scenarios
Personal
photo
involved
data
CDVP
memory
focus
organisation
contextual
volume
develop
wearing
recorders
sensecam
understanding
creation
possible
Media
human
discuss
video
usage
organise
novel
along
challenges
findings
domain
Managing
devices
Content
technologies
Context
Archives
collections
specifically
DCU
organising
employed
effective
retrieval
cameras
techniques
retrieving
captured
memories
users
Analysis
today's
utilising
automatically
talk
pose
long-term
many
digital

Context and Content Analysis for Managing Personal Media Archives

Dr. Cathal Gurrin
Centre for Digital Video Processing,
Dublin City University, Ireland.

Centre for Digital Video Processing

- 6 Faculty, 17 PostDoc, 4 RA, 20 PhD
- CDVP Summary Research Interests:
  - Digital multimedia analysis and organisation
    - Video : TV, News, Educational, CCTV
  - Personal Media Management
    - Human Digital Memories
    - Personal Photo Management
  - Sensor Networks & SensorWeb Technologies
  - Text IR / Search Engines

Human Digital Memories

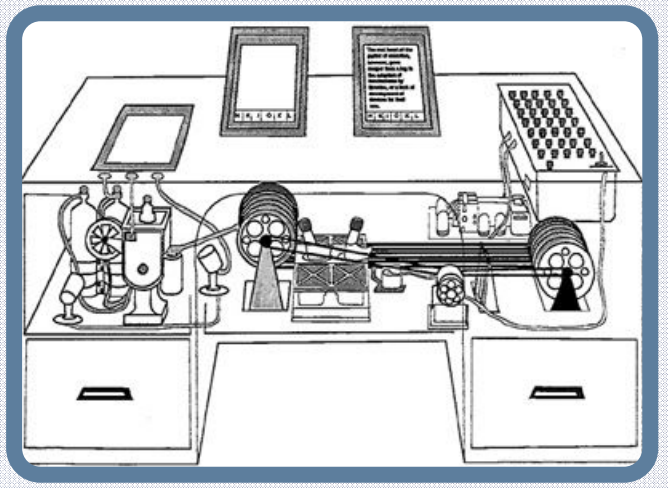
- A Human Digital Memory (HDM) is a surrogate of your own memory, though in digital form.
- Although it may sound like Science Fiction, it is being done now, though not quite as intrusively as you would think
- So
  - –
  - –
- So
  va ns;
- Sometimes its because we can, and we are still researching how we can use it

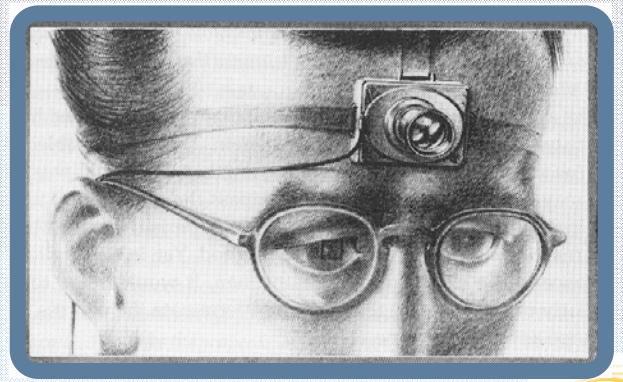
"Some day we will be able to record everything we see and hear."

Bill Gates, in "The Road Ahead", 1995

What comprises a HDM?

- Humans have been finding ways to substitute memory for a long time
  - Druids, books, computers…
  - MEMEX from 'As We May Think'
- For HDMs we can…
  - Log our environment;
    - Traffic, pollution, **people near us**, etc…
  - Log our **location**
  - Log what we see and do
    - **Visually**, aurally,
  - Recording biometrics;
    - Log our health, fitness, activity, etc…
  - Log our actions
    - Communications
    - Data interactions





People have worked on this…

In this talk…

- I will focus on the visual aspects of a Human Digital Memory
  - Capturing what we see, and organising this data
- Agenda
  - CDVP Research on Personal Media Archives
  - Constructing a Human Digital Memory
  - Organising a Visual Human Digital Memory
  - Challenges for Human Digital Memories
  - Human Digital Memories, My thoughts

Prior CDVP Experience in Managing Personal Media Archives

Personal Photo Collections

Personal Video Collections

Managing Personal Photos

- Most photo search relies on date/time and manual annotation/organisation
- Recently location (one type of context) has been used to support organisation (Flickr)
- Some key points for automatic organisation
  - Photo capture is bursty in nature
    - Helps EVENT segmentation
  - Context of capture can drastically reduce the search space
    - Date/Time/Location/People/biometric data
  - Content Analysis can also help (feature detectors)
    - Baby/dog/face/building

http://mediassist.cdvp.dcu.ie/demo/frameset.php

Google

MediAssist

Tools for organising, browsing and retrieving from a personal electronic picture collection

TOTAL #PHOTOS: 2128

Enter Text Query Here

LOCATION

Select the place where the photos were taken. RESET

COUNTRY: CANADA | STATE/COUNTY: Quebec | CITY/TOWN: Any

TIME RANGE

Set start and end time for your search.

96 97 98 99 00 01 02 03 04 05 06

RESET

SELECTION: 7 EVENTS 100 PHOTOS SHOW

ADVANCED

PERSONS 0 FEW CROWD

NAMES

BUILDING YES NO ANY

TIME FILTER

Month: Jan Mar May Jul Sep Nov

Day: 1 5 10 15 20 25 31

Day of Week: M T W T F S S

Hour: 0 6 12 18 24

LIGHT STATUS ANY

INDOOR / OUTDOOR IN OUT ANY

WEATHER ANY

RESET

SEARCH SUMMARY

Browse the events below and view more photos in each event. Click on a photo to see full-size. [INDIVIDUAL PHOTOS]

Information about photos relevant to the query:

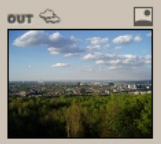
Photos: 100 INDOOR 30 WEATHER: 4 96 0 0 LIGHT STATUS: 0 98 0 2

Events: 7 OUTDOOR 70

EVENT: Montreal, Quebec, CANADA. (16-5-2004) OUT SHARED 37 Photos VIEW ALL

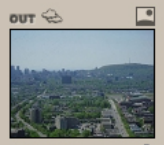
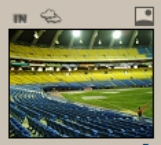






EVENT: Hochelaga County, Quebec, CANADA. (17-5-2004) IN/OUT SHARED 28 Photos VIEW ALL

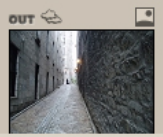


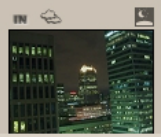
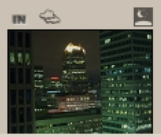




EVENT: Montreal, Quebec, CANADA. (18-5-2004) IN SHARED 1 Photo VIEW ALL

Content

EVENT: Montreal, Quebec, CANADA. (18-5-2004) IN/OUT SHARED 9 Photos VIEW ALL

My Friends' Faces

Search and browse people's faces in your photos and share with friends in a fun and easy way

Logged in as: **Hyowon**
834 Photos in **106** Events

UPLOAD • MY ACCOUNT • SIGN OUT • HELP

On the table

Dublin, Ireland
5 June 2005

Owner: **Hyowon**
834 Photos in **106** Events

Search for people, location, or event in annotation

GO

ADVANCED

PEOPLE LIST

86 people annotated. To search by person, select & click GO button (Hold Ctrl for multiple selection).

NAME	PHOTO	EVENT
Fabrice	23	3
Paul	234	43
Georgina	76	20
Cathal B	1,430	95
Mary	15	3
Hyowon	9	2
Bart	53	6
Neil B	1,316	103
Alan B	635	25
Sandra	128	14

MORE

GO

RECENT EVENTS

INDOOR | CITYSCAPE | PEOPLE IN THIS PHOTO

Paul
Georgina
Cathal B
Mary

On Cathal's birthday party on Friday afternoon in The Arc, Liffey Valley. People were happy and day was bright, restaurant was pleasant. We liked the day very much! The meal was also very nice.

The Arc: http://www.thearc.ie
Dublin tourist guide: http://www.dublin-tour.ie

COMMENTS

Joyce
Nice photo! I hope you guys enjoyed the event. (7 Jun 2005)

Christopher
Very nice shot with nice lighting! What restaurant is this? Let me know if you visit this place again.. (22 Jul 2005)

Cathal B
Ah, thanks for uploading this photo! (23 Jul 2005)

↳ **Hyowon**
No problem, I had it for a long time but hadn't uploaded. (24 Jul 2005)

Aoife
I like the colour - well done! (1 Aug 2005)

Dave
Ireland must be a good place for celebrating a

NAVIGATE EVENT

Cathal's birthday Party (27 Photos)

PEOPLE LINK

See these people in other photos in:
This Event | **All My Photos** | **All My Buddies**

Paul (234 Photos)

Georgina (76 Photos)

Cathal B (1,430 Photos)

MORE

EVENT LINK

See similar Events in:
My Photos | **All My Buddies** | **All Users**

MORE

Simon
EVENT **Dublin,** Aug 06 (24 Photos)
EVENT **Dublin,** Feb 06 (72 Photos)

Joe
EVENT **Dublin,** Apr 06 (59 Photos)

What we have learned?

- Event segmentation is important to organise the content
- Context is key for accessing content based on a person's memory of the event
  - A person's query will be incomplete/incorrect
  - When an event occurred is not easily recalled
    - Where the event took place is much easier to recall!
- Content analysis tools will inevitably fail some/much of the time, so need to be able to manually fix.
  - When they work, they can be very useful if the correct suite of tools are chosen
  - What are the correct tools for HDMs?

Personal VIDEO Archives

Integration of Content Analysis, specifically visual feature extraction tools

Video Search

- Video data organised into video shots and scenes for easier organisation
- Extensive use of Metadata for searching
- Also, research effort on visual content search
  - Automatically extract features from each keyframe/shot/…
    - Image/Video low-level features
    - Derived Attributes
      - People, Location, Objects, Events
      - Likely trained using an SVM over the low-level features
  - CMU estimate that about 5,000 concepts detected with minimum accuracy of 10% is needed to provide web search quality results for domain limited video

QUERY PANEL

Enter term(s) and click SEARCH button. Added shots will be used together for searching.
[**clear text**] [**clear all**]

rocket launch | SEARCH

- [ ] GLOBAL COLOUR
- [ ] LOCAL COLOUR
- [x] EDGE
- [x] MOTION
- [ ] TEXTURE
- [ ] FACE FILTER

had failed repeatedly and publicly and there're tests...

REMOVE ✕

powerful rockets that in nineteen sixty one launched during...

REMOVE ✕

are shuttle launch can be built in wrestling but...

REMOVE ✕

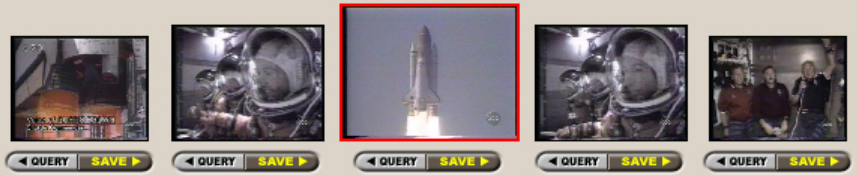
SEARCH RESULT

Search found **250** matching shots. Following is the **ranked list** of the search result. Click an image to play the segment, click SAVE button to add to your answers, or click QUERY button to improve subsequent query.

RESULT PAGE: 1 2 3 4 5

1: ABC News (30 Oct 1998)

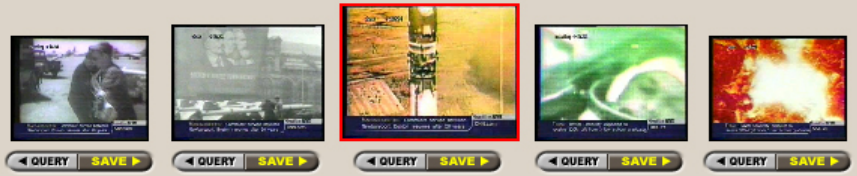
> MORE MATCHES IN THIS BROADCAST



that manages about we've got big he'd be in on board discovery yesterday to **launch** would have looked like this when is in the middle of this picture are shuttle **launch** can be built in wrestling but it was far gentler than lives at least one we've that the astronauts back in space again that was quite different than a ride i got before of course back a long time ago

2: CNN News (26 Oct 1998)

> MORE MATCHES IN THIS BROADCAST



those were the coldest years of the cold war the soviet military arsenal included powerful rockets that in nineteen sixty one launched during the government the first person in space this all u. s. **launch** rockets

3: CNN News (26 Oct 1998)

> MORE MATCHES IN THIS BROADCAST

the government the first person in space this all u. **nch** rockets had failed repeatedly and publicly and there're tests pressure was building dev american investments these are the past there have been

4: CNN News (26 Oct 1998)

> MORE MATCHES IN THIS BROADCAST

TASK 1

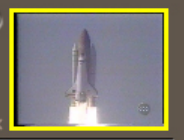
SEE DESCRIPTION

FINISH THIS TASK

SAVED SHOTS

You saved **2** shots, shown below - these are your **answers** to the current task. Add more, delete, or use them for subsequent query.

QUERY



DELETE ✕

QUERY

DELETE ✕

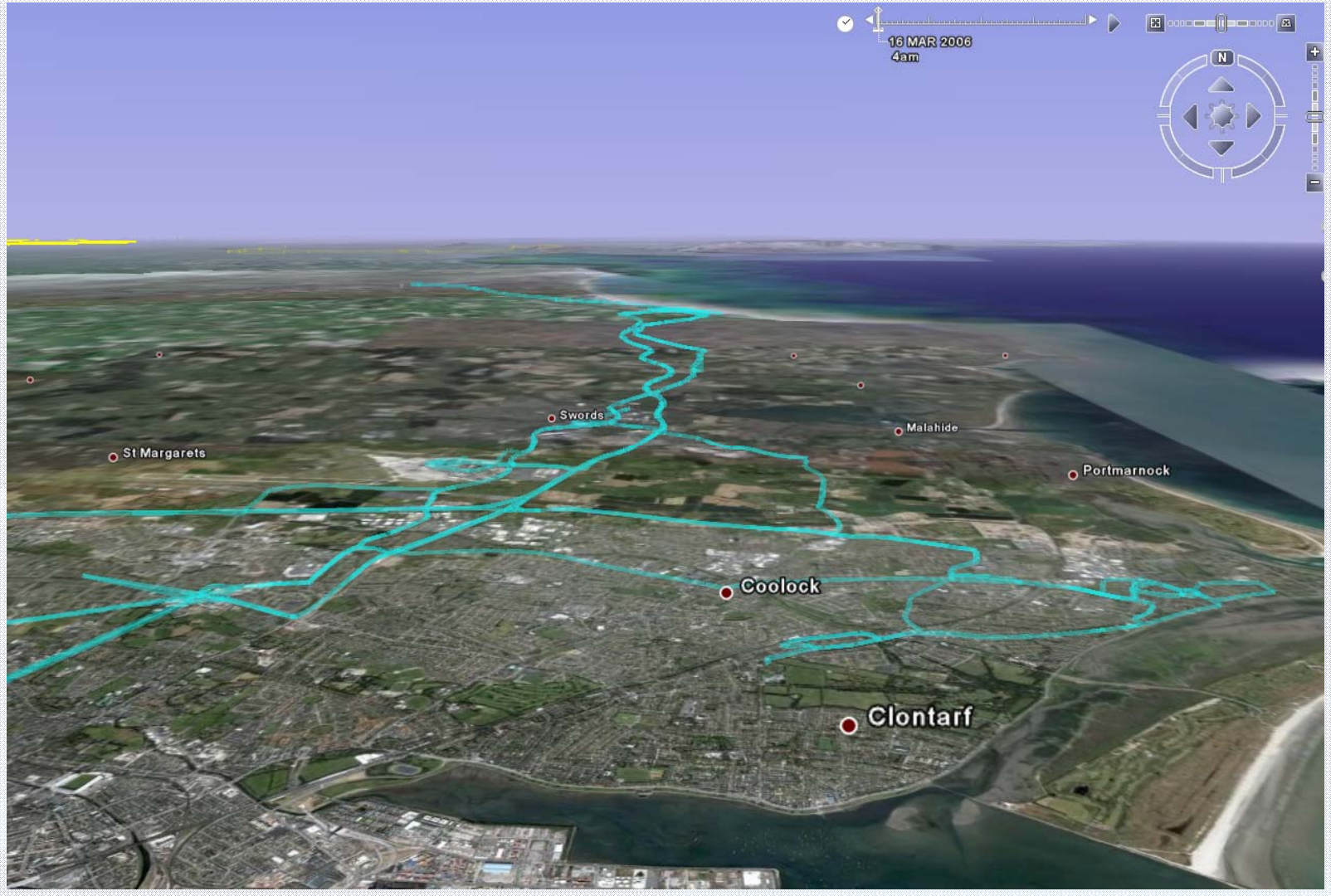
What have we learned?

- Shot Boundary Detection and scene detection is important to organise the video content into meaningful units
- Content analysis tools can help
  - Either a lot of tools (,000s) or carefully chosen domain dependent tools (e.g. for sports summarisation)
  - Are the same HDM tools suitable for HDMs?
- Textual surrogates (not visual content) are still are key access and organisation methodologies
- Keyframe size and layout help to guide a user to the most important parts of the content

Constructing a Human Digital Memory

Movement
People we meet
Our Interactions
What one sees

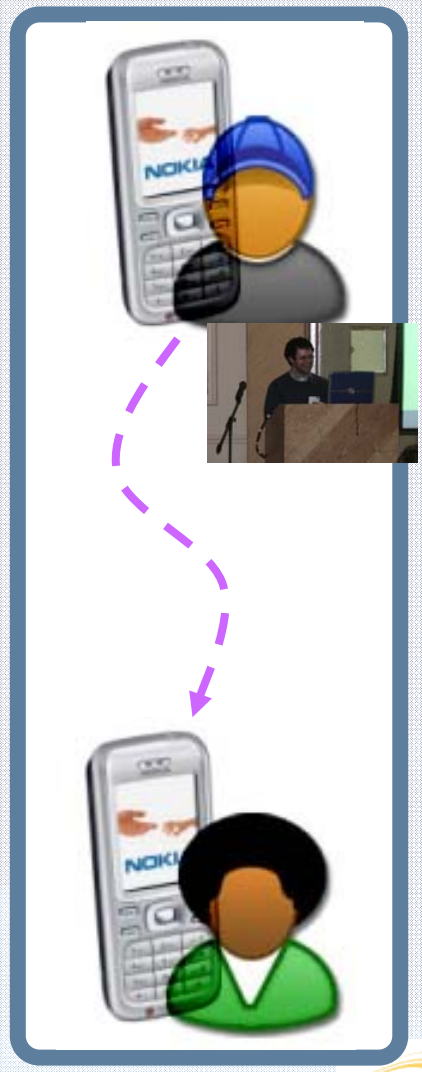
Logging Movement



Cathal Gurrin (2008)

Logging People Interactions

- Social Network Generation
  - Based on real-world interactions using Bluetooth on mobile devices
- This allows us to log who is near to us at any one time
- Can help to organise our HDM by automatically annotating events with people likely to be there



Logging our Actions

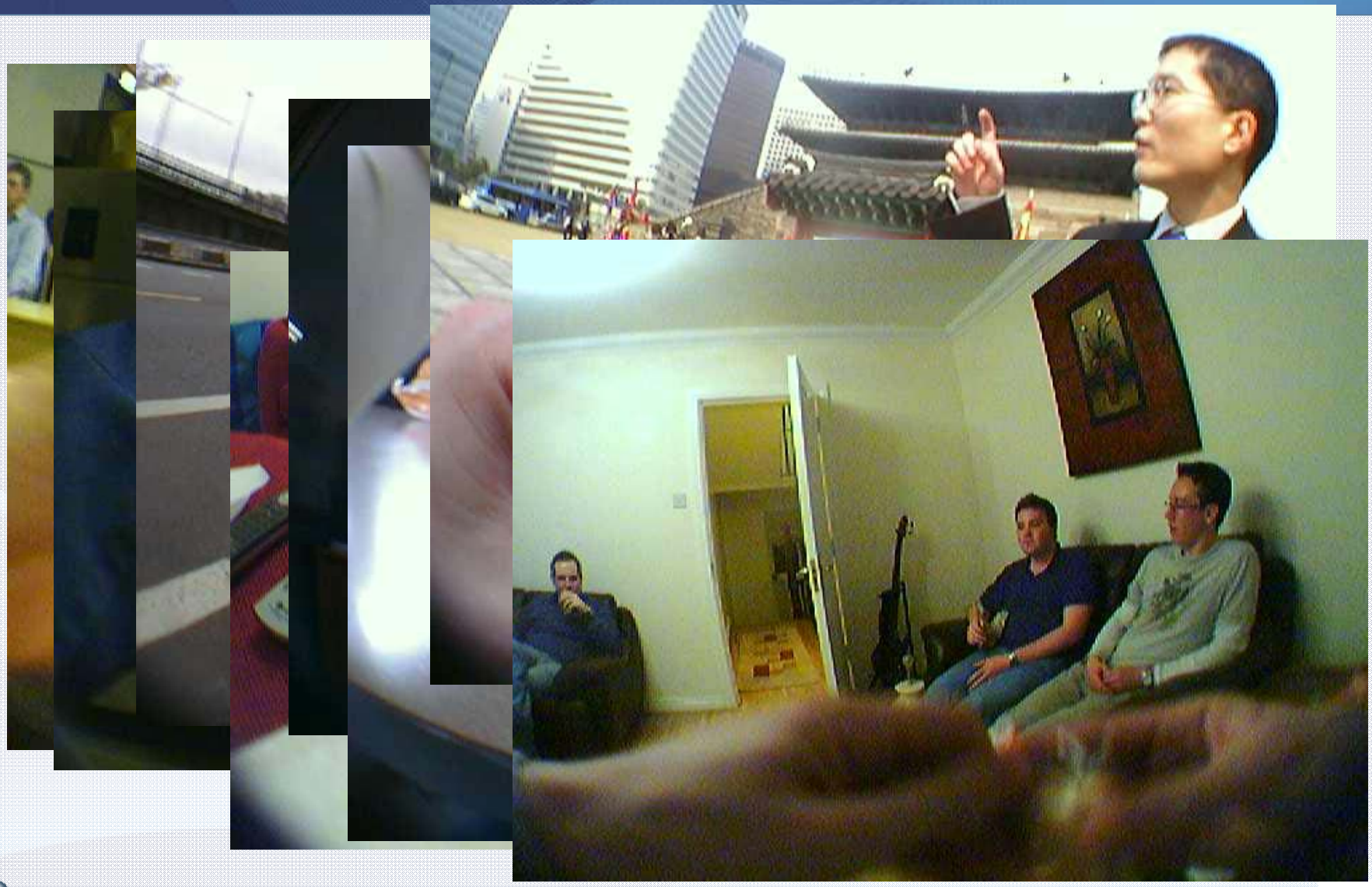
- Biosensors can log our actions and even log aspects of our emotion
  - Require wearing sensors and not likely to be comfortable to wear
    - E.g. heart rate, excitement, respiration,…
  - Sensecam gathers environmental temperature
- Logging a person's communications and data actions
  - Phone calls
  - Text messages
  - Emails
  - Web pages visited
  - Content of documents worked on

SenseCam: Logging what we see

- SenseCam is a Microsoft Research Prototype
- Multi-sensor device:
  - colour camera (vga with fisheye lens)
  - Sensors
    - 3 accelerometers
    - light meter
    - passive infrared sensor

- 1GB flash memory storage of over a week
- Smart image capture ~3 images/min
- CDVP have seven Sensecams
  - We see them as general HDM creation tools
  - Other research is more domain specific (e.g. health)

SenseCam Photos



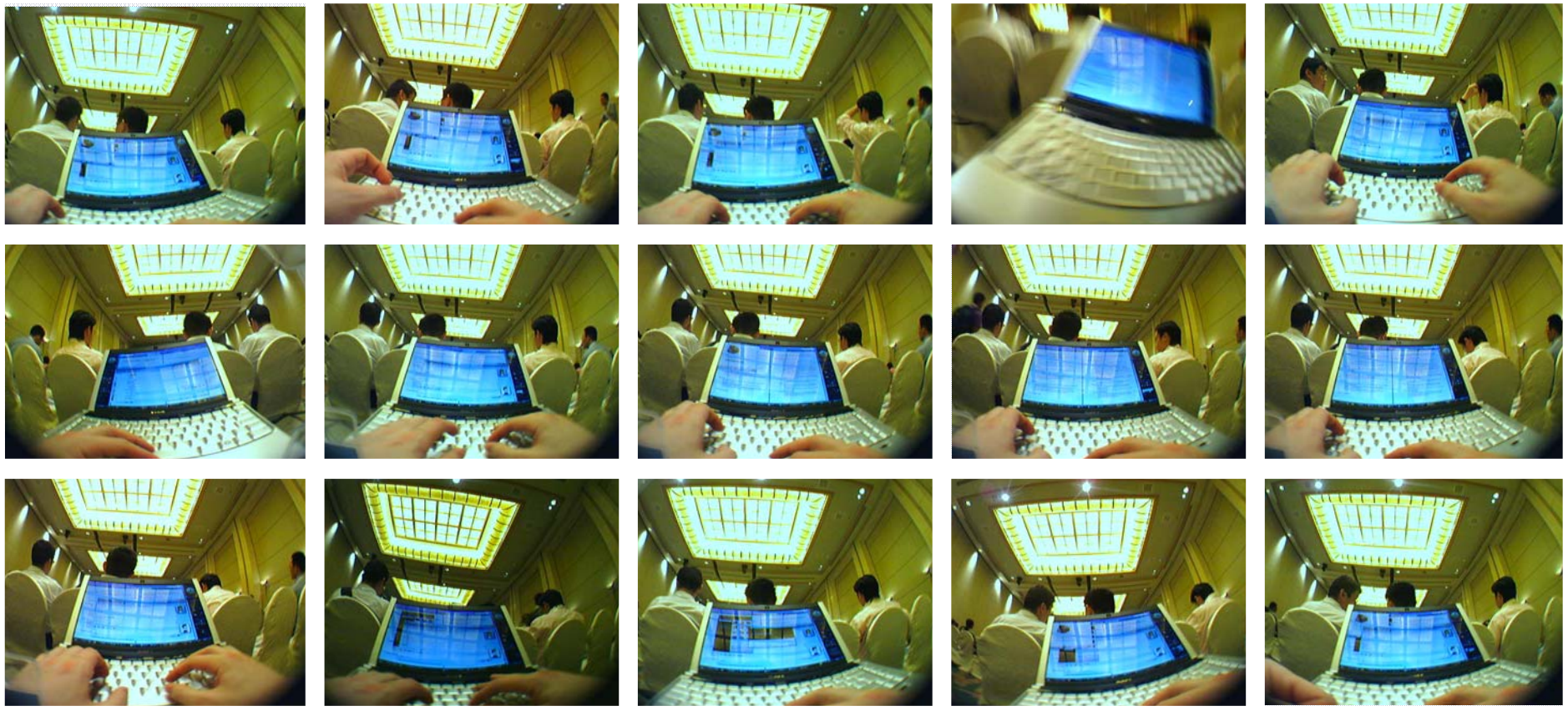
So, what does a day look like?

Analysing the Visual HDM

- Wearing a Sensecam will produce over a million photos in one year
  - Assuming about 16 hours per day
- These photos:
  - Have huge levels of redundancy
  - Vary in quality from unusable to photo-album
  - Captured scenes differ from conventional photos
    - Non-standard photo capture environments, like inside a car, or in my office.
  - Don't typically have a salient object
    - Sometimes people are captured 'headless'
  - Capture the hands of the wearer

Analysis: Redundancy

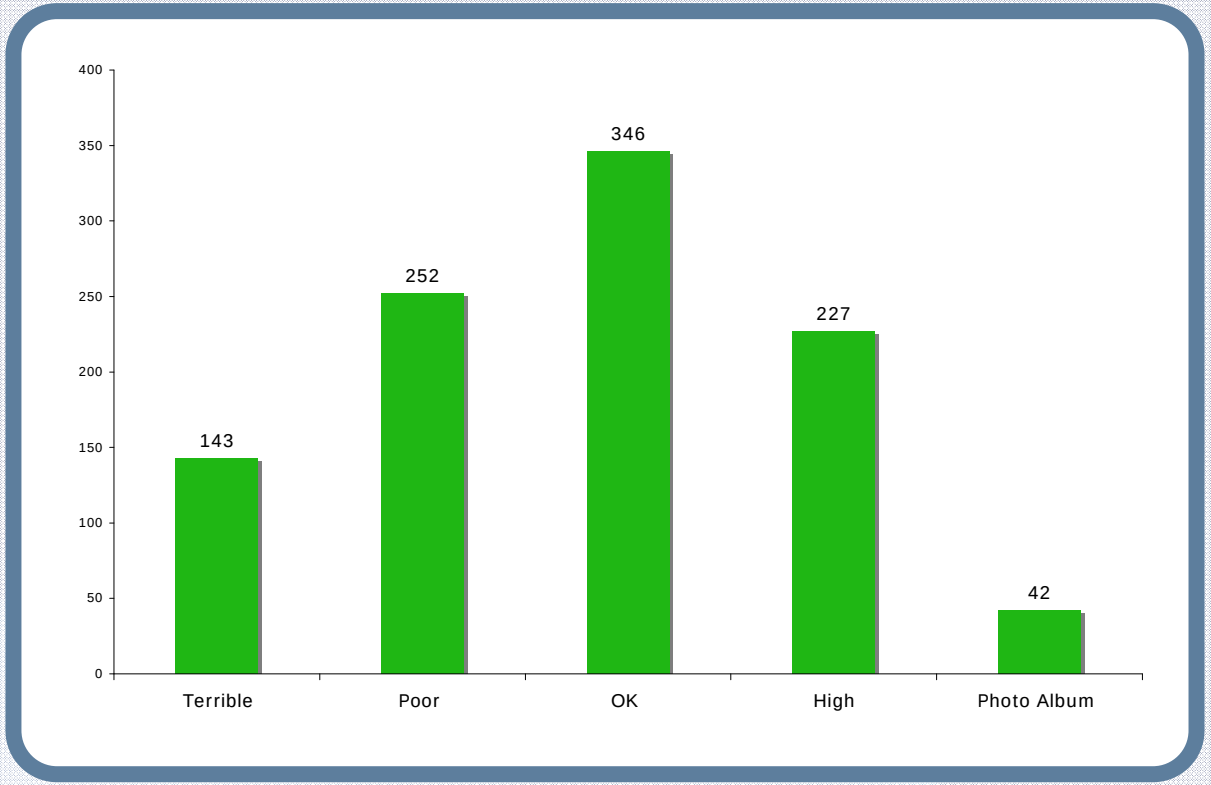
- Huge redundancy in visual HDM data.



Analysing the Visual HDM

- Wearing a Sensecam will produce over a million photos in one year
  - Assuming about 16 hours per day
- These photos:
  - Have huge levels of redundancy
  - Vary in quality from unusable to photo-album
  - Captured scenes differ from conventional photos
    - Non-standard photo capture environments, like inside a car, or in my office.
  - Don't typically have a salient object
    - Sometimes people are captured 'headless'
  - Capture the hands of the wearer

Analysis: Quality



- Randomly selected 1000 images from 1 million
- Manually annotated for quality

Analysing the Visual HDM

- Wearing a Sensecam will produce over a million photos in one year
  - Assuming about 16 hours per day
- These photos:
  - Have huge levels of redundancy
  - Vary in quality from unusable to photo-album
  - Captured scenes differ from conventional photos
    - Non-standard photo capture environments, like inside a car, or in my office.
  - Don't typically have a salient object
    - Sometimes people are captured 'headless'
  - Capture the hands of the wearer

Analysis: Comparison to Normal Photos

	HDM (1000 from 1M)	Personal Photos (10,523)
People	29.9	*30.5*
Buildings	3.5	*35.0*
Indoor	73.4	*15.2*
Outdoor	5.1	*84.8*
Cityscape	2.1	*22*
Landscape	1.1	*23.8*
Computer Screen/TV	26.7	0.0
Conversation Scenario	13.3	< 1.0
In a Vehicle	16.2	0.0
Work Scenario	24.6	0.0
Back view of People	6.8	2.0

Analysing the Visual HDM

- Wearing a Sensecam will produce over a million photos in one year
  - Assuming about 16 hours per day
- These photos:
  - Have huge levels of redundancy
  - Vary in quality from unusable to photo-album
  - Captured scenes differ from conventional photos
    - Non-standard photo capture environments, like inside a car, or in my office.
  - Don't typically have a salient object
    - Sometimes people are captured 'headless'
  - Capture the hands of the wearer

Organising Visual HDMs

Event Detection

Integrating Context

Content Analysis

Suggested Organising Process

1. Event detection and segmentation
   - To divide a HDM into a set of discrete events
   - Fixed events or dynamic events?
2. Integration of content analysis
   - To add automatic semantic annotations to the content
3. Integration of context information
   - Additional semantic enrichment
4. Indexing and presentation
5. Possible human annotation
   - Amend / add to the automatic semantic annotation and enrichment

1. Event Detection & Segmentation

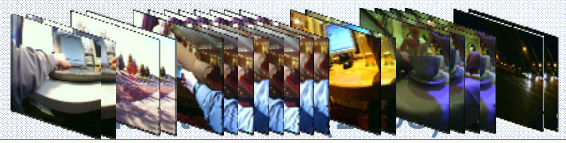
- Segment 4,500 photos per day into a set of events (F 0.62 for visual features & sensors, 0.55 for visual alone)
  - We employ visual features and output of on-device sensors

Event Segmentation

Multiple Events

Finishing work in the lab | At the bus stop | Chatting at Skylon Hotel lobby | Moving to a room | Tea time | On the way back home

Summarisation



2. Integration of Content Analysis

- We know that analysis content to identify visual features can work for both photo and video retrieval
  - May not accurate when examined independently, but useful within an overall content organisation scenario
- For HDMs, content analysis can:
  - Provide useful organisation functionality for a visual HDM
    - Quality detection of photos
    - Near duplicate detection
  - Support indexing and retrieval
  - Support linkage among events based on concept co-occurrence
- However a new suite of content analysis tools will need to be developed.

Concept Detection Tools

- Like conventional photos, visual concept detectors can be used for semantic enrichment
  - However, the concepts will be different

Concept	Frequency	Concept	Frequency
Vehicles (external view)	0.93%	Vegetation	*1.5%*
Road	1.95%	Screen	***20.51%***
Inside of vehicle	**4.33%**	Paper/Book (reading)	*1.79%*
Indoors	37.39%	Meeting	*4.63%*
Door	1.12%	Office	***15.06%***
Outdoors	6.37%	Food (eating)	*4.15%*
Building	3.58%	Hands	20.6%
Tree	1.85%	Holding a cup/glass	0.81%
Grass	0.86%	Faces	5.84%
Sky	2.93%	People	12.83%

Concept Detector Evaluation

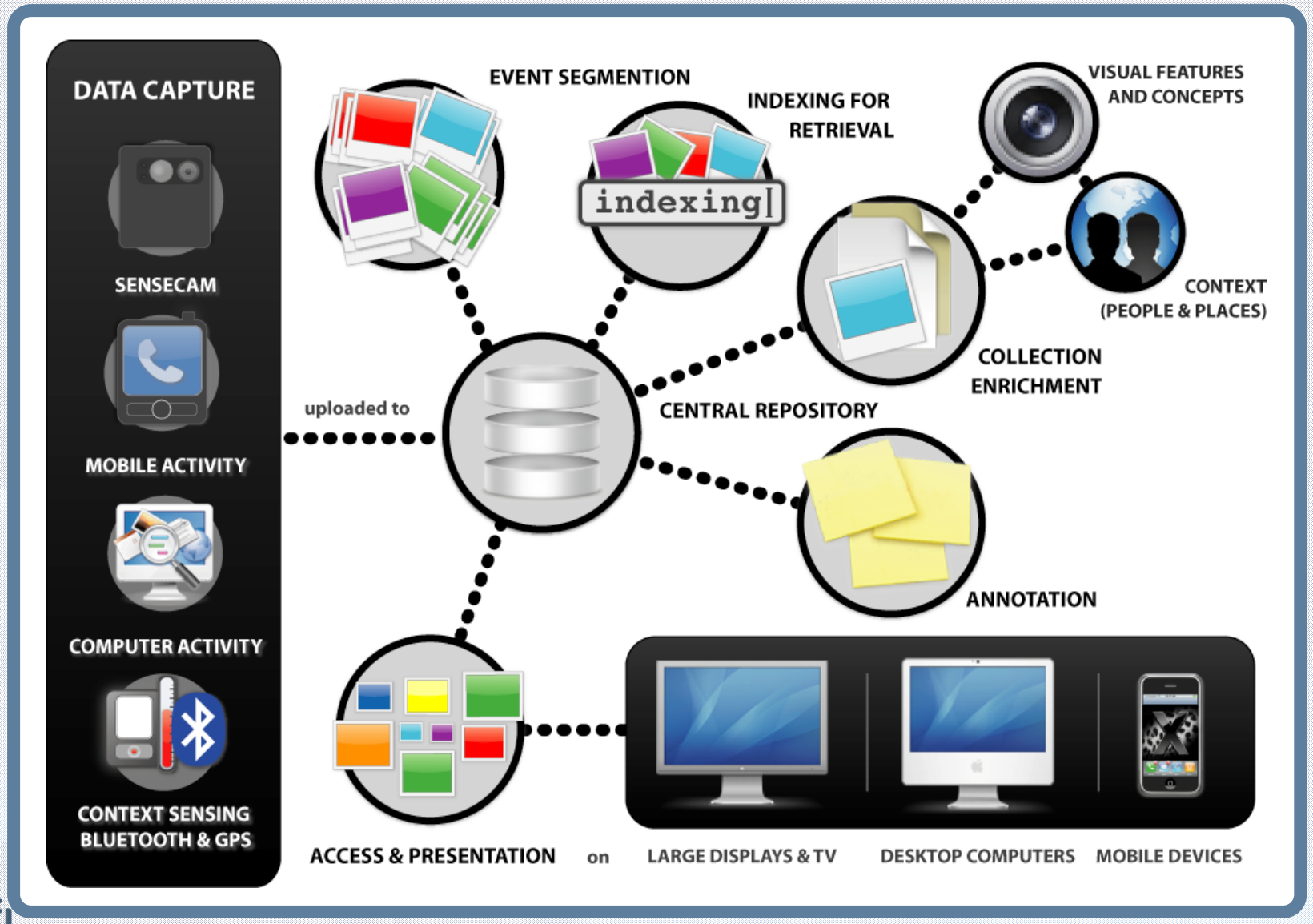
Suite of Sensecam specific concept detectors under development (with MediaMill):

- Steering wheel (72%)
- Shopping (75%)
- Inside of vehicle when not driving (airplane, taxi, car, bus) (60%)
- Toilet/Bathroom (58%)
- Giving Presentation / Teaching (29%)
- View of Horizon (23%)
- Door (62%)
- Staircase (48%)
- Hands (68%)
- Holding a cup/glass (35%)
- Holding a mobile phone (39%)
- Eating food (41%)
- Screen (computer/laptop/tv) (78%)
- Reading paper/book (58%)
- Meeting (34%)
- Road (47%)
- Vegetation (64%)
- Office Scene (72%)
- Faces (61%)
- People (45%)
- Grass (61%)
- Sky (79%)
- Tree (63%)

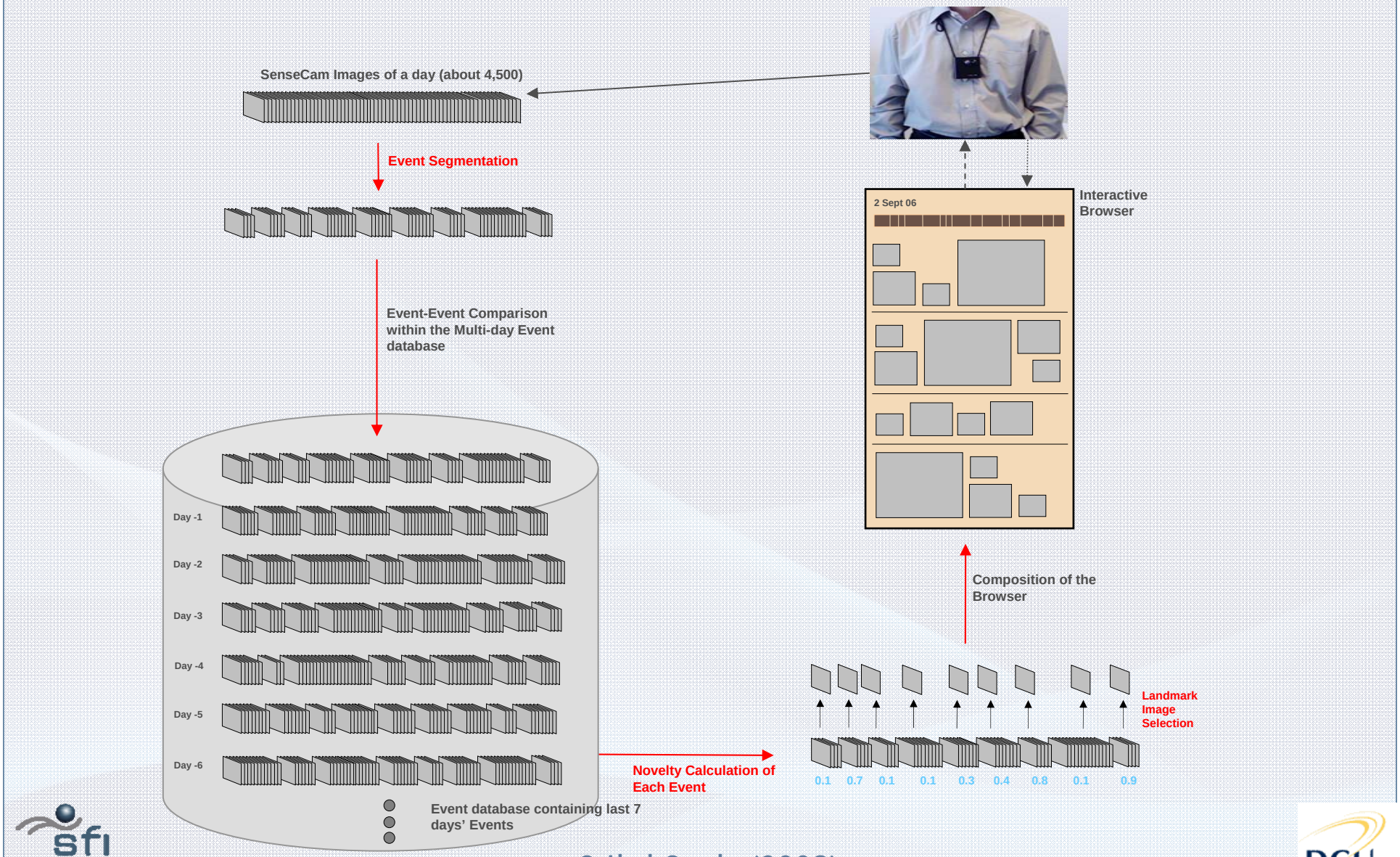
3. Integration of Context

- Shown to work for photo search
- We integrate:
  - GPS location
  - Social context, e.g. people near me,
  - Environmental context, e.g. temperature, light, acceleration
  - Biometric metadata
  - Communication interactions

4. Indexing and Presentation



4. Indexing and Presentation



My Visual Diary with SenseCam

hlee: 76,430 photos (25 days)

MY ACCOUNT | SIGN OUT | ABOUT

CALENDAR

◀ MAY ▶ 2006

S	M	T	W	T	F	S
30	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31	1	2	3
4	5	6	7	8	9	10

DURATION ▶

CAPTION SEARCH

WEEKLY SUMMARY

Selected day is shown below in the context of whole week. Move mouse cursor over to see other similar Events in the week

S
M
T
W
T
F
S

29 May 2006

19 EVENTS

Drag the slider bar to adjust the number of Important Events

0 am | 6 am | 12pm | 6pm | 0am

I was chatting with Gareth on the conference in July. Quite a few chats today !

ADD TO FAVE | FIND SIMILAR

F FAVOURITE (25)

SIMILAR EVENTS

92 Similar Events have been found. Click on the photo to replay all photos within the Event.

| 1 | 2 | 3 | 4 | 5 | 6 |

Sort by: TIME | SIMILARITY | #PEOPLE

16:20 (Duration: 08m 43s)
14 APR 2006 ▶

13:45 (Duration: 14m 05s)
14 APR 2006 ▶

F
10:02 (Duration: 23m 56s)
13 APR 2006 ▶

14:39 (Duration: 15m 30s)
12 APR 2006 ▶

F
11:25 (Duration: 06m 21s)
12 APR 2006 ▶

09:52 (Duration: 01m 03s)
12 APR 2006 ▶

15:19 (Duration: 21m 10s)

Other CDVP Sensecam Activities

Event Augmentation

Location Mapping

Event augmentation – Croke Park

Here's an image from a SenseCam after a Irish football game in Croke Park, Dublin. Can we see other people's pictures of this match.

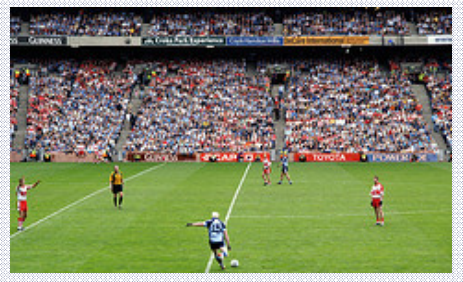
Let's search by location…

Cathal Gurrin (2008)

Event augmentation – Croke Park

- Receive the following pictures…
- Then filter out to just those results from the same day







Cathal Gurrin (2008)

Location Mapping Sensecam

- Experimental integration of location stamped visual lifelog with Visual Mapping Software

- Typical Scenario:

> *I recently visited Asia, find me a sequence of events where I was eating with other people in both Korea and China.*

  - Requires:
    - GPS location stamping
    - Feature detectors for (people, food, eating, etc…)
    - Event Segmentation
      - With Key photo selection

Challenges for HDMS

Indexing & Retrieval Challenges

Challenges for HDMs

- Where to get the data?
  - Very private data, not easily available
- How to evaluate?
  - The data evaluation is inherently subjective to the data gatherer
- Need to:
  - minimise the cognitive load during indexing.
  - organise into events (dynamic?).
  - automatically apply semantic enrichment techniques
  - provide fast search facilities for millions of photos per year, many tens of millions of sensor data, and tens of thousands of human interactions.
  - automatically link between related events
  - provide high precision of retrieval.
  - support context sensitive retrieval

Human Digital Memories
My thoughts…

Experiences
Thoughts
Who would do this?
Why?

Experiences of wearing a Sensecam

- I have been gathering a HDM (visual & location) since June 2006
  - Over 2.5 million SenseCam images
  - Each with GPS position
- Summary of my experiences:
  - Most people don't notice the camera
    - Those that do always remember!
    - Most people don't mind the camera
  - About 40% of photos captured are low quality, more are stop-photos (boring photos of typical scenes like driving or working at desk).
  - Need extremely understanding family, girlfriend and friends

Millionth Image

Favourite Image

Some comments…

- Importance of contextual semantic enrichment
- Event browsing is key
  - Too many photos to browse, need event summary and then 'drill down' to view event in detail if required
  - Stop events, (like work desk and driving) can be hidden.
- 'Total Recall', little sign of 'Event Decay'
  - I remember nearly every (non stop-) event when I see it…
- How do I want to interact with my HDM?
  - I want seamless integration with my context (recommendation)
  - I want powerful and fast search
    - Not only time/date, but location, people, biometrics, …
  - I want blog-style sharing of important events
    - Ideally directly from the device

Why would people do this?

- Because we can:
  - My life on 1TB of photos, audio another TB.
  - In 2020 2TB = €10
  - People blog, so this is an extreme extension of blogging
- Health Benefits of Lifelogging
  - The Frammington Study has shown the benefits of logging an aspect of peoples lives.
  - Much research into Sensecam wearing for Alzheimers sufferers.
- Who would do this?
  - Security/Health professionals
  - people with memory difficulties
  - Bloggers & people who want an automatic diary of their life

Xie Xie

http://www.cdvp.dcu.ie

Acknowledgements:

Microsoft

Irish Research Council for Science Engineering and Technology

Science Foundation Ireland (grant 07/CE/I1147)

Questions ?