CLU Digitizing on MrSID Compressed Imagery

USDA-FSA

Aerial Photography Field Office

February 2001

Table of Contents

CLU Digitizing on MrSID Compressed Imagery

Executive Summary

At the present time CLU digitizing has been done using seamless, tone matched 7.5' MDOQ tiles produced by APFO. These tiles are at the original resolution of 1 meter pixels. There has been some interest by county offices to digitize using the MrSID compressed county mosaics as a back drop rather than the MDOQ's. When imagery is compressed using the MrSID software some of the data is lost resulting in the possibility of error being introduced into the final compressed image. Whether or not this possible error justified not using the compressed imagery for digitizing had never been extensively evaluated. The purpose of this study was to do just that. The theory that digitizing differences by individuals could actually introduce more error than the compression process was the main focus of this research. If the error due to compression could be shown to be within county office standards then it was thought that the compressed imagery could be used for digitizing.

To determine the level of error introduced through the digitizing process 10 people with varying levels of GIS and digitizing experience were selected at APFO to digitize a field. Each person digitized the same 32 acre field 3 times, first using the original uncompressed imagery as a back drop and then using the MrSID compressed imagery. These field polygons were saved as shapefiles with tabular of the polygon area and distance around the polygon. The average area measured using the original imagery had a 0.1 acre difference than the average area using the MrSID compressed imagery. That is a difference of 1/3 of 1 percent. Making careful comparisons of the compressed and uncompressed imagery no offset was found between the two type of imagery.

Recommendation

It is the recommendation of APFO that FSA service centers and contractors be allowed to use MrSID compressed imagery for CLU digitizing. This will allow them to work with much smaller data file sizes and fewer images. A typical county needs about 15 7.5' MDOQ's for full county coverage. Using compressed imagery of the same county requires only one image. Both uncompressed MDOQ tiles and compressed MrSID county mosaics should be available for purchase. A no cost viewer such as ArcExplorer2 or MapSheets Express could be included on the CD when imagery is sold.

Digitizing on MrSID Compressed Imagery

Background

APFO for the last several years has been producing seamless, tone matched 7.5' tiles from DOQQ's for use by FSA service centers, other government agencies, private business, and the general public. From these tiles APFO has also produced a compressed mosaic of each county using MrSID compression software. The compressed mosaics have been produced to be used by the county offices as a backdrop or general reference image. Because the MrSID wavelet compression is a lossy algorithm some of the original image data is lost. Other compression methods such as ECW and JPEG are also lossy.

Purpose of this study

Because of the loss of data during the compression process it is possible for inaccuracies to be introduced into the imagery. Differences in tonal and/or positional attributes could possibly arise through the compression process. Whether or not this compressed imagery could be used for official measurements of distance or area by the county offices was a point of concern. At the present time the people receiving the 7.5' MDOQ's and compressed county images have been advised by APFO to not use the compressed imagery for taking measurements. Interest has been shown by several counties and other users of the data in being able to use the compressed imagery for digitizing field boundaries. CLU boundaries as well as other types of spatial data such as soils, vegetation, hydrology, and man-made features can all be easily digitized using ortho imagery as a backdrop. From this digitized data useful attributes can be extracted such as distance or area measurements.

There are several reasons for wanting to use the compressed imagery instead of the full resolution tiles. For one thing, if using compressed imagery you only need to work with one image to view an entire county. If using full resolution images you may need to load 15 or more MDOQ's to view an average size county. This requires keeping track of many more files as well as using much more hard drive space. A typical county using 15 MDOQ tiles would require almost 3 GB of disk space for the full resolution imagery. This same county imagery in compressed format would need only about 150 MB, compressed 20:1. Another advantage to using compressed imagery is that it will generally display or refresh on the computer monitor much quicker than does the original uncompressed TIFF image.

Although it had been assumed that the possibility of error being introduced into the imagery from the compression process would make it unusable for official field measurements, no thorough tests had actually be performed. In the later part of 2000 it was decided at APFO that an evaluation needed to be made to determine if the compressed imagery was acceptable or not for CLU digitizing. A hypothesis brought forth at this time was that any error caused by compression was actually less than the error introduced in the digitizing process. Because different people will digitize the same field slightly different it was thought that the human factor would be greater than the software factor for introducing error or variation into field digitizing on compressed imagery. Many things can affect how well a project is digitized. These factors can include such things as image quality, leaf on/leaf off, image interpretation skill and experience, as well as the attitude or professionalism of the person digitizing.

It was decided that the best method of determining this was to have various people digitize the same field using both a full resolution and a MrSID compressed image. Before beginning the actual digitizing tests some standards and currently used methods for field digitizing were investigated. The procedure generally used in the county offices when digitizing a field to determine acreage is to measure the field three separate times and then take the average. Another county office standard is to digitize at a scale no larger than 1:4800 (Common Land Unit 8-CM page 6-1).

Data

The field digitizing test was performed using the following software, images, and settings:

- Full resolution MDOQ M3909548
- MrSID 20:1 compressed image mosaic_1-1_s_ks103
- ArcView3.1 UNIX version
- Viewing scale 1:4800
- UTM zone 15
- Field location UTM 318148, 4356689
- Imagery of Leavenworth County, KS

Methodology

For the testing process ten people with varying degrees of experience with GIS digitized the field boundary. Each person digitized the field three times using the full resolution MDOQ as a background and three times using the MrSID compressed image as a background. Each person was asked to try their best to digitize the field exactly the same each time. These digitized field boundaries were saved for each person. The digitized polygons were saved as an ArcView shape file. Each person created a total of six polygons, three from the full resolution image and three from the compressed image. Tabular data was also created for each of the polygons. This tabular data shows the perimeter of the field in meters and the area of the field in square meters.

From this data other statistics were created. For example the average area and the average perimeter for each persons digitizing on the two types of imagery were noted. The variation between the high and low values for both area and perimeter for each person were also calculated. The averages were determined for each person as well as for the entire group. The average area measured by all ten people was 129117.9 sq. meters on the MDOQ and 128706.4 sq. meters on the compressed image. This data was converted to acres since that is the common land unit used by the county offices. The average area from all of the digitizing was 31.9 acres for the full resolution image and 31.8 acres for the compressed imagery. The difference between the two was only 0.1 acres or 1/3 of 1 percent.

On January 22, 2001 Dale Fowers of the Weber/Davis/Morgan and Bruce Lundquist of the Cache county FSA offices were contacted to obtain information regarding accuracy standards and procedures used by the county FSA offices. Both confirmed the practice of digitizing a field three times and taking the average to determine field acreage. Both also noted that slight differences in digitizing styles and accuracy levels had been seen when different people performed the field digitizing.

Cache county has no official standards for determining if a field boundary needs to be redigitized or not. After working many years in the office they get a feel for when a

measurement is too big or too small as compared to previously measured fields. If a significant difference is noted then they will either field check the area or examine the image and remeasure it. In the Weber/Davis/Morgan office they use an accuracy standard of six percent. If a new measurement varies more than six percent from the previous measurements they will investigate the problem. Any variance less than six percent is deemed to be acceptable.

Results

In the APFO digitizing study the difference noted between the area calculated from the full resolution and the compressed imagery was only 1/3 of 1 percent. This is eighteen times more accurate than the 6 percent standard used at the Weber/Davis/Morgan office. In the APFO study an average difference of 0.1 acre on a 32 acre field was noted. Using the county office standard of 6 percent the allowable difference on this same 32 acre field would be 1.9 acres.

The possibility of error or offset introduced into an image through the compression process was also examined. Earlier studies were done at APFO and by MRJ consultants regarding offset noted in MrSID compressed imagery. Both of these tests showed that there was an approximate ½ meter offset to the southeast on MrSID compressed images. This occurred when viewing MrSID compressed images in ArcView. The problem has since been fixed by LizardTech, the maker of MrSID. Testing was done this month as part of the digitizing test to make sure that a MrSID compressed image when viewed in ArcView does not exhibit any offset. Using tic marks and other areas of the image where the same pixel could be identified on both the full resolution and and the compressed image no offset was found. The selected pixels displayed the same UTM coordinates on both the compressed and full resolution images.

A separate study was done by the USDA to evaluate the horizontal accuracy level of MrSID compressed imagery. This study found that the radial error found on a MrSID compressed image was 1 meter worse than the original DOQ imagery. Their study found that the MrSID compressed imagery "is perfectly acceptable for virtually all uses of USGS raster data". (USGS MrSID Accuracy Testing, can be found on LizardTech website). This study was done in 1998. Prior to mid-2000 MrSID imagery when viewed in ArcView exhibited a ½ to 1 meter offset. This problem has since been corrected.

Conclusions

Sufficient evidence appears to exist in support of allowing the FSA County Service Centers to use compressed imagery for field digitizing. Factors in favor of using compressed imagery include:

- The DOQQ's from which the MDOQ tiles and compressed county images are made can have 10 meters or more of offset and still satisfy USGS accuracy requirements. Studies done at APFO and other locations have found only 0 to 1 meter of offset caused by the image compression.
- The hardcopy photo measuring system currently being used by the counties can have as much as a 1 percent error from the start. FSA requires that aerial photographic enlargements, representing cropland areas, be rectified to scale with an accuracy of plus or minus 1 percent (1-AP page 51). Image compression does not cause more than 1 percent error.
- No image offset was detectable in the latest APFO test of MrSID compressed imagery. A previous study had found an average offset of ½ meter.
- The difference for the average area of the digitized field on the full resolution image and the MrSID compressed image was only 0.1 acre or 1/3 of 1 percent. The Weber/Davis/Morgan service center allows up to 6 percent difference.
- If the service centers wanted to sell digital imagery of a particular parcel of land in their county they would not have to determine which CD the parcel is located on. All that would be required would be to sell the entire compressed county mosaic on 1 CD.
- A no cost viewer such as ArcExplorer2 or MrSID GeoViewer could also be included on the CD. The data could then be displayed and queried on the clients own computer.
- Less file management and disk space would be needed in the county offices.
- MrSID compressed imagery can be viewed in ArcView, ArcInfo, ArcExplorer, and the majority of GIS and image processing software.
- MrSID compressed imagery can be used in ArcPad running on any Windows CE hand held computer device.

Recommendations

- CLU digitizing on compressed imagery should be allowed by the FSA County Service Centers and contractors.
- Imagery should be available for purchase in either 7.5' MDOQ format or compressed county mosaics.
- A no cost viewer software should be included on the image CD. Since imagery could be in either TIFF format (MDOQ) or MrSID format (compressed county) the no cost viewer should be capable of displaying both of these formats. The potential is also there for providing vector data of field boundaries, soils, vegetation, or other information in shapefile format. The no cost viewer should also be able to display shapefiles. ArcExplorer2 from ESRI and MapSheets Express from ERDAS can display all of these data formats.

Compressed Image Digitizing Test

	Area	*Perimeter*
Test 1		
MDOQ		
Average	131970	1542
Variation	538	12
MrSID		
Average	130798	1543
Variation	872	2

	Area	*Perimeter*
Test 2		
MDOQ		
Average	126948	1527
Variation	1453	3
MrSID		
Average	125955	1525
Variation	901	10

	Area	*Perimeter*
Test 3		
MDOQ		
Average	130998	1545
Variation	776	12
MrSID		
Average	131453	1543
Variation	847	9

	Area	*Perimeter*
Test 4		
MDOQ		
Average	125938	1527
Variation	620	1
MrSID		
Average	125757	1522
Variation	396	9

	Area	*Perimeter*
Test 5		
MDOQ		
Average	128688	1541
Variation	1789	20
MrSID		
Average	127895	1540
Variation	1578	8

	Area	*Perimeter*
Test 6		
MDOQ		
Average	129764	1539
Variation	1925	13
MrSID		
Average	130894	1538
Variation	1756	7

	Area	*Perimeter*
Test 7		
MDOQ		
Average	129252	1536
Variation	570	13
MrSID		
Average	129543	1545
Variation	1300	4

	Area	*Perimeter*
Test 8		
MDOQ		
Average	129296	1530
Variation	663	6
MrSID		
Average	129535	1533
Variation	2134	6

	Area	*Perimeter*
Test 9		
MDOQ		
Average	129684	1538
Variation	4919	13
MrSID		
Average	128038	1532
Variation	862	6

	Area	*Perimeter*
Test 10		
MDOQ		
Average	128641	1536
Variation	1073	9
MrSID		
Average	127747	1541
Variation	931	12

General Statistics

Average Area

MDOQ average area of all ten tests
29117.9 sq. meters
31.9 acres

MrSID average area of all ten tests
128706.4 sq. meters
31.8 acres

MDOQ – MrSID difference
411.5 sq. meters
0.1 acres

Percent difference
0.32 %
1/3 of 1 percent

Average Perimeter

MDOQ average perimeter of all tests
1536.067 meters

MrSID average perimeter of al tests
1536.017 meters

Average area variation

MDOQ average area variation all tests
1432.6 sq. meters

MrSID average area variation all tests
1157.7 sq. meters

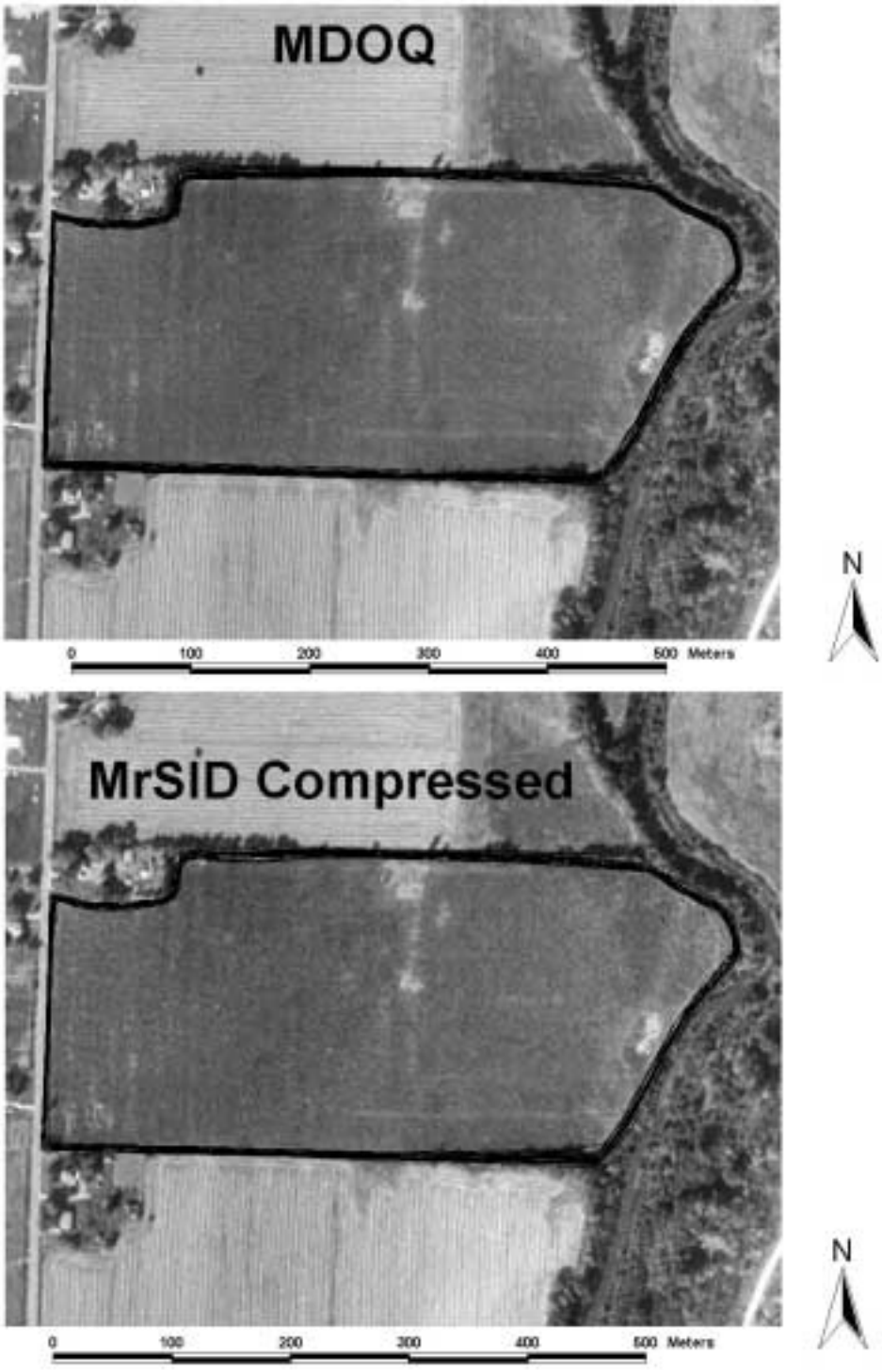
Average perimeter variation

MDOQ ave. perimeter variation all tests
10.2 meters

MrSID ave. perimeter variation all tests
7.3 meters

Units: Meters
Acreages determined using these conversion values:
1 sq. meter = 1.1960292 sq. yards
1 acre = 4840 sq. yards
Digitizing done at a scale of 1:4800
Digitizing done with ArcView3.1 UNIX version
Images used for digitizing: MDOQ M3909548
mosaic_1-1_s_ks103
Images are of Leavenworth County, KS
UTM coordinates of field used in digitizing test: 318148
4356689

Digitized polygons for the MDOQ and compressed county imagery



Disclaimer: The mention of software or products in this report does not constitute an endorsement or guarantee by the USDA, FSA, or APFO over other comparable products.