



IMT 4891 - Digital Workflow Fundamentals
August 26, 2009

#2: Collecting and describing data

Kjell Are Refsvik

Welcome!

Questions so far?

Last lecture...

We apologize for the lack of sound on the published lecture videofile from last week.



This week we will be trying our best to capture sound using the Azden WMS Pro lavalier microphone system.

E-mail

Please send us an email this week to verify that you can be reached by e-mail. Please include the following information in your e-mail to us:

Name

Class

Exchange student or not

Full-time/Part-time student

Campus/Distant student

Send to kjell.refsvik@hig.no

Laboratory hours today 13:30-15:15 in A211. Feel free to join us to start working on your weekly exercises.

Frontier room/web page will be updated thursday (tomorrow) with exercises and links to more suggested reading material.

Today

	Principles	Formats	Standards	Tools/ devices
Capture / produce data				
Describe data				

Today

	Principles	Formats	Standards	Tools/ devices
Capture / produce data				
Describe data				

Digital photos

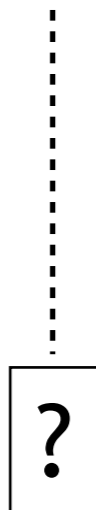


Capture data

Introduction

Capture/creation of data

By sensors or computer programs



Introduction

Capture/creation of data

By sensors or computer programs



Keyboard
Mouse

Introduction

Capture/creation of data

By sensors or computer programs



Keyboard
Mouse

Sound
Image
Gyro
GPS

Principles (and general facts)

Sensors today can capture large amounts of data:

15 million pixel images or
48000 sound cycles per second

Data can then hold a lot of quality

Think carefully before you let your device process and compress the recording for you. RAW>JPEG loses 60-80% of data

Capture and store raw data if possible (Aperture Demo)

Structure and Backup is highly important
("It does not exist before it is stored at least 3 places")

Metadata is becoming more important

Standards

Few open and shared capture standards.
Often vendor-specific.

Photo: Wide array of raw formats
Sound (aiff, wav), GPS (nmea, gpx)
and other devices have more standard
output formats.

Tools/devices

Video, sound, photo, gyro, geo, time

Some special purpose (gps), some all-round devices (phone)

Some of these devices has also been called context-aware devices due to the fact that they can be aware of their context and this information can be used to help the user

Date, time, location, direction, altitude,

http://en.wikipedia.org/wiki/Context_awareness

Add-ons like good microphones is then a necessity
(example: equipment used in this lecture)



image
 date
 time
 exposure data incl.
 gyro



image
 date
 time
 exposure data incl.
 gyro
 [sound + video]



longitude
 latitude
 altitude
 elipsoid
 date
 time
 direction
 speed

sound
 image
 video
 exposure data

longitude
 latitude
 altitude
 elipsoid
 date
 time
 direction
 speed



sound
 date
 time



Describe data

Introduction

Why describing data is important
and how you can do it.

<http://en.wikipedia.org/wiki/Metadata>

Principles

Types of metadata:

Administrative

Metadata used in managing and administering information resources.

Descriptive

Used to describe or identify information resources

Preservation

Used in the preservation of data for the future

Technical

Metadata relating to how the object was born or needs to be played back

Use

Metadata that relates to how the object can be used.

From: "Introduction to Metadata - Pathways to Digital Information", Getty Information Institute

Principles

Automate collection of machine-readable data

Example: lon, lat, elipsoid, alt, direction

From that: country, region, state, county

Collect metadata as early as possible in the process, preferably in the device

4 places to store metadata:

Included in data (visible or not) (Note: will destroy original data)

As part of the filename (useful)

In a sidecar file (easy to lose)

Embedded in a header (exif and iptc) (useful)

Formats

Photos (tiff, jpeg): EXIF and IPTC IIM

Music/Sound (.mp3): ID3

Other file formats: A range of vendor-specific ones

Standards

EXIF and IPTC (Photos)

Example: XMP

ISO 8601 (date/time naming)

ISO 3166 (country codes)

LOM (Learning Object Metadata)

Dublin Core (everything)

Future and the semantic web

http://en.wikipedia.org/wiki/Metadata_standards

Tools

The Operating system

Special programs

Critical to choose tools for editing data that does not destroy metadata

Example: Flickr and their removal of metadata

ExifTool

Exiftool

<http://www.sno.phy.queensu.ca/~phil/exiftool/>

Exercises

1. Find a few of your own digital photos, alternatively download some from the Internet. Look inside the files using the “exiftool” program. What metadata do you find and what groups do they belong to?
2. Find out how to put metadata into your pictures and insert your name into a appropriate field. Show the content of the metadata before and after.
3. Extract metadata from a range of file-types on your harddrive. What types of files have the most metadata?
4. Do you find any significant difference in terms of metadata when looking inside files created by computer software as opposed to files created by different types of devices (cameras, sound recorders, mobile phones, gps-systems)?

Solving atleast 2 of these exercises before next weeks lecture will be mandatory. You are free to select what exercises to solve.

We would like you to upload your answers as a single utf-8 encoded plain text file (no Microsoft Word-files, please) into Fronter>imt4891...>Task>Week 35.

Please name the file imt4891_week35_<your_surname>.txt