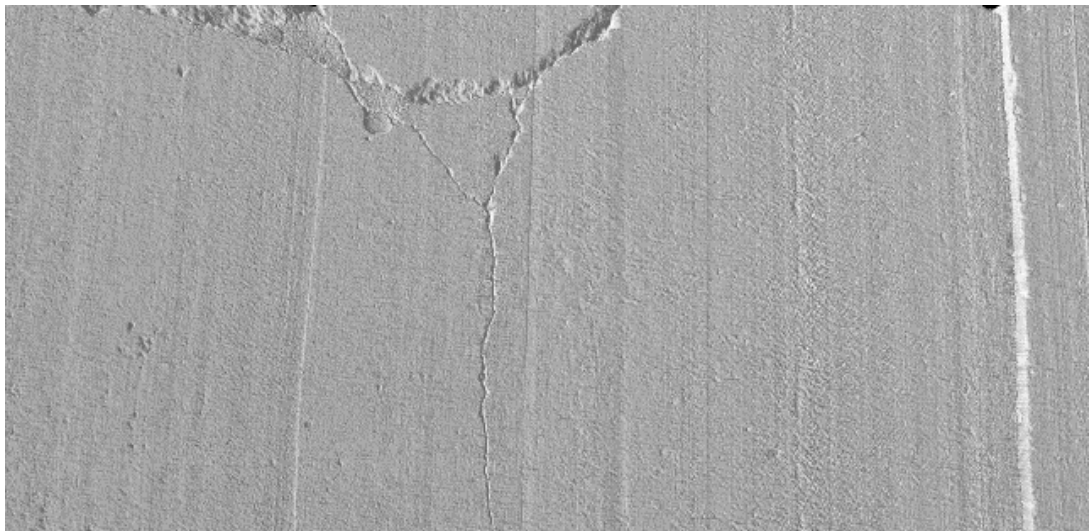




U.S. Department  
of Transportation  
**Federal Highway  
Administration**

# **Development of Standard Data Format for 2-Dimensional and 3-Dimensional (2D/3D) Pavement Image Data that is used to determine Pavement Surface Condition and Profiles**

## **Task 3 - Evaluate Data Items and Formats**



**Office of Technical Services  
FHWA Resource Center  
Pavement & Materials  
Technical Services Team**

**December 2016**

## **Notice**

This document is disseminated under the sponsorship of the U.S. Department of Transportation in the interest of information exchange. The U.S. Government assumes no liability for the use of the information contained in this document. This report does not constitute a standard, specification, or regulation. The U.S. Government does not endorse products or manufacturers. Trademarks or manufacturers' names appear in this report only because they are considered essential to the objective of the document, the transfer of technical information.

## **Quality Assurance Statement**

The Federal Highway Administration (FHWA) provides high-quality information to serve Government, industry, and the public in a manner that promotes public understanding. Standards and policies are used to ensure and maximize the quality, objectivity, utility, and integrity of its information. The FHWA periodically reviews quality issues and adjusts its programs and processes to ensure continuous quality improvement.

<b>1. Report No.</b>	<b>2. Government Accession No.</b>	<b>3. Recipient's Catalog No.</b>
<b>4. Title and Subtitle</b> Development of Standard Data Format for 2-Dimensional and 3-Dimensional (2D/3D) Pavement Image Data that is used to determine Pavement Surface Condition and Profiles		<b>5. Report Date</b> 12-12-2016
<b>7. Author(s)</b> Wang Kelvin C. P., Qiang "Joshua" Li, and Cheng Chen		<b>6. Performing Organization Code</b>  <b>8. Performing Organization Report No.</b>
<b>9. Performing Organization Name and Address</b> Oklahoma State University (OSU) 207 Engineering South Stillwater OK 74078		<b>10. Work Unit No. (TRAIS)</b>
<b>12. Sponsoring Agency Name and Address</b> Federal Highway Administration Office of Technical Services - FHWA Resource Center Howard Street, Suite 4000 Baltimore, MD 21201		<b>11. Contract or Grant No.</b> DTFH6115P00103
<b>15. Supplementary Notes</b> Mergenmeier, Andrew, FHWA, contract COR.		<b>13. Type of Report and Period Covered</b>
<b>16. Abstract</b>  This project was funded by the Transportation Pooled-Fund study, TPF-5(299). This report summarizes the contract Task 3 work.  This report documents the work performed by the research team on the evaluation of the commonly collected data items and existing data formats for pavement surface condition and profiles, and recommends the desired data items to be included in the proposed standard data formats for both 2D and 3D pavement image data.		<b>14. Sponsoring Agency Code</b>
<b>17. Key Words</b> Standard Data Format; 2D/3D; Pavement Image Data	<b>18. Distribution Statement</b> No restrictions. This document is available from the National Technical Information Service, Springfield, VA 22161	



## SI\* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS				
SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
<b>LENGTH</b>				
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
<b>AREA</b>				
in <sup>2</sup>	square inches	645.2	square millimeters	mm <sup>2</sup>
ft <sup>2</sup>	square feet	0.093	square meters	m <sup>2</sup>
yd <sup>2</sup>	square yard	0.836	square meters	m <sup>2</sup>
ac	acres	0.405	hectares	ha
mi <sup>2</sup>	square miles	2.59	square kilometers	km <sup>2</sup>
<b>VOLUME</b>				
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft <sup>3</sup>	cubic feet	0.028	cubic meters	m <sup>3</sup>
yd <sup>3</sup>	cubic yards	0.765	cubic meters	m <sup>3</sup>
NOTE: volumes greater than 1000 L shall be shown in m <sup>3</sup>				
<b>MASS</b>				
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
<b>TEMPERATURE (exact degrees)</b>				
°F	Fahrenheit	5 (F-32)/9	Celsius	°C
<b>ILLUMINATION</b>				
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m <sup>2</sup>	cd/m <sup>2</sup>
<b>FORCE and PRESSURE or STRESS</b>				
lbf	poundforce	4.45	newtons	N
lbf/in <sup>2</sup>	poundforce per square inch	6.89	kilopascals	kPa
APPROXIMATE CONVERSIONS FROM SI UNITS				
SYMBOL	WHEN YOU KNOW	MULTIPLY BY	TO FIND	SYMBOL
<b>LENGTH</b>				
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
<b>AREA</b>				
mm <sup>2</sup>	square millimeters	0.0016	square inches	in <sup>2</sup>
m <sup>2</sup>	square meters	10.764	square feet	ft <sup>2</sup>
m <sup>2</sup>	square meters	1.195	square yards	yd <sup>2</sup>
ha	hectares	2.47	acres	ac
km <sup>2</sup>	square kilometers	0.386	square miles	mi <sup>2</sup>
<b>VOLUME</b>				
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m <sup>3</sup>	cubic meters	35.314	cubic feet	ft <sup>3</sup>
m <sup>3</sup>	cubic meters	1.307	cubic yards	yd <sup>3</sup>
<b>MASS</b>				
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
<b>TEMPERATURE (exact degrees)</b>				
°C	Celsius	1.8C+32	Fahrenheit	°F
<b>ILLUMINATION</b>				
lx	lux	0.0929	foot-candles	fc
cd/m <sup>2</sup>	candela/m <sup>2</sup>	0.2919	foot-Lamberts	fl
<b>FORCE and PRESSURE or STRESS</b>				
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in <sup>2</sup>

\*SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380.  
(Revised March 2003)



## TABLE OF CONTENTS

<b>1. INTRODUCTION.....</b>	<b>1</b>
<b>2. ASSESSMENT OF CURRENT PRACTICES.....</b>	<b>2</b>
<b>3. RECOMMENDATION OF DESIRED DATA ITEMS AND DATA FORMATS.....</b>	<b>3</b>
Introduction.....	3
Assessment of Existing File Formats .....	3
Recommended File Structure.....	4
Recommended Data Items .....	7
Additional Discussions .....	9
<b>4. CONCLUSIONS.....</b>	<b>11</b>
<b>REFERENCES .....</b>	<b>12</b>

## LIST OF FIGURES

Figure 3.1 Layout of the File Structure .....5

## LIST OF TABLES

Table 3.1 Metadata Example .....7  
Table 3.2 Metadata Entries (Partial Listing).....7  
Table 3.3 File Trailer.....7  
Table 3.4 Recommended Data Items .....8



## 1. INTRODUCTION

Task 3 of this project calls for the assessment of existing data items collected and data formats of pavement image data, which includes three sub-tasks in the original proposal:

- *Subtask 3.1 Identify and Select Best Practices of Image Data Collection:* Based on the literature review and interviews with TPF-5(299) participating State departments of transportation (DOTs), the detailed image data collection practices of the selected DOTs will be obtained and the data items collected will be evaluated. Any existing relevant data format standards will be assessed.
- *Subtask 3.2 Evaluate Data Items Collected and Data Formats:* The detailed image data collection practices including the data items collected and their data formats will be obtained from the selected DOTs representing best practices and investigated. Finally, the inclusion of data items into the common data format will be developed in consultation with the TPOC, the panel, the participating State DOTs, and the industry.
- *Subtask 3.3 Prepare and Submit Interim Report*

This report documents the work performed by the research team for Task 3 of this project. In this Task, the team evaluates the common data items collected and existing data formats for pavement surface condition and profiles, and recommends the desired data items to be included in the standard data formats for both 2D and 3D pavement image data. Because Task 3 is closely tied to and conducted based on the literature review and the survey of the TPF-5(299) States and relevant industry vendors (Task 2), many findings and assessments that were presented in the submitted Task 2 report but relevant to the work in Task 3 are not included in this report. In addition, due to the intertwining nature of Task 3 and Task 4, some of the deliverables of Task 4 are included in the report for Task 3. The report is organized as follows:

- Chapter 1 outlines the sub-tasks included in Task 3 of this project;
- Chapter 2 summarizes the assessment of the current practices of image data collection; and,
- Chapter 3 recommends the desired data items to be included in the standard data formats for both 2D and 3D pavement image data.

## **2. ASSESSMENT OF CURRENT PRACTICES**

Based on the survey conducted by the Oklahoma State University (OSU) research team, 2D and 3D data sets are used as the primary visual representations of pavement surface data in the industry of pavement surface data collection. It is further determined that data collection of 2D data is generally based on the JPEG format within the 8-bit grey-scale range. The data storage of 3D data with systems provided by the limited number of technology suppliers in the US is based on a dynamic range between 12-bits and 16-bits. It is assumed that for the sake of covering all possibilities 16-bit range is used for 3D pixels for this project.

The industry practices on how to compress and encode the 2D, 3D, and other metadata sets for pavement surfaces vary greatly and are considered proprietary. The research team has not been able to obtain specific data compression and encoding methods for pavement data from other technology suppliers and DOT users. However, such information is not critical for the project delivery and its subsequent implementations. This is because the data encoding format of 2D and 3D data sets once defined can be implemented through writing a relatively small software program without technical difficulty.

Based on the development of the demonstration software for pavement 2D and 3D image data encoding and decoding during the fall project meeting at the 2015 Road Profile Users' Group (RPUG) conference, the OSU research team recommends that through this project a standard data format be defined for 2D data, 3D data, and metadata sets that can be conveniently implemented by technology suppliers or users. Therefore, other specific technical work items described in the proposal for Sub-task 3.1 are not documented in the report.

### 3. RECOMMENDATION OF DESIRED DATA ITEMS AND DATA FORMATS

#### Introduction

Based on the literature review and survey conducted in Task 2, and subsequent discussions with the FHWA Technical Point of Contact (TPOC) and the project panel, existing data items and data formats are evaluated in this chapter. Specifically, the file structures of existing data formats are overviewed to assess the suitability for application across multiple platforms and the development of standard data formats for both 2D and 3D pavement image data for this project. Subsequently, the desired or recommended data format structures and data items are presented.

#### Assessment of Existing File Formats

##### General vs. Specific File Format

Based on the literature review conducted in Task 2, the existing file formats can be classified as general formats and specific formats. The general formats are designed with high flexibility to accommodate as many types of data as possible for a wide range of general applications. Example formats include the ASTM standard E57 (Huber, 2011) and the Universal 3D (U3D) (ECMA, 2007). However, the requirement of highly flexible formats may result in very complex file definitions and data structures with specifically designed libraries for data encoding and decoding.

The specific formats are designed for specific applications. This format can be designed in a relatively simple and effective manner since the data items are well-organized in structure. Example file formats include the File Exchange Format (FEF) proposed by Tsai and Wang (2014) for the storage of line laser imaging data.

Considering pavement image data contains only organized "grid" data in the form of a matrix of points (pixels), specific data format is recommended for pavement 2D/3D imaging data to facilitate optimal data representation and compression methods. Meanwhile, this specific data formats can be supplemented by integrating the principles in the general file formats. For example, as adopted in the E57 format, an open source reference library can be used that is capable of reading, writing, and validating files in the TPF-5(299) standard file format.

##### Self-Contained File Format

Generating and collecting large data sets are becoming a necessity in many domains that also need to keep the data for long periods and across multiple platforms (Simona et al., 2014). Examples include genomics, medical records, astronomy, atmospheric science, photographic archives, video archives, large-scale e-commerce, pavement imaging data sets, etc. While this presents significant opportunities, a key challenge is the creation of vendor-neutral storage containers that can provide economically scalable storage systems and be interpreted over time to preserve the data in the long run. This includes both the data itself as well as semantic metadata necessary to enable search, access, and analytics on that data for future usage and analysis. For this purpose, Simona et al. (2014) recommends that the data format should be self-describing (namely it can be interpreted by different systems and in different points in time) and self-contained (all data needed for the preservation objects' interpretation is contained in the container). The metaphor is a closed bottle that includes all the information needed to understand the bottle's contents at a future point in time (Simona et al., 2014).

However, among the existing data formats, many of which do not follow the rule of being self-contained. For example, the XML-based “ElementaryData” format (Gajewski, 2012) itself does not content the original data, but only provides the information of the folder where the photographic documentation is saved. Another example is the FEF format (Tsai and Wang, 2014), where metadata and original data are stored in two separated files: an index file and corresponding data frame files respectively. To be self-contained, the information in the index file and data frame file should be kept within a single file.

Therefore, for this project it is recommended a self-contained data format be used, where all metadata and raw image data are grouped into one single file, so that users can share individual pavement images as convenient as sharing photos from their family album. This self-contained design will make the following data analysis and management straightforward.

### **Recommended File Structure**

A data file consists of many data items. The file structure defines how these data items are organized and stored into a file. Therefore, a file structure is a crucial component for a data format standard. A desirable file structure should be “broadly transferrable” with backward compatibility and forward compatibility. For example, the Europe HiSPEQ project team (2015) suggests the “broadly transferrable” concept in its data format, which means that the core components of the format remain the same whilst having some customization for each administration.

As pavement imaging technology will continue to evolve, vendors will foster innovation by developing better image data collection and processing methods, and agencies will also continue to require new capabilities. It is important to incorporate the “broadly transferrable” concept into this TPF-5(299) pavement image data standard and to design pavement image file format with both backward and forward compatibility.

In order to obtain these desired proprieties, a five-section file structure is proposed, which can be further grouped into two parts. As shown in the Figure 3.1, the first part consists of the core sections including File Header, 2D Image Data, 3D Range Data, and Right of Way Data (optional). These data items defined in the core sections are not supposed to be changed in the future. The second part primarily contains the optional metadata section, which can be extended and modified as needed by vendors or highway agencies to accommodate the needs from individual data collection practice and equipment. All five sections can be written into one single file sequentially as shown in Figure 3.1.

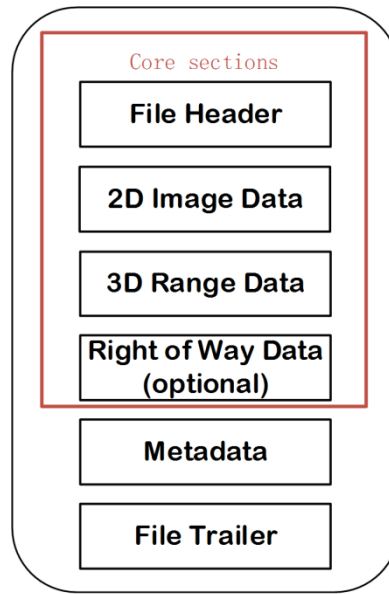


Figure 3.1 Layout of the File Structure

### File Header Section

The file header identifies and declares the core properties of the 2D/3D data stored in a file. Each property is represented by a data item, which will be denoted by a variable. The variable will have a predefined data type with fixed byte length. These variables can be classified into two categories, data reading required and fast archiving/retrieving required:

- Data reading required: such as File format version, image size, resolution, bit depth, compression algorithm name, storage offset, and so on,
- Fast archiving and retrieving required: such as location and time.

The first category of data items (variables) is critical for successful 2D/3D data reading. The second category of data items is specially designed for fast archiving and retrieving in database-based pavement data management.

The header of an image file may be encoded in either ASCII or binary format. For ASCII text format, each header entry specified in a file, may be separated using new line label, such as “\n” or be listed in one line. Even though both binary and ASCII text file formats have been used in existing data formats, the binary file format has at least two distinctive advantages:

- it occupies less storage; and
- it is faster to read (significant less time-consuming for string parsing).

Efficient storage space and fast reading/writing speed are highly desired properties for the pavement data storage, process, analysis and management; the binary format is therefore recommended.

### 2D Intensity Data and 3D Range Data Sections

Pavement surface visual data sets include 2D and 3D. The 2D data captures pavement intensity information, while 3D range data provide depth information of pavement surfaces. Both 2D intensity and 3D range data will be the key data components in the data format

standard. Considering some vendors or agencies are depending only on 2D or 3D systems while others implement both of them, the desired file format should be able to handle all possible scenarios.

A specific implementation issue is the data compression for both 2D and 3D data sets. It is recommended that in the file header specific data items are assigned to identify the compression algorithms for 2D and 3D data. Further in the metadata section, optionally, technology suppliers can provide their own detailed information for data compression, such as their recommended compression libraries used for the 2D or 3D data compression. Additionally, the metadata section can also be used to provide coding algorithms for compression if a technology supplier or user is willing to share such information.

In addition, a software library in the form of Dynamic Link Library (DLL) can be provided by the industry vendors and users so that their proprietary compression and decoding algorithms can be included in the data format for metadata. However, this particular feature is not required in the project deliverables by the OSU team.

#### Right-Of-Way (ROW) Image Section (Optional)

Right-Of-Way (ROW) images can provide important visual reference information for pavement image data. Moreover, the collection of ROW images in JPEG format has a long history for many highway agencies. The ROW image section is recommended by the research team to be optional, which can be selected based upon the needs of vendors or highway agencies, or left empty. JPEG, which is predominantly used for the storage of ROW images, is recommended as the standard image format. However, it was later agreed by the FHWA TPOC, project panel, and the research team to exclude ROW data from the standard data format since these data sets are out of the scope of this project.

#### Metadata Section (Optional)

Metadata is structured, descriptive information about a resource, or data about data (Nogueras-Iso et al., 2004). The proposed file header will contain the critical metadata about the pavement image data in a fixed-data size. In order to obtain some desired properties, such as “broadly transferrable” property with backward compatibility and forward compatibility, an optional metadata section is highly recommended, which can be extended and modified as desired by vendors or highway agencies to accommodate the needs from individual data collection practice, equipment provider, or a user.

Various metadata data formats are available, including Computer Graphics Metafile (CGM), Enhanced Metafile (EMF), Windows metafile (WMF), “IPTC Core” Schema for Extensible Metadata Platform (XMP), Dublin Core Metadata Initiative (DCMI), and Picture Licensing Universal System (PLUS) (Riley and Becker, 2010), which can be generally implemented in two types of approaches:

- Rigid format - employs fixed-size fields and records. The fields are frequently populated with special codes that can be interpreted only through the use of external lookup tables. The Weigh-In-Motion (WIM) data formats defined in the FHWA Traffic Monitoring Guide (FHWA, 2001) is an example of this approach.
- The second type of format uses flexible, self-documenting structures with variable lengths and structure. The Extensible Markup Language (XML) standard is an example of this implementation.

The first approach is simple to understand and implement, but has the disadvantage of not being flexible. The second approach is more flexible, but can lead to increased implementation complexity. As a result, hybrid methods combining the merits of both approaches have been developed. For example, the standard pavement profile data format defined in the ASTM E2560-13 (*Standard Specification for Data Format for Pavement Profile*) implements hybrid metadata definitions and syntax. Based on a comprehensive review of the ASTM E2560-13 standard and the team’s experience in developing the standard, it is found that the technical methods in the ASTM E2560-13 format and syntax meet the needs of the storage of pavement image metadata, and therefore are recommended for this project.

Specifically, the first value in the Metadata portion in Figure 3.1 provides the number of metadata entries (MDE) as shown in Table 3.1. Table 3.2 shows the information to construct an example MDE. The specific metadata data entries and their tags can be designed to meet various specific needs from the needs of vendors or highway agencies.

Table 3.1 Metadata Example

Variable Name	Data Type	Data
Number of MDEs	Int32	Number of MDEs

Table 3.2 Metadata Entries (Partial Listing)

Variable Name	Data Type	Data
Tag of MDE	Int32	Metadata tag
Data type of MDE	Int32	Data type index of MDE
Array size	Int32	“-1” if not an array. “0” if array is empty. Numbers greater than 0 specify the number of elements in the array
Count	Int32	For data types “String” and “Array (String)”, count = the number of bytes in the string. For other data types, count = 1.
Name length	Int32	For predesigned metadata entries, this is 0. For user-defined entries, this value is the length of the name.
Name	String	Name of the metadata
MDE	varies	Information associated with tag of MDE

### File Trailer Section

The file trailer is used to signal the end of the file (Table 3.3).

Table 3.3 File Trailer

Variable Name	Data Type	Data	Default Value
End of file	3-byte String	Indicates the end of the file	“@@@”

### **Recommended Data Items**

Based on the survey results, feedback from industry vendors and highway agencies, and the hands-on experience gained from the development of writing the encoding and decoding software for 2D data, 3D data, and metadata sets, the following data items are recommended to be included in the standard data format as shown in Table 3.4.

Table 3.4 Recommended Data Items

Variable Name	Data Type	Data Details
Version	4-byte String	Version number of the file format
SW version	4-byte String	Identifier of the software that produced the data file
State Name	2-byte String	FIPS State Code
Route Name	12-byte String	Name of the highway in HPMS standard
Direction	1-byte String	Direction of travel
Lane identification	2-byte String	Lane index
File Serial Number	Int32	File serial number in continuous data collection
GPS Longitude	Float 32bit	GPS longitude value IEEE 754 binary32
GPS Latitude	Float 32bit	GPS latitude value: IEEE 754 binary32
DMI Pulse	Int32	Distance Measuring Instrument (DMI) pulse counting index
Date	8-byte String	Date data was collected—(yyyymmdd)
Time	6-byte String	Time data was collected—(hhmmss)
Time stamp	Long long int	Milliseconds since UNIX Epoch: Jan 1, 1970 00:00:00
Event Mark ID	Int32	Marker of special data starting point or ending point
2D Data Offset	Int32	Offset in bytes from the beginning of the file to the beginning of the 2D data
2D Data Size	Int32	The compressed 2D data size
3D Data Offset	Int32	Offset in bytes from the beginning of the file to the beginning of the 3D data
3D Data Size	Int32	The compressed 3D data size
ROW Data Offset	Int32	Offset in bytes from the beginning of the file to the beginning of the ROW data in JPEG format
ROW Image Size	Int32	The compressed ROW Image size
Metadata Offset	Int32	Offset in bytes from the beginning of the file to the beginning of the metadata
2D Compression Method	4-byte String	Identifies compression algorithms, such as 1: PNG; 2: JPEG; 3:JPEG2000.
2D Compression Quality Index	Float 32bit	Compression quality level
2D Resolution Longitude Direction	Float 32bit mm	Distance between two data rows in longitude direction in millimeters.
2D Resolution Transverse Direction	Float 32bit mm	Distance between two data columns in transverse direction in millimeters
2D Width	Int32	Pixel numbers in transverse direction
2D length	Int32	Pixel numbers in longitude direction
2D Data Bit Depth	Int32	The bit depth for each data point
3D Compression Method	4-byte String	Identifies compression algorithms, such as 1: PNG; 2: JPEG; 3:JPEG2000
3D Compression Quality Index	Float 32bit	Compression quality level
3D Resolution Longitude Direction	Float 32bit mm	Distance between two data rows in longitude direction in millimeters
3D Resolution Transverse Direction	Float 32bit mm	Distance between two data columns in transverse direction in millimeters
3D Resolution Elevation Direction	Float 32bit mm	Units for range data value in millimeters
3D Width	Int32	Pixel numbers in transverse direction
3D length	Int32	Pixel numbers in longitude direction
3D Data Bit Depth	Int32	The bit depth for each data point
Speed	Float 32bit mm/s	Average vehicle speed associated with data
Vehicle Number	8 byte String	Vehicle identification



Variable Name	Data Type	Data Details
Operator Name	32 byte string	Operator identification
Reserved Item1	4 byte string	For the future & vendor specific
Reserved Item2	Int32	For the future & vendor specific
Reserved Item3	Float 32bit	For the future & vendor specific
Reserved Item4	8 byte String	For the future & vendor specific
File Header Trailer	4 byte String	For error checking of file header

## Additional Discussions

### Thumbnails Preview

Thumbnail preview is useful for fast searching and skim of images with dramatic scene changes, which has been provided in JPEG files in Windows Operating Systems (OS) and specialized software. The common sizes of thumbnails vary from 75x75, 100x100, 125x125, 150x150, 200x200 400x400 to 600x600. However, since pavement surfaces do not generally have distinctive features, down-sampling full lane width 2D or 3D pavement images into thumbnails with small sizes may fail to provide visually distinguishable images for fast searching. Increasing the thumbnail image size can help, but at the cost of extra storage and high data loading time. In addition, since pavement 2D and/or 3D images are saved using custom file types, special programming is needed to coerce Windows file explorer to automatically show a thumbnail for each custom file. The Microsoft Shell Extension Handler technique (Ackley, 2005) could be implemented to develop the thumbnail-based preview software. As a result, the research team does not recommend the thumbnail preview function to be included in the pavement image data format, but in the future development of pavement viewer software.

### JPEG Compression

Based on the survey results, JPEG is the most widely used compression algorithm, which trades image quality for compression, for the compression of 2D intensity image and 3D range data sets. The JPEG compression is typically controlled by a percentage of the quality level, which controls the amount of loss in visual quality during the compression and the size of the compressed file. Based on the team's extensive testing and the requirements of precision, accuracy, and tolerances within highway agencies, and due to the lossy nature of JPEG compression algorithms which would result in loss of pavement texture details, lossless algorithms are recommended for 3D pavement surface data.

### GPS Data

Several comments on the GPS spatial data are obtained from the survey performed in Task 2. Many GPS latitude/longitude data formats are used by highway agencies including:

- IEEE-754 FP signed degrees format - "±DDD.ddd" preceded South latitudes and West longitudes with a minus sign (such as 36.125564, -97.073851);
- Degrees or radians stored as a fixed point value in an 32 or 64 bit integer (such as 361255640, -970738510);
- A fixed structure format in the form of +xxx.xxxxxxxx or -xxx.xxxxxxxx (such as +036.12556400, -097.0738510);
- "DMS + compass direction" formats that use degrees, minutes, and seconds along with symbols for compass directions (N, S, E or W) (such as 41°25'01"N, 120°58'57"W).

Considering the IEEE 754 Single-precision floating-point is one of the most common used data types in various programming languages such as C++/C, and SQL/MySQL, it is recommended using IEEE-754 FP float for the storage of GPS data with 6 to 9 decimal digits.

#### Compression Library and Demo Software

A compression library can be defined in the Metadata section of the data format so that a technology supplier or user can develop and use their own compression library as a software implementation that can be directly applied by a software coder who writes a utility to convert a 2D and/or 3D data sets into the standard data format. This conversion can be implemented in any data vehicle, or as a post-processing capability. The Metadata can either describe the algorithms of encoding and decoding, or identify a software library file that directly conducts encoding and decoding for 2D and 3D data sets, such as a DLL file. A demo software would be helpful to technology providers and users who would like to test their data format, understand the technical process, and eventually develop their own implementations.

Both the compression library file and the demo software are not defined in the deliverables for the project.

## **4. CONCLUSIONS**

In Task 3, the research team examined the data formats and practices of the industry and recommended the general data format for the 2D and 3D pavement image data. Due to the proprietary nature of data formats in the pavement industry, the team was not successful in getting the specific and technical details of those data formats. The described approach in Task 3 would satisfy the needs of the participating States per the final project deliverables.

Software delivery is not part of the project deliverables. It is therefore recommended that a suitable compression library and demo software be developed in a separate project along with future support services, so that the industry and users would have a clearer and more efficient path to test and adopt the standard data format recommendations.

## REFERENCES

- Ackley Ryan. Microsoft Office Thumbnails in SharePoint. Microsoft, 2005.  
[https://msdn.microsoft.com/en-us/library/aa289172\(VS.71\).aspx](https://msdn.microsoft.com/en-us/library/aa289172(VS.71).aspx), accessed on Feb 25, 2016.
- ECMA International. Standard ECMA-363: Universal 3D File Format. Rue du Rhône, Geneva, 2007. <http://www.ecma-international.org/publications/files/ECMA-ST/ECMA-363%204th%20Edition.pdf>
- Gajewski Wojciech. “ElementaryData” format – way to high quality of road condition measurement data: application results. Proceeding of 4th European Pavement and Asset Management (EPAM) conference, Sweden, 2012.
- HiSPEQ Project Team. HiSPEQ: Guidance for Road Administrations for Specifying Network Surveys. Conference of European Directors of Roads (CEDR). National Road Administrations of Denmark, Germany, Ireland, Netherlands, UK and Slovenia, 2015.
- Huber Daniel. The ASTM E57 file format for 3D imaging data exchange. Proceeding of SPIE Vol 7864: Three-Dimensional Imaging, Interaction, and Measurement. 2011.
- Nogueras-Iso J., F.J. Zarazaga-Soria, J. Lacasta, R. Béjar, P.R. Muro-Medrano. Metadata standard interoperability: application in the geographic information domain. Computers, Environment and Urban Systems, Volume 28, Issue 6, November 2004, PP: 611–634.
- Riley Jenn and Devin Becker. Glossary of Metadata standards Indiana University Libraries, 2010.  
[http://www.dlib.indiana.edu/~jenlrile/metadatamap/seeingstandards\\_glossary\\_pamphlet.pdf](http://www.dlib.indiana.edu/~jenlrile/metadatamap/seeingstandards_glossary_pamphlet.pdf), accessed on Feb 25, 2016.
- Simona Rabinovici-Cohen, Roger Cummings and Sam Fineberg, "Self-contained Information Retention Format For Future Semantic Interoperability", Proceedings of the 4th International Workshop on Semantic Digital Archives (SDA), September 2014, London, UK.
- Tsai Yichang (James) and Zhaohua Wang, A Remote Sensing and GIS-based Asset Management System (RS-GAMS) Phase 2 Final Report. Contract Number: RITARS-11-H-GAT, April 2014.