue)).

Also examined is image segmentation and object-based classification. In this approach image objects are derived from input information and are then classified based on various object features such as shape, color and texture attributes. In the case of this work, the six multispectral bands from the NAIP/Landsat fused image were used in addition to derived products such as NDVI; principal components analysis, specifically components 1, 2, 3, and 5; and Tassled Cap components 4 and 6. In addition, a 10-foot spatial resolution digital elevation model derived from LiDAR data was used. Using the Definiens Developer 7.0 image analysis software, these input data layers, weighted equally, were applied to the multiresolution segmentation process within the software to generate image objects. These image objects were exported to ESRI's ArcMAP 9.2 GIS software for training. During the training process, several image objects were selected that represented a land cover category of interest. Once training was completed for each land cover class, the object features (including mean reflectance for each layer, standard deviation for each layer, several shape attributes and several texture attributes) were applied to the See5 data mining tool. See5 is used to identify the image object attribute that best identifies the land cover class of interest based on all of the input image object attributes. The result of the See5 analysis is a classification decision tree. This decision tree was recreated in Definiens Developer to generate the final classification. Figure 8 provides an example of this object-based classification. While spatially coarse the possibility of improved thematic information exists. Again this basic analysis provides evidence that the derived fused image has the potential to provide an improved land cover product. It is expected that for the final classification effort, some form of a hierarchical classification process utilizing both the per-pixel and image object approaches will be used.

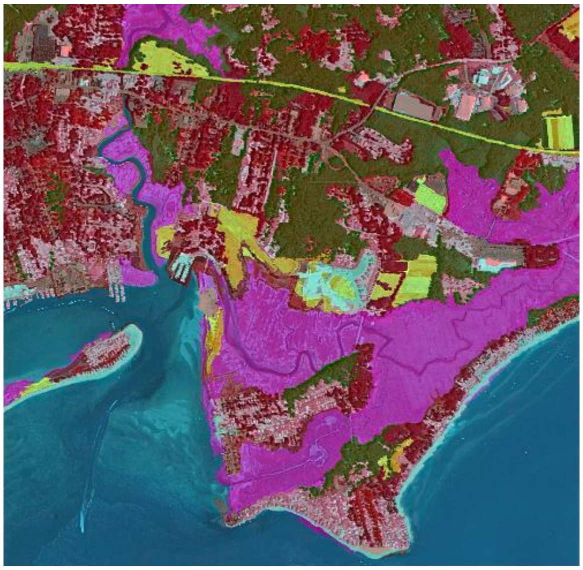


Figure 8. Example of object-based classification (transparency draped over fused image) based on segmentation of NAIP/Landsat fused image and derivative products. Pinks and reds represent built-up areas, grey barren area, green forested areas, yellow grassy areas, blue water, and magenta tidal wetlands.

DISCUSSION

The goal of image fusion is typically to derive a multispectral image that maintains the spectral qualities of the lower spatial resolution multispectral image with the spatial qualities of the higher resolution panchromatic image. Essentially the method utilized should not distort the spectral qualities of the multispectral image used (Erdogan *et al.*, 2008). As stated earlier, the fusion technique used has its benefits and detractions depending on the expected use of the fused image. That is to say, different fusion techniques will produce different results. The user must carefully choose the fusion procedure that will produce the desired outcome.

Figure 9 provides scatter plots comparing the near infrared (band 4), and shortwave infrared (bands 5 and 6) of the NAIP/Landsat 10-foot spatial resolution fused image with the original Landsat 80-foot spatial resolution image and the Landsat 40-foot spatial resolution multispectral and panchromatic fused image (the true color bands were not examined since they were supplied directly from the NAIP imagery without any fusion techniques applied). What one would like to see in these scatter plots is good correlation among each band. This would indicate that minimal spectral distortion has occurred. As is apparent, a one-to-one correlation is not present in any of these band comparisons, and although there appears to be good grouping of reflectance values (blue, green, yellow and red areas) in similar regions of each reflectance axis among some of the bands (*i.e.* NAIP/Landsat fused band 4 and Landsat band 4), there definitely exists some differences in other band comparisons (*i.e.* NAIP/Landsat fused band 6 and Landsat Fused Crisp band 6). Additionally, there is significant scattering of small numbers of pixels (magenta areas) throughout the entire scene. This indicates that the fusion technique was not successful. In the purest sense of image fusion this may be true, but for application purposes, the image still provides useful information for classification. This is apparent in the two land cover approaches examined thus far.

CONCLUSION

While there is still more work to be completed, this proof-of-concept studies appears to indicate that the process of fusing Landsat image data with NAIP data is plausible for generating an image that is of high enough spatial and spectral resolution to derive a land cover map of improved spatial and thematic detail. While in its purest sense, there is still room for improvement in the fusion of NAIP and Landsat imagery, but for application purposes the fused image shows some merit. Work is still needed to fully address the classification capabilities of the fused image to produce acceptable thematic information. Overall, this process could be beneficial for many researchers and analyst given the availability of Landsat imagery, currently being made available free through the USGS Global

Visualization Viewer website[1], and the availability of NAIP imagery, freely accessible through the USDA Geospatial Data Gateway.[2]

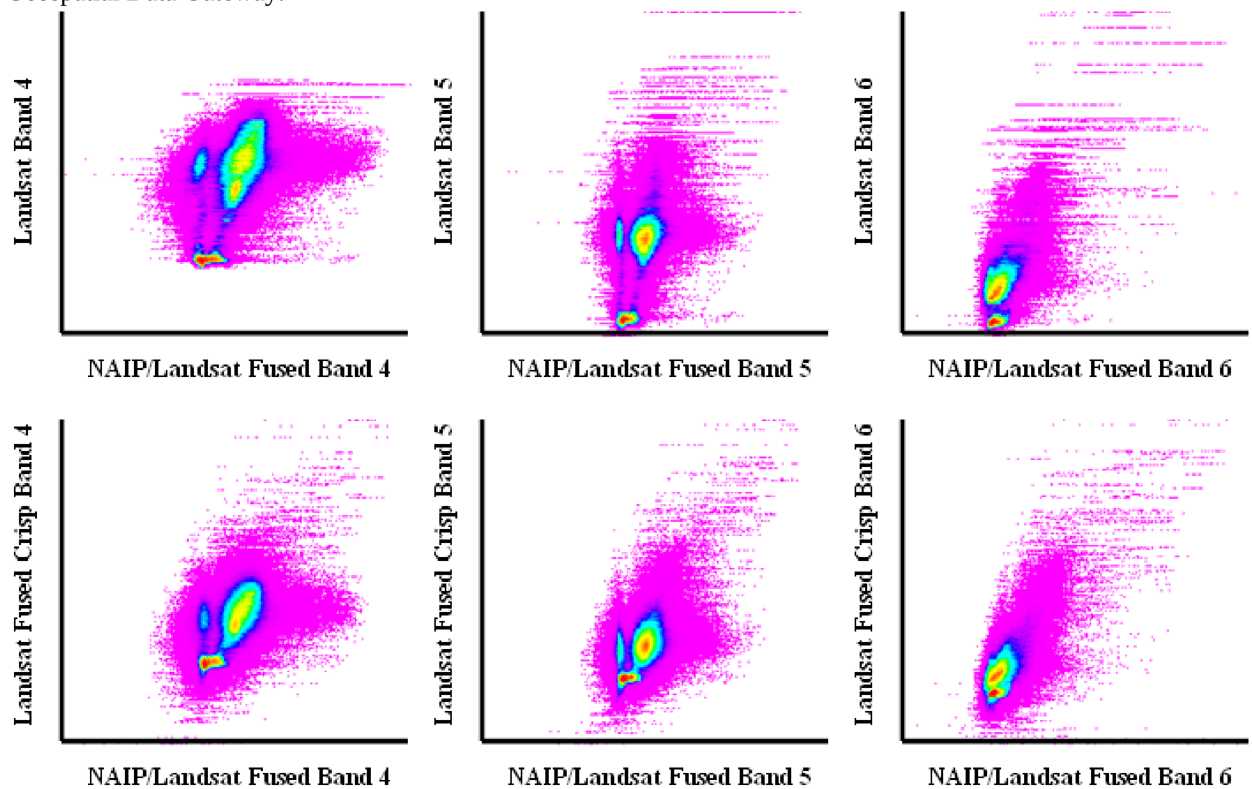


Figure 9. Scatter plots showing the distribution of multispectral bands 4, 5, and 6 of the NAIP/Landsat fused image with the original Landsat bands (top row), and fused Landsat bands (bottom row).

ACKNOWLEDGEMENTS

The research reported in this paper was supported by the National Aeronautics and Space Administration under the project *Incorporating NASA's Applied Sciences Data and Technologies into Local Government Decision Support in the National Application Areas of Coastal Management, Water Management, Ecologic Forecasting and Invasive Species*. [CLEAR Publication Number 081219.1]

REFERENCES

Chavez, P.S., 1986. Digital merging of Landsat TM and Digitized NHAP Data for 1:24,000-scale image mapping, *Photogrammetric Engineering and Remote Sensing*, 52(10): 1637-1646.

Chavez, P.S., Jr., and A.Y. Kwarteng, 1989. Extracting spectral contrast in Landsat Thematic Mapper image data using selective principal component analysis, *Photogrammetric Engineering and Remote Sensing*. 55(3): 339-348.

Chavez, P.S., S.C. Sildes, and J.A. Anderson., 1991, Comparison of three different methods to merge multiresolution and multispectral data: Landsat TM and SPOT panchromatic, *Photogrammetric Engineering & Remote Sensing*, **57,** 295± 303.

---

[1] USGS Global Visualization Viewer website: http://glovis.usgs.gov/
[2] USDA Geospatial Data Gateway website: http://datagateway.nrcs.usda.gov/

Ehlers, M., 2004. Spectral characteristics preserving image fusion based on Fourier domain filtering, In: Ehlers, M., Posa, F., Kaufmann, H.J., Michel, U., de Carolis, G. (Eds.), Remote Sensing for Environmental Monitoring, GIS Applications, and Geology IV, in *Proc. SPIE*, Vol. 5574, Bellingham, WA: 1-13.
Ehlers, M., A. Greiwe and D. Tomowski, 2006. On segment based image fusion, In *Proc of SPIE 2006*, the International Society for Optical Engineering. Stockholm, Sweden. 11p.
Ehlers, M., 2007. Segment based image analysis and image fusion, In *Proc. 2007 ASPRS Annual Convention*, Tampa, FL. 12 p.
Emami H., 2002, Evaluation and decomposition of mixed pixels in remotely sensed hyperspectral images for accuracy improvement of classification results, MSc thesis, K.N.Toosi University, Tehran, Iran.
Erdogan, M., H.H. Maras, A. Yilmaz and O.T. Ozerbil, 2008. Resolution merge of 1:35000 scale aerial photographs with Landsat 7 ETM imagery, *Int. Archives Photogrammetry, Remote Sensing, and Spatial Information Sciences*, Vol 37 Part B7, Bejing, China: 1281-1285.
Kanniah, K.D., N.S. Wai, A. Shin and A.W. Rasib, 2007. Per-pixel and sub-pixel classifications of high-resolution satellite data for mangrove species mapping, *Applied GIS*, 3(8): 1-22.
Klonus S. and M. Ehlers, 2007. Image fusion using the Ehlers spectral characteristics preservation algorithm, *GIScience and Remote Sensing*, 44(2): 93-116.
Kotsiantis, S.B., 2007. Supervised machine learning: a review of classification techniques, *Informatica*, 31(3): 249-268.
Pohl, C. and J.L. Van Genderen, 1998. Multisensor image fusion in remote sensing: Concepts, methods and applications, *Int. J. Remote Sensing*, 19(5): 823-854.
Sanjeevi, S., K. Vani and K. Lakshmi, 2001. Comparison of conventional and wavelet transform techniques for fusion of IRS-1C LISS-III and pan images, In *Proc. 22nd Asian Conference on Remote Sensing*, Singapore. 6p.
Siddiqui, Y., 2003. The modified IHS method for fusing satellite imagery, In *Proc. 2003 ASPRS Annual Convention, Anchorage*, AK. 10 p.
Wald, L., 1999. Definitions and terms of references in data fusion, *International Archives of Photogrammetry and Remote Sensing*, Vol. 32. part 7-4-3 W6. Valladolid, Spain.
Vrabel, J., 1996. Multispectral imagery band sharpening study, *Photogrammetric Engineering and Remote Sensing*, 62(9): 1075-1083.
Zhou, J., D. Civco and J. Silander, 1998. A wavelet transform method to merge Landsat TM and SPOT panchromatic data, *Int. J. Remote Sensing*, 19(4): 743-757.