FamilySearch Digital Image Standard: A Key to Quality

Richard J. Laxman; FamilySearch International; Salt Lake City, Utah/USA

Abstract

FamilySearch is a non-profit organization that captures images and metadata from documents in archives worldwide and hosts them for individuals doing genealogical research. FamilySearch developed a "Digital Pipeline" to capture 50 million images and metadata per year from original records. The pipeline also includes scanning 65 million images from the existing 2.4 million roll microfilm collection. These images and metadata are processed, cataloged, indexed, hosted and preserved. FamilySearch developed hardware and software such as a digital camera and microfilm scanning systems that are used in the Digital Pipeline. The organization develops and implements internal standards for images and metadata. These standards are presented to AIIM/ANSI (Association for Information and Image Management/American National Standards Institute) and ISO (International Organization for Standards) for consideration as national and international standards.

This session will explain how FamilySearch developed and uses its Digital Image Specification to produce quality images for its Digital Pipeline and why quality standards are important for an imaging process. Participants in this presentation will learn about the characteristics of a digital image such as tonal range, contrast and tonal and spatial resolution. The importance of each characteristic as it relates to image quality will be explained. The tools FamilySearch uses to create and implement its Digital Image Specification and measure quality as part of the organization's Digital Pipeline also will be explained and demonstrated.

The benefits of creating high-quality images according to standards vs. the consequences of using lower-quality images in an imaging process can be clearly demonstrated in a digital process. The presenter will show those benefits and consequences encountered throughout the Digital Pipeline processes. The importance of using quality standards especially when dealing with older, damaged or faded documents will be shown and become particularly evident for any digital process. A discussion on why image creation using quality standards is absolutely necessary for the preservation of those images will be included in the presentation.

FamilySearch

FamilySearch, formerly the Genealogical Society of Utah, is a non-profit organization that gathers records for the purpose of genealogical research. Today this is done using a combination of digital camera systems and microfilm camera systems developed by FamilySearch. The microfilm is subsequently scanned creating digital surrogates.

Digital Pipeline

FamilySearch developed a Digital Pipeline to create digital images from microfilm scanning and from original documents captured with digital cameras. See Figure 1. Included in this pipeline is the receipt and processing of images, waypointing and indexing of images, hosting and distribution of images and metadata and preservation of images and metadata. Quality and standards play a prominent role throughout the Digital Pipe line. Figure 1 shows the Digital Pipeline developed by FamilySearch.

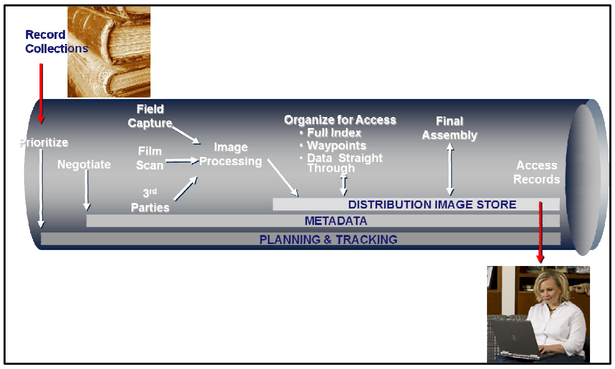


Figure 1. *FamilySearch Digital Pipeline*

Purpose of Standards

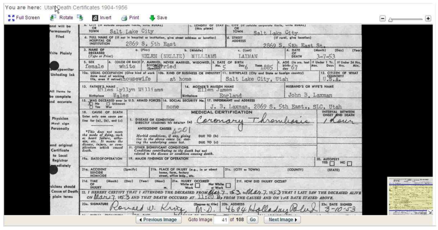
The purpose of standards is to evaluate the overall quality of delivered work and products. These services and products may be delivered commercial enterprises, and governmental entities, libraries, cultural heritage institutions, etc. Standards also promote interchangeability and quality control of products and can eliminate misunderstandings or confusion between service providers and users. Standards can benefit domestic and international trade, communication and understanding.
Standardization applies to terminology, definitions, sizes, formats, quality, methods of measurement and procedures for the creation, use, storage and retrieval and preservation of structured and unstructured documents, records and related metadata.

Purpose of FamilySearch Standards

The original records imaged by FamilySearch are found in various states of organization and deterioration and in many formats such as loose papers and bound material. FamilySearch has a program to digitize and preserve these records. FamilySearch started microfilming records in 1938. The quality of the exposures on the microfilm created prior to the development of microfilm quality standards is at times substandard. These two challenges, poor original documents and poor early microfilming quality, require that imaging of original records and microfilm be done according to standards that will ensure readable images.

The purpose of the FamilySearch Standard is to ensure readable facsimile of the original document vs. an actual representation or exact copy of the document itself. To this end the Digital Image standard was developed to get from the source records (Figure 2) to the readable images (Figure 3).

Figure 2. FamilySearch source records for digital imaging

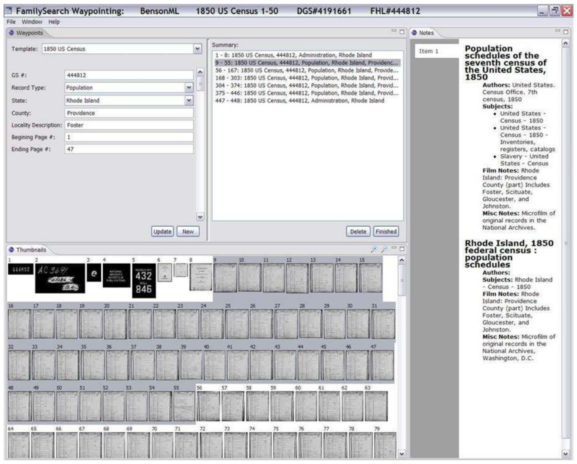


You are here: Utah Death Certificates 1904-1956

CITY OR TOWN: Salt Lake City
CITY OR TOWN: Salt Lake City
FULL NAME OF HOSPITAL OR INSTITUTION: 2869 S. 5th East
STREET ADDRESS: 2869 S. 5th East St.
3. NAME OF DECEASED: HELEN (NELLIE) WILLIAMS LAXMAN
4. DATE OF DEATH: 3-7-53
5. SEX: female
6. COLOR OR RACE: white
7. MARRIED, NEVER MARRIED, WIDOWED, DIVORCED: married
8. DATE OF BIRTH: Mo. 5 Day 6 Year 1885
9. AGE: 67
10a. USUAL OCCUPATION: housewife
10b. KIND OF BUSINESS OR INDUSTRY: at home
11. BIRTHPLACE: Salt Lake City, Utah
12. CITIZEN OF WHAT COUNTRY? U.S.A.
13. FATHER'S NAME: Miles Lyllyn Williams; Birthplace: Wales
14. MOTHER'S MAIDEN NAME: Ellen James; Birthplace: England
HUSBAND'S OR WIFE'S NAME: John B. Laxman
16. SOCIAL SECURITY No.: none
17. INFORMANT and ADDRESS: J. B. Laxman, 2869 S. 5th East., SLC, Utah
18. CAUSE OF DEATH: Coronary Thrombosis
INTERVAL BETWEEN ONSET AND DEATH: 1 hour
22. I HEREBY CERTIFY THAT I ATTENDED THE DECEASED FROM Mar 7, 1953 to Mar 7, 1953 THAT I LAST SAW THE DECEASED ALIVE ON Mar 7, 1953 AND THAT DEATH OCCURED AT 11:20 P.M. FROM THE CAUSES AND ON THE DATE STATED ABOVE.
23a. SIGNATURE: Ronald W. King, M.D.
23b. ADDRESS: 4694 Holladay Blvd
23c. DATE SIGNED: 3-10-53

Figure 3. FamilySearch processed image

Other functions in the Digital pipeline, Waypointing and Internet Indexing, are dependent on adherence to the standard for quality image creation. Individuals involved in Waypointing, the grouping of images by logical beginning and end points such as locality, need focused, properly cropped image to tag them using thumbnail derivatives. Field identification templates are created for document types delivered to volunteers involved in the Internet Indexing program. If images are skewed, the templates will not line up and accurately point the indexer to field to be extracted from the document in the image. Figures 4 and 5 are samples of the Waypointing and Internet Indexing applications respectively.



FamilySearch Waypointing: BensonML 1850 US Census 1-50 DGS#4191661 FHL#444812

Template: 1850 US Census
GS #: 444812
Record Type: Population
State: Rhode Island
County: Providence
Locality Description: Foster
Begining Page #: 1
Ending Page #: 47

Summary:
1 - 8: 1850 US Census, 444812, Administration, Rhode Island
9 - 55: 1850 US Census, 444812, Population, Rhode Island, Providenc...
56 - 167: 1850 US Census, 444812, Population, Rhode Island, Provide...
168 - 303: 1850 US Census, 444812, Population, Rhode Island, Provid...
304 - 374: 1850 US Census, 444812, Population, Rhode Island, Provid...
375 - 446: 1850 US Census, 444812, Population, Rhode Island, Provid...
447 - 448: 1850 US Census, 444812, Administration, Rhode Island

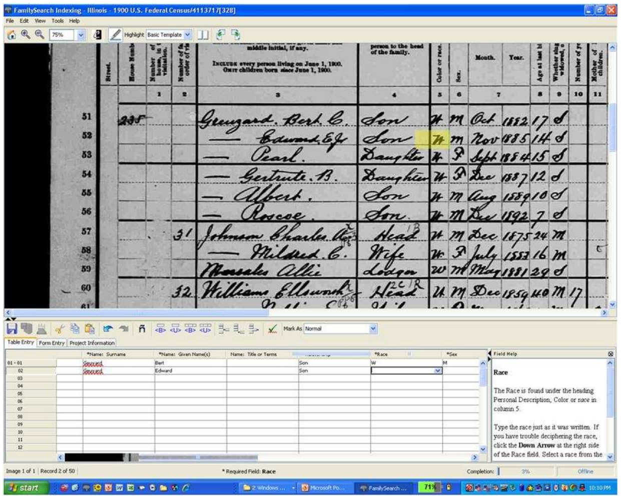
Item 1

Population schedules of the seventh census of the United States, 1850
Authors: United States. Census Office. 7th census, 1850
Subjects:
- United States - Census - 1850
- United States - Census - 1850 - Inventories, registers, catalogs
- Slavery - United States - Census

Film Notes: Rhode Island: Providence County (part) Includes Foster, Scituate, Gloucester, and Johnston.
Misc Notes: Microfilm of original records in the National Archives.

Rhode Island, 1850 federal census : population schedules
Authors:
Subjects: Rhode Island - Census - 1850
Film Notes: Rhode Island: Providence County (part) Includes Foster, Scituate, Gloucester, and Johnston.
Misc Notes: Microfilm of original records in the National Archives, Washington, D.C.

Figure 4. FamilySearch Waypointing application



FamilySearch Indexing - Illinois - 1900 U.S. Federal Census/4113717[328]

	*Name: Surname	*Name: Given Name(s)	Name: Title or Terms		*Race	*Sex
01 - 01	Geuyard	Bert		Son	W	M
02	Geuyard	Edward		Son		
03						
04						
05						
06						
07						
08						
09						
10						
11						
12						

Race

The Race is found under the heading Personal Description, Color or race in column 5.

Type the race just as it was written. If you have trouble deciphering the race, click the **Down Arrow** at the right side of the Race field. Select a race from the

Image 1 of 1 | Record 2 of 50 | * Required Field: **Race** | Completion: 3% | Offline

Figure 5. FamilySearch Internet Indexing application

The Digital Image standard also drives the requirements for the development and enhancement of the hardware and software created by FamilySearch. The requirement to meet the standard and still maintain high production necessitated that FamilySearch develop digital camera and microfilm scanning systems. See Figure 6.

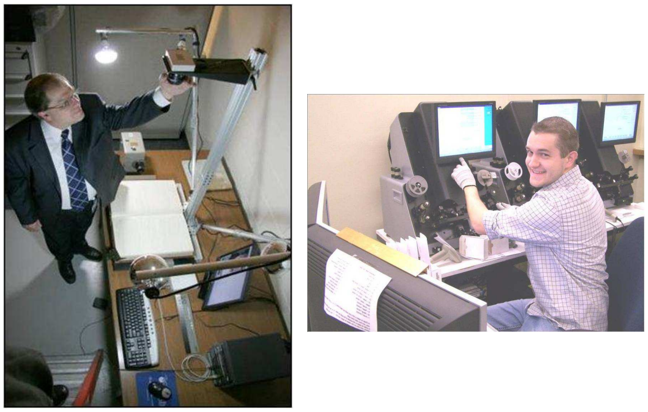


Figure 6. FamilySearch digital camera and microfilm scanning systems

The FamilySearch Digital Image standard

The Digital Image Research Team (DIRT) at FamilySearch selected the digital image characteristics to be included in the standard. The characteristics selected were deemed germane to the creation of the digital images that would meet requirements of all the functions in the Digital Pipeline and the needs of customers viewing the images on FamilySearch.org. The DIRT team is comprised of Pipeline operations managers, front-line Pipeline team members and imaging engineers.

The DIRT team selected the following characteristics to be included in the original Digital Image Specification standard:

Image Characteristics	
Tonal Range	Image File Format
Grayscale Contrast	Image Compression
Tonal Resolution	File Name
Even Exposure	Image Dimension
Spatial Resolution	File Size
Color Space	Complete Capture
Focus	Image Orientation
Blur	Image Skew

In addition to the above characteristics, three additional image characteristics are being added to the standard. Work should be completed on these additions by Q3 2010.

- Fixity
- Attribution
- Watermarking

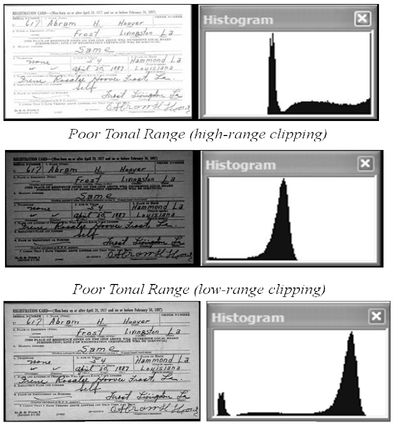
The remainder of this paper will present the definition of each characteristic, its desired attribute range, acceptable measure and implementation methods.

Tonal Range

Definition	The luminance range of a document from the deepest shadow (black) to the brightest highlight (white).
Desired attribute range	The darkest and brightest pixels of an image will be no closer than 3% to the extremes of the tonal range. Acceptable lower and upper pixel values at approximately 8 and 247, using an 8-bit tonal resolution (i.e. 0-255)
Measure	The tonal range of a digital image is measured by determining the pixel value of the brightest and darkest pixels

.

The first two images displayed in Figure 7 are examples of overexposed and underexposed images. The accompanying histograms graphically display the tonal range of the images being clipped on the upper and lower ends. The third image was properly exposed and the histogram is correctly positioned.



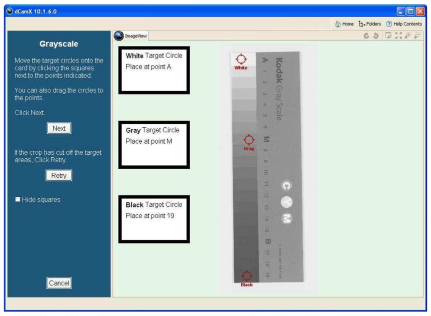
Poor Tonal Range (high-range clipping)

Poor Tonal Range (low-range clipping)

Good Tonal Range (no clipping)

Figure 7. *Incorrect and correct tonal range with histogram display*

The digital camera software, DCamX, automatically adjust the tonal range using a grayscale target shown in Figure 8.



Grayscale

Move the target circles onto the card by clicking the squares next to the points indicated.

You can also drag the circles to the points.

Click Next.

If the crop has cut off the target areas, Click Retry.

■ Hide squares

White Target Circle
Place at point A

Gray Target Circle
Place at point M

Black Target Circle
Place at point 19

Figure 8. *Tonal resolution calibration in DCamX software*

Grayscale Contrast

Definition	The ratio between the foreground and background luminance intensities in an image. The human eye is logarithmically sensitive to brightness, implying that for the same perception, higher brightness requires higher contrast.
Desired attribute range	Currently being reviewed and updated
Measure	Ratio between measured pixel value between foreground and background.

The grayscale contrast is measured by determining the ratio between pixel values of the foreground (writing) and the background. See Figure 9.

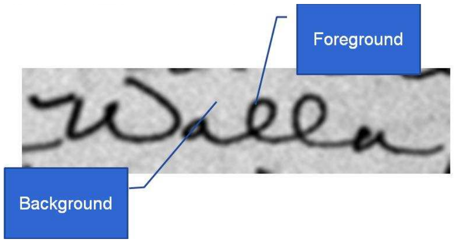


Figure 9. Grayscale contrast measurement

Tonal Resolution

Definition	Sometimes referred to as bits per pixel or bit depth, this value determines the range of intensities a pixel can assume. The greater the tonal resolution, the more possible intensities available at capture.
Desired attribute range	FamilySearch images possess 256 luminance levels per image channel used. Tonal resolution may be higher at capture (e.g., 12-bits, or 4,096 possible intensities) to allow the application of computer algorithms such as contrast stretching, exposure balancing, or auto-focus using more data, but these additional levels

	are eventually discarded as the human eye cannot detect them.
Measure	Each pixel of a digital image contains information to define its luminance (and hue, if a color image). In a grayscale image, the tonal resolution determines how many steps from minimum (pure black) to maximum (pure white) are defined for that luminance or, for color images, how many colors can be represented

The top grayscale strip in Figure 10 is divided into eight smaller sections, each slightly higher in intensity from left to right. The bottom grayscale strip is similar, but exhibits a higher tonal resolution—more pixel intensities are possible. Grayscale images produced by FamilySearch have over five times the tonal resolution of the bottom rectangle.

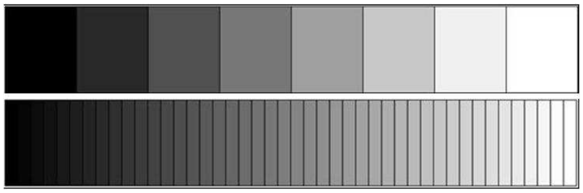
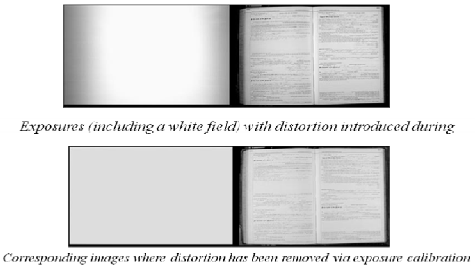


Figure 10. *Grayscale strips displaying tonal resolution*

Even Exposure

Definition	Even exposure is a condition where no distorted image intensities exist as a result of the capture system (lenses, CCD arrays, lighting, software, etc.).
Desired attribute range	No distortion over 0.5% in pixel luminance introduced by capture system.
Measure	Even exposure is measured by sensing variation in luminance intensity at the camera in order to identify where illumination is uneven. Areas of the captured image which vary in intensity represent an imbalance in a capture situation. This uneven illumination is measured by sampling and comparing pixel intensity across a document that has known and uniform luminance properties.

The digital camera software automatically corrects uneven exposure on images caused by unbalanced lighting or the light-gather characteristics of a camera lens. Figure 11 displays results of the software corrections.



Exposures (including a white field) with distortion introduced during

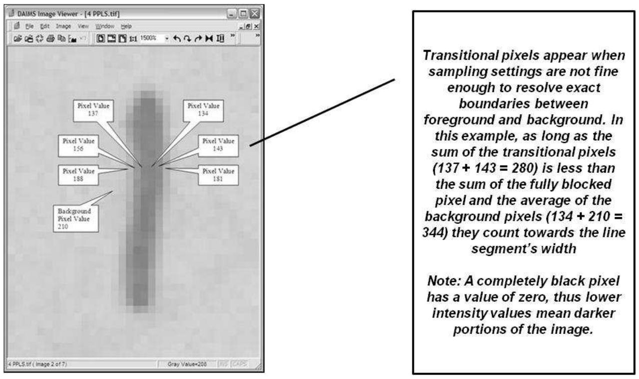
Corresponding images where distortion has been removed via exposure calibration

Figure 11. *DCamX uneven exposure correction*

Spatial Resolution

Definition	Spatial resolution is a calculation of the number of samples made from a source document and is often described as pixels per unit of measure. Ability to show fine detail accurately—to capture all the important information from a document. The more pixels used, the more detail can be seen and the higher the image's resolution.
Desired attribute range	Sufficient dots-per-inch (DPI) to provide 4-3 pixels per thinnest line segment (PPLS) in representative writing from the document. Ensures that sufficient detail to be able to read/transcribe information from any original document captured in a digital image.
Measure	The pixels-per-line-segment measure is based on determining the number of pixels captured across the thinnest line of a representative character from a document.

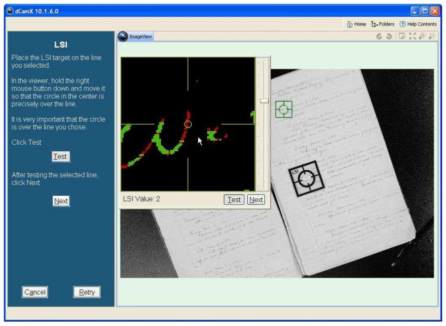
The formula for measuring spatial resolution using the pixels per line segment method is shown in Figure 12. The digital camera software determines if a selected line segment passes the requirement for spatial resolution. See Figure 13.



Transitional pixels appear when sampling settings are not fine enough to resolve exact boundaries between foreground and background. In this example, as long as the sum of the transitional pixels (137 + 143 = 280) is less than the sum of the fully blocked pixel and the average of the background pixels (134 + 210 = 344) they count towards the line segment's width

Note: A completely black pixel has a value of zero, thus lower intensity values mean darker portions of the image.

Figure 13. *Measuring spatial resolution using pixels per line segment method*



LSI

Place the LSI target on the line you selected.

In the viewer, hold the right mouse button down and move it so that the circle in the center is precisely over the line.

It is very important that the circle is over the line you chose.

Click Test

Test

After testing the selected line, click Next

Next

Cancel Retry

LSI Value: 2

Test Next

Figure 13. *DCamX software verify spatial resolution*

Color Space

Definition	A color space (sometimes referred to as image channels) represents the type of intensity information recorded for any given pixel. Bi-tonal and grayscale images contain a single channel recording the intensity of luminance. A typical RGB (color) image has three channels recording the intensity of the red, green, and blue channels respectively.
Desired attribute range	One channel for all documents where hue does not play a genealogical role in conveying information, three channels otherwise.
Measure	Number of color spaces or image channels used: one channel: black and white information only (often encoded in eight bits of grayscale, see Tonal Resolution) for all documents where hue (i.e. color) does not hold genealogical significance, three channels (in red, green, and blue) otherwise.

Color channels can be graphically represented using histograms. Figure 14 provides that illustration.

Figure 1. *FamilySearch Digital Pipeline*

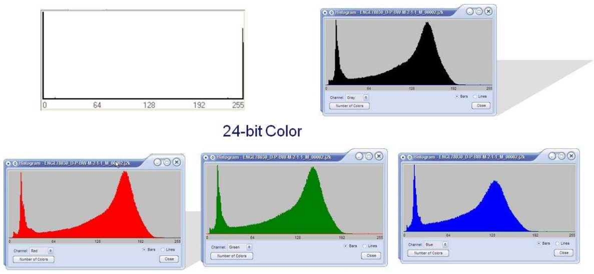
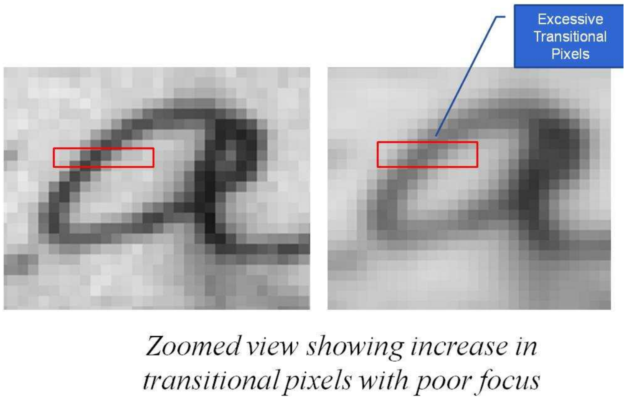


Figure 14. *Color channels represented through histograms.*

Focus

Definition	The quality of maximum sharpness of an image. Maximum sharpness is achieved by adjusting the relative positions of the capture medium (e.g., CCD array, lens) and the subject matter.
Desired attribute range	Proper alignment of the focal plane, as shown by no more than two transitional pixels within a high contrast boundary between foreground and background Assuming no out-of-focus characteristics exist in the original, e.g., a microfilm document captured with poor focus
Measure	Pixel count of transitional line. As an image moves to an out-of-focus state the number of transitional pixels expands

The difference in the number of transitional pixels between an in-focus and out-of-focus image can be clearly seen in Figure 15.



Excessive Transitional Pixels
Zoomed view showing increase in transitional pixels with poor focus

Figure 15. *Transitional pixel – in-focus vs. out-of-focus*

Blur

Definition	This is a function of motion during exposure of the document and is typically unidirectional in nature. The result of the blur may not be uniform to the entire document. Contributing factors of this are movement of the source document and scanner transport issues.
Desired attribute range	Transitional pixels in a single direction shall not be more than one pixel. The lines or characters in an image shall not appear as a double image
Measure	Pixel count of transitional line in a single direction. Appearance of multiple adjacent instances of the same line or character Aid to digital camera operator - Sensor in camera head

The causes and appearances of an out-of-focus and blur are distinct – the former created by improper focal length; the latter by movement. Figure 16 is provided to show the differences between the characteristics.

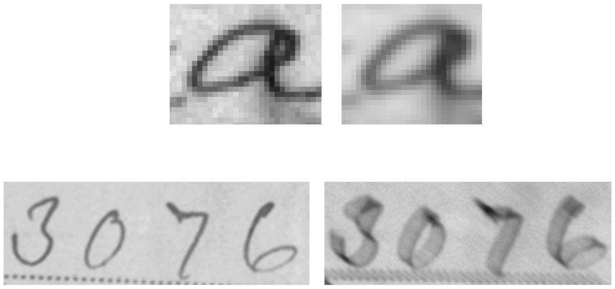


Figure 16. Top – out-of-focus; Bottom - blur

Image File Format

Definition	File formats are intended for different purposes and are often associated with a particular software program. Several are specifically designed for images. Commonly used file format names may in reality be image compression formats but also used to denote a file format.
Desired attribute range	TIFF 8-bit grayscale, JPG, JP2 lossless, PDF, etc.
Measure	Use of the desired file formats listed above is shown by the proper file extension plus conformance to the industry standards established for a particular format. Future Validation using JHOVE

Image Compression

Definition	A method used to encode data, including images, to minimize the size of the file. Losslessly encoded images are pixel-for-pixel equivalents to the original image, yet when stored take up less space. Lossy encoded images are smaller as they discard image data that ideally will not be noticed by the human eye. As the aggressiveness of the lossy compression increases, the data loss becomes more apparent - artifacts.
Desired attribute range	Currently being reviewed Compression artifacts do not visually degrade the readability of the document/image
Measure	Over-compression is the point at which compression artifacts visually degrade the readability of the document image.

Compression of a document image file increases artifacting or blocking around the lines which create the script. Too much compression can render the writing unreadable. Figure 17 is an example of image compression with the resulting artifacting.

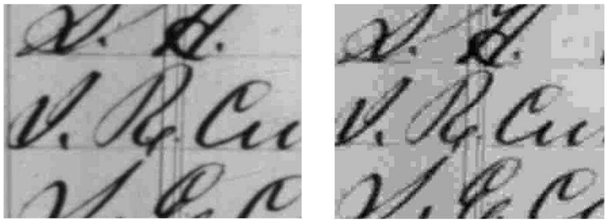


Figure 17. FamilySearch Digital Pipeline

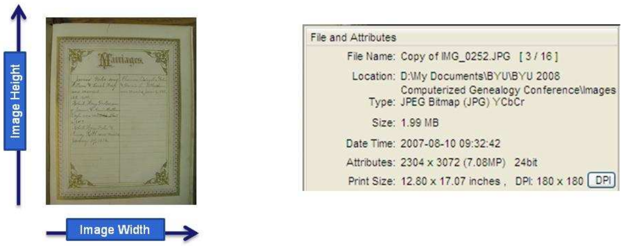
File Name

Definition	A file name is an identifier that is given to label a file. Filenames range from the simple to the complex, depending on the schema being used as well as external factors such as an operating system. Typically, filenames can be no longer than 255 characters long and can be comprised of both numbers and characters
Desired attribute range	Standard filenames based on business rules
Measure	Conformance to a particular set of business rules which allow for unique identification of image files

Image Dimension

Definition	The dimensions as measured in pixels along the X and Y coordinates of an image. These measurements define the pixel area of the image. These dimensions affect the file size of the image. No constraint. Determined by document size and spatial resolution
Desired attribute range	Image dimension is a result of original document size and spatial resolution settings which are determined by the original document characteristics Oversized documents that do not fit within an imaging device's capture area are scanned in sections, resulting in multiple digital images for one original document. Measure The number of pixels wide by the number of pixels high; measured in pixels along X and Y coordinates.
Measure	The number of pixels wide by the number of pixels high; measured in pixels along X and Y coordinates. Any image content beyond document edges showing complete capture (i.e., non-document space) is included in the measurement

Checking the file attributes in an image viewer will provide the X and Y pixel dimensions of an image along with other information about the image as shown in Figure 18.



File and Attributes

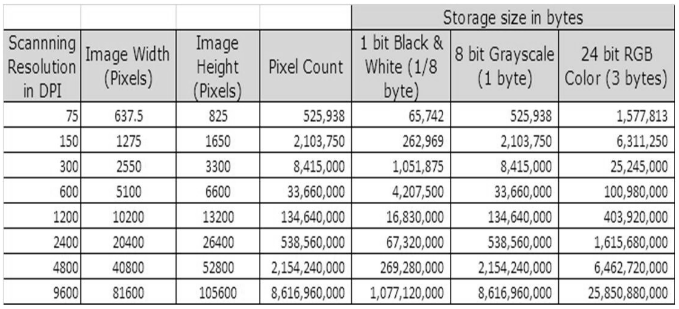
File Name: Copy of IMG_0252.JPG [3 / 16]
Location: D:\My Documents\BYU\BYU 2008 Computerized Genealogy Conference\Images
Type: JPEG Bitmap (JPG) YCbCr
Size: 1.99 MB
Date Time: 2007-08-10 09:32:42
Attributes: 2304 x 3072 (7.08MP) 24bit
Print Size: 12.80 x 17.07 inches , DPI: 180 x 180 DPI

Figure 18. *Image attributes as displayed in a image viewer.*

File Size

Definition	File size is simply defined as the size of an image file required to store the digital image and its associated metadata (if any). File size will vary depending on original image dimensions, spatial resolution, signal-to-noise ratio, compression methodology, etc.
Desired attribute range	No fixed number, though images are captured so that file size is no larger than straightforward legibility requires.
Measure	For uncompressed images – file size is computed by multiplying tonal resolution by the number of samples captured along an image's width, multiplied by the number of samples captured along an image's height, plus any space devoted to the storage of image metadata. Images employing compression use various schemes to represent redundancy with fewer data points than would be required to store data for each pixel. File size is usually stated in kilobytes (also KB), megabytes (MB), or gigabytes (GB)

The table in figure 19 shows the effect of original document/image size and spatial resolution on file size.



				Storage size in bytes		
Scannning Resolution in DPI	Image Width (Pixels)	Image Height (Pixels)	Pixel Count	1 bit Black & White (1/8 byte)	8 bit Grayscale (1 byte)	24 bit RGB Color (3 bytes)
75	637.5	825	525,938	65,742	525,938	1,577,813
150	1275	1650	2,103,750	262,969	2,103,750	6,311,250
300	2550	3300	8,415,000	1,051,875	8,415,000	25,245,000
600	5100	6600	33,660,000	4,207,500	33,660,000	100,980,000
1200	10200	13200	134,640,000	16,830,000	134,640,000	403,920,000
2400	20400	26400	538,560,000	67,320,000	538,560,000	1,615,680,000
4800	40800	52800	2,154,240,000	269,280,000	2,154,240,000	6,462,720,000
9600	81600	105600	8,616,960,000	1,077,120,000	8,616,960,000	25,850,880,000

Figure 19. *File-size growth with increased document size and resolution..*

Complete Capture/Cropping

Definition	Capture is complete when the full contents of one facing of a document is represented in the corresponding digital facsimile (image).
Desired attribute range	Only one leaf of a folio (i.e., a sheet of paper once folded resulting in two leaves or four pages) per image unless document is too large to capture at appropriate spatial resolution, then captured in pieces. Two leafs or pages of an open book may be captured in one image if sufficient spatial resolution is achieved. Otherwise all document edges are visible in the digital facsimile.
Measure	The image is captured so that all document edges (where the physical object ends and the "capture surface" begins) are visible, meaning some capture surface area is present as boundary in the digital image. Complete capture is measured by how adequate a boundary area exists along all document edges. No extraneous materials should appear in the image

Image Orientation

Definition	The reference axis of a digital image in relation to the document in represents. Sometimes referred to as right reading.
Desired attribute range	Primary reference axis of a digital document matches the reference axis of its corresponding digital image. In other words, a document is not captured "on its side."
Measure	The direction of intended viewing of the primary content of a document, e.g., lines of text or the orientation of imagery captured in an image

Figure 20 shows two image – one with an obvious orientation; the other needing to view the opposing side to determine orientation

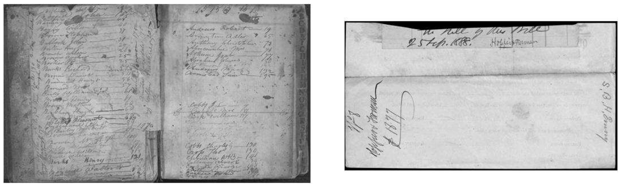


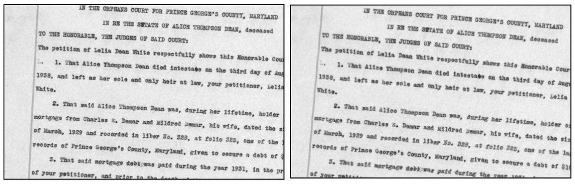
Figure 19. *File-size growth with increased document size and resolution..*

Image Skew

Definition	Image skew exists when the angle of the primary reference axis of an image is different from the reference axis of the resulting image. Typically skew is a slight phenomenon caused by misalignment of a document to the imaging system at capture.
Desired attribute range	Image skew no greater than 3.0 degrees.
Measure	Skew is measured in degrees – any angle between the reference axis of the digital image and

	the reference axis of the original document (skews of greater than +/- 45 degrees) are additionally classified as image orientation issues).

Image skew can create challenges for creating indexing templates and obtaining accurate optical character recognition (OCR). Figure 20 compares an acceptable skew on an image to an over-skewed image.



IN THE ORPHANS COURT FOR PRINCE GEORGE'S COUNTY, MARYLAND

IN RE THE SSTATE OF ALICE THOMPSON DEAN, deceased

TO THE HONORABLE, THE JUDGES OF SAID COURT:

The petition of Lelia Dean White respectfully shows this Honorable Cou

1. That Alice Thompson Dean died intestate on the third day of Aug 1938, and left as her sole and only heir at law, your petitioner, Lelia White.

2. That said Alice Thompson Dean was, during her lifetime, holder mortgage from Charles E. Demar and Mildred Demar, his wife, dated the s of March, 1929 and recorded in liber No. 328, at folio 325, one of the records of Prince George's County, Maryland, given to secure a debt of $

3. That said mortgage debt was paid during the year 1931, in the pr of your petitioner, and prior to the death

Further skew: Acceptable on left (3.0 degrees) unacceptable on right (5.0 degrees).

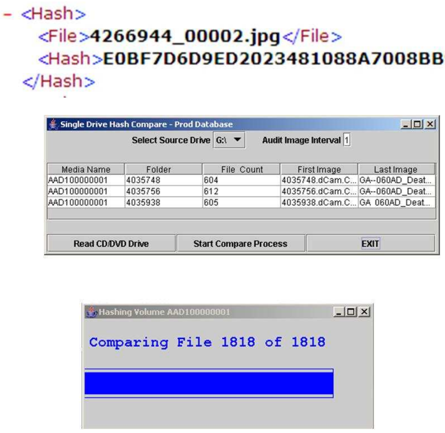
Figure 20. *Acceptable vs. unacceptable skew in an image*

Future Attributes

Fixity

Definition	A 'fingerprint' for each image file created to for a Digital Archive. As each file is scanned and a unique "fixity key" is created for the file. If anything is altered within the file, a subsequent fixity scan will generate a different key to flag a change.
Desired attribute range	Subsequent regeneration of fixity key (digital signature) of an image file matches the original key exactly. ;(image file [container] or bitmap)
Measure	As image files are copied to a new location, another fixity key is generated and compared to the original key to validate the authentic transfer of the image copy. (include bitmap fixity) MD5 hash algorithm used as Fixity for field imagery.

Fixity check is accomplished by recreating a MD5 hash signature and comparing the original to the stored signature as shown in Figure 21.



```
- <Hash>
    <File>4266944_00002.jpg</File>
    <Hash>E0BF7D6D9ED2023481088A7008BB91B2</Hash>
  </Hash>
```

Single Drive Hash Compare - Prod Database

Select Source Drive G:\ Audit Image Interval 1

Media Name	Folder	File Count	First Image	Last Image
AAD100000001	4035748	604	4035748.dCam.C...	GA--060AD_Deat...
AAD100000001	4035756	612	4035756.dCam.C...	GA--060AD_Deat...
AAD100000001	4035938	605	4035938.dCam.C...	GA 060AD_Deat...

Hashing Volume AAD100000001

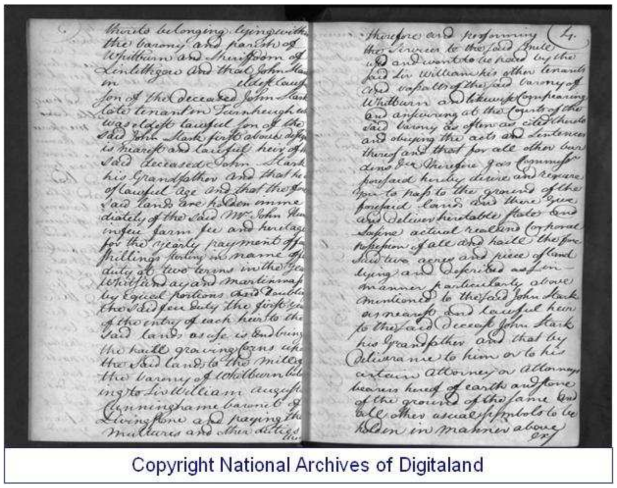
Comparing File 1818 of 1818

Figure21. *MD5 hash signature and hash comparison utility*

Attribution

Definition	The acknowledgment of the ownership or possession of the original object represented in a distribution image.
Desired attribute range	The attribution should be placed outside of the boundaries of object shown in the image. The user interface could contain the attribution in the form of a banner. The attribution will not be part of the image itself.
Measure	The attribution meets the copyright holder's requirements for acknowledgement of ownership or other requirement.

An attribution banner could be displayed next to the image as shown in Figure 22 or anywhere in the viewing area



thirls belonging lying with
the barony and parish of
Whitburn and Sheriffdom of
Linlithgow And that John Sta
in eldest lawf
son of the deceased John Stark
late tenant in Turnheugh w
was eldest lawful son of the
said John Stark first above desig
is nearest and lawful heir of th
said deceased John Stark
his Grandfather and that he
of lawful age and that the fo
said lands are holden imme
diately of the said Mr John Stir
in feu farm fee and heritag
for the yearly payment of fo
shillings sterling in name of fe
duty at two terms in the yea
Whitsunday and martinmas
by Equal portions and doublin
the said feu duty the first ye
of the entry of each heir to the
said lands as use is and bein
the haill grassing forms up
the said lands to the mill of
the barony of Whitburn bel
ing to Sir William augus
Cunninghame barnet of
Livingstone and paying the
multures and other duties

therefore and performing 4.
the Services to the said Mule
use and wont to be paid by the
said Sir William his other tenants
and vassalls of the said barony of
Whitburn and likewise Compearing
and answering at the Courts of the
said barony as often as cited thereto
and obeying the acts and sentences
thereof and that for all other bur
dens Yea therefore you Commissr
foresaid hereby desire and require
you to pass to the ground of the
foresaid lands and there give
and deliver heritable state and
Sasine actual real and Corporal
possession of all and haill the fore
said two acres and piece of land
lying and described as in
manner particularly above
mentioned to the said John Stark
as nearest and lawful heir
to the said deceast John Stark
his Grandfather and that by
deliverance to him or to his
certain attorney or attorneys
bearers hereof of earth and stone
of the ground of the same and
all other usual symbols to be
holden in manner above

Copyright National Archives of Digitaland

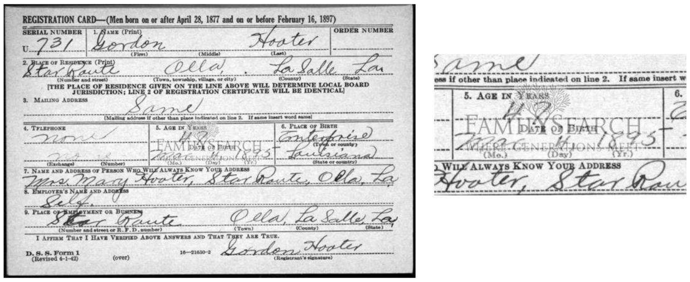
Figure 22. *Sample of an image attribution*

Visible Watermarking

Definition	An image, both visible and translucent, which is superimposed or overlaid on the primary image. The watermark image may consist of the logo or seal of the organization which holds the rights to the primary image. The primary image may be viewed, but is marked clearly as the property of the organization owning the copyright.
Desired attribute range	The visible watermark is overlaid on the primary image in a way which makes it difficult to remove, in order to achieve the goal of indicating ownership. Watermark structure, opacity, size and position must all be considered to achieve this goal Watermarking may be best accomplished in real-time as an image is delivered for viewing, downloading or printing rather than done to each image stored on a server. This approach will reduce storage requirements.

Measure	The content underneath the watermark overlaid on the primary image remains readable.

Figuere 23 diplays an image with the FamilySearch logo as a visible watermark by way of example


REGISTRATION CARD—(Men born on or after April 28, 1877 and on or before February 16, 1897)

SERIAL NUMBER U 731	1. NAME (Print) Gordon (First) (Middle) Hooter (Last)	ORDER NUMBER

2. PLACE OF RESIDENCE (Print)
Star Route (Number and street) Olla, (Town, township, village, or city) La Salle (County) La (State)
[THE PLACE OF RESIDENCE GIVEN ON THE LINE ABOVE WILL DETERMINE LOCAL BOARD JURISDICTION; LINE 2 OF REGISTRATION CERTIFICATE WILL BE IDENTICAL]

3. MAILING ADDRESS
Same
(Mailing address if other than place indicated on line 2. If same insert word same)

4. TELEPHONE none (Exchange) (Number)	5. AGE IN YEARS 47 DATE OF BIRTH Mar (Mo.) 4 (Day) 1895 (Yr.)	6. PLACE OF BIRTH Enterprise (Town or county) Louisiana (State or country)

7. NAME AND ADDRESS OF PERSON WHO WILL ALWAYS KNOW YOUR ADDRESS
Mrs. Mary Hooter, Star Route, Olla, La

8. EMPLOYER'S NAME AND ADDRESS
Self.

9. PLACE OF EMPLOYMENT OR BUSINESS
Star Route (Number and street or R. F. D. number) Olla, (Town) La Salle, (County) La (State)

I AFFIRM THAT I HAVE VERIFIED ABOVE ANSWERS AND THAT THEY ARE TRUE.

D. S. S. Form 1
(Revised 4-1-42) (over) 16—21630-2 Gordon Hooter
(Registrant's signature)

Same
ess if other than place indicated on line 2. If same insert w
5. AGE IN YEARS 47 6.
DATE OF BIRTH
Mar 4 1895
(Mo.) (Day) (Yr.)
WILL ALWAYS KNOW YOUR ADDRESS
Hooter, Star Rou

Figure 23. Visible watermark

.

Author Biography

Richard Laxman championed digital technology and imaging processes development for FamilySearch. He developed the first digital camera system for the organization and today manages its digital data processing center. Mr. Laxman received a master’s degree in Business Information Systems (2005) from Utah State University. He serves as committee chair of C24, Electronic Imaging in the AIIM standards program, project editor for an ISO metadata project and serves on the AIIM Standards Board.