

Providing Effective Memory Retrieval Cues through Automatic Structuring and Augmentation of a Lifelog of Images

Aiden Doherty

(Supervisor: Prof. Alan F. Smeaton)

Centre for Digital Video Processing (CDVP) &
CLARITY: Centre for Sensor Web Technologies,
Dublin City University

CLARITY: Centre for Sensor Web Technologies

- Recently announced CSET (Centre for Science Engineering & Technology)
- Funded by Science Foundation Ireland (SFI) with industry contributions
- 5 year duration, following on from previous 4-year “Adaptive Information Cluster”
- Administrative centre in UCD, researchers in DCU, UCD and Tyndall Institute
- Within DCU involves CDVP (Computing & EE), NCSR, Health & Human Performance

CLARITY What ? “The Sensor Web”

- Increasing availability of cheap, robust, and deployable sensor technologies ushering in a wave of new information sources;
- Ubiquitous, dynamic, noisy, reactive and yielding unstructured data-streams == sensor web
- Realizing the sensor web demands a large-scale, multi-disciplinary research effort == CLARITY
- Moving beyond our research silos to novel research interactions;
- Demonstrator projects in:
 - Personal health and wellness;
 - Environmental monitoring;

Principal Investigators

Prof. Barry Smyth	- <i>Personalization, recommender systems, mobile computing</i>
Prof. Alan Smeaton	- <i>Content-based information retrieval</i>
Prof. Dermot Diamond	- <i>Materials research, wearable sensors</i>
Prof. Noel O'Connor	- <i>Audio-visual analysis, multi-modal information processing</i>
Mr. Gregory O'Hare	- <i>Ubiquitous computing, multi-agent systems</i>

Associate PIs

Prof. Paddy Nixon	- <i>Pervasive computing, middleware, security, trust, privacy</i>
Prof. Niall Moyna	- <i>Sports Science, wearable sensing</i>
Dr. Simon Dobson	- <i>Middleware, pervasive computing</i>
Dr. Cian O'Mathuna	- <i>Sensor devices, energy-aware hardware</i>
Dr. Brian Caulfield	- <i>Physiotherapy, therapeutic gaming, wearable sensors</i>

Funded Collaborators

Chris Bleakley (UCD), **Conor Brennan (DCU)**, Rem Collier (UCD), **Brian Corcoran (DCU)**, **Cathal Gurrin (DCU)**, Neil Hurley (UCD), Lorraine McGinty (UCD), **Kieran Moran (DCU)**, **Kieran Nolan (DCU)**, Brendan O'Flynn (TNI), **Donal O'Gorman (DCU)**, **Brett Paull (DCU)**, Emanuel Popovici (TNI), Aaron Quigley (UCD), **Mark Roantree (DCU)**

Overview

- **WHAT IS LIFELOGGING?**
- THE HUMAN MEMORY SYSTEM
- AUTOMATED PROCESSING TO SUMMARISE SENSECAM DATA
 - Segmenting Sequences of Images into Events
 - Retrieval of Similar Events
 - Determining Important Events
 - Selecting Optimal Keyframe
 - Augmenting Events
 - Lifestyle Signatures
- CONCLUSIONS
 - Future Work: Aggregating Data Sources

Lifelogging

Lifelogging is about digitally recording your daily life

Sometimes its for a reason

Work e.g. security personnel, medical staff, etc.

Personal e.g. diaries, etc.

Sometimes its for posterity

Recording vacations, family gatherings, social occasions

Sometimes its because we can

And we're not yet sure what we'll do with it e.g. MyLifeBits

Memory Aids Through the Ages: Cave Paintings (approx. 30,000 years ago)



Cave painting, Lascaux, France, 15,000 to 10,000 B.C.

Memory Aids Through the Ages: **Storytelling**

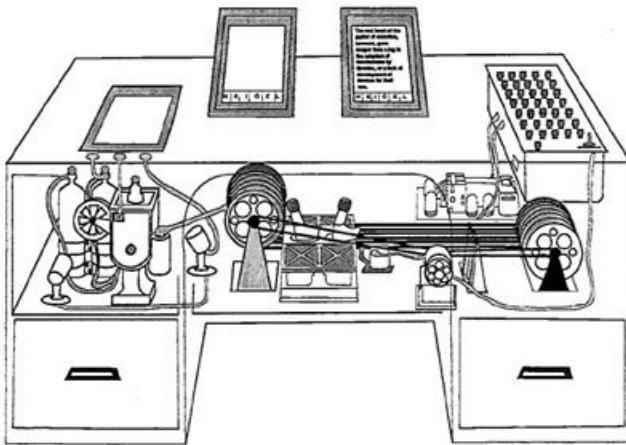


Memory Aids Through the Ages: **Books** (approx 5,000 years ago)

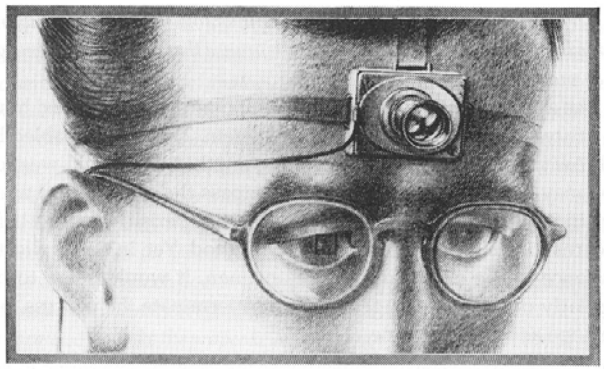


MEMEX

The first lifelogging scenario



In 1945, Vannevar Bush wrote 'As We May Think'... a prophetic view of computing technology:



- Hyperlinks & WWW
- Modelling associative memory
- Miniature wearable camera
- Lifelogging

Lifelogging Devices

Tano *et. al.* University of Electro-Communications, Tokyo, Japan



Lifelogging Devices

Lin & Hauptmann, Carnegie Mellon, PA, USA



SenseCam

Multi-sensor device

- Colour camera
- 3 accelerometers
- Light meter
- Passive infrared sensor

1GB flash memory storage

Smart image capture ~3 images/min

We've had access to 7 SenseCams in last 2 years



How to Review Images?

Make a 2 minute movie of your day!



Lifelogging Aiding Memory

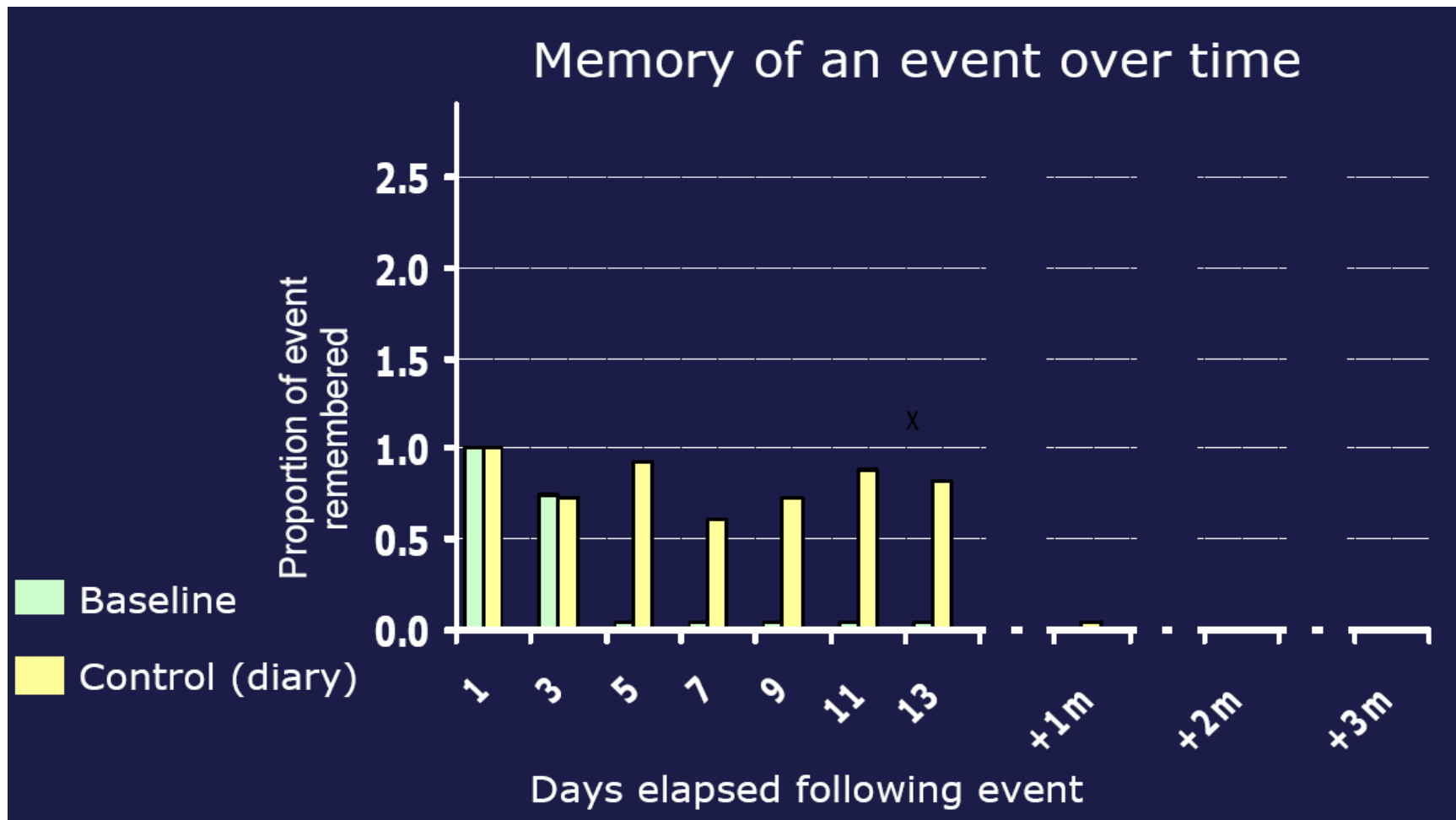
- Preliminary Study carried out by Cambridge Memory Clinic, Addenbrooke's Hospital
- 63 year old, well-educated married woman, with limbic encephalitis (usually has no memory a few days after an event)
- Attends events along with her partner

SenseCam as a Memory Aid



Microsoft Research Cambridge presentation: http://research.microsoft.com/~shodges/presentations/UBICOMP_senseCam.pdf

SenseCam as a Memory Aid



Microsoft Research Cambridge presentation: http://research.microsoft.com/~shodges/presentations/UBICOMP_senseCam.pdf

SenseCam as a Memory Aid



Microsoft Research Cambridge presentation: http://research.microsoft.com/~shodges/presentations/UBICOMP_senseCam.pdf

Overview

- WHAT IS LIFELOGGING?
- **THE HUMAN MEMORY SYSTEM**
- AUTOMATED PROCESSING TO SUMMARISE SENSECAM DATA
 - Segmenting Sequences of Images into Events
 - Retrieval of Similar Events
 - Determining Important Events
 - Selecting Optimal Keyframe
 - Augmenting Events
 - Lifestyle Signatures
- CONCLUSIONS
 - Future Work: Aggregating Data Sources

Memory Systems of Interest

- SENSORY
- SHORT-TERM
- LONG-TERM
 - PROCEDURAL
 - DECLARATIVE
 - Semantic
 - EPISODIC/ AUTOBIOGRAPHICAL
 - “Cued Recall” better than “Free Recall” (*Purdy, '01*)
 - Encoding has strong effect on retrieval (*Godden, '75*)

Cued Recall & Visual Encoding

- Visual encodings are very strong (*Brewer, '88*)
- Encoding from same perspective/environment as viewer is powerful (*Vargha-Khadem, '01*)
- Memories can be temporally encoded (*Larsen, '96*)
- Distinct memories are more strongly encoded (*Purdy, '01*)
- Memories stored by association (*Baddeley, '04*)

Our Take...

To effectively provide memory retrieval cues using SENSECAM we need to automatically:

- *“Chunk” similar images into distinct events*
- *Suggest more “distinctive” events*
- *“Associate” related events*
- *Provide potentially additional retrieval cues from other sources*

A 2 MINUTE BREATHER!

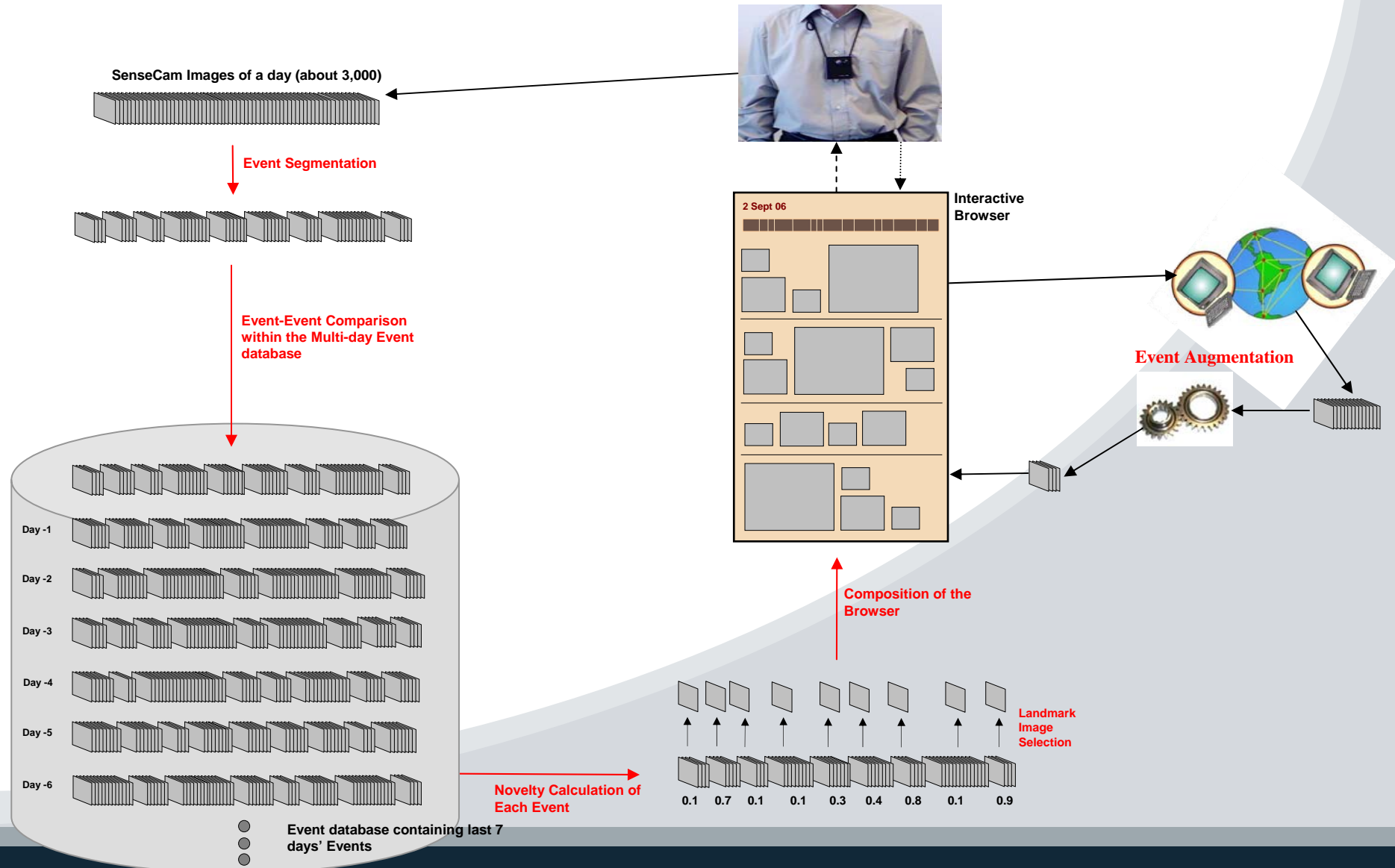
BBC Science Program – October 2008

James May's Big Ideas: Man-Machine

Overview

- WHAT IS LIFELOGGING?
- THE HUMAN MEMORY SYSTEM
- **AUTOMATED PROCESSING TO SUMMARISE SENSECAM DATA**
 - Segmenting Sequences of Images into Events
 - Retrieval of Similar Events
 - Determining Important Events
 - Selecting Optimal Keyframe
 - Augmenting Events
 - Lifestyle Signatures
- CONCLUSIONS
 - Future Work: Aggregating Data Sources

Daily Browser Overview



Event Segmentation

Breakfast



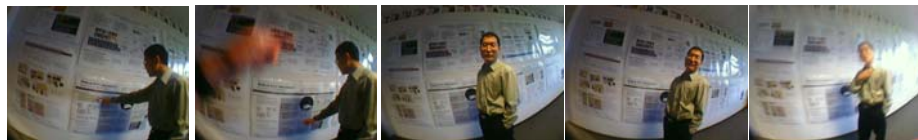
Work



Car



Talking to colleague

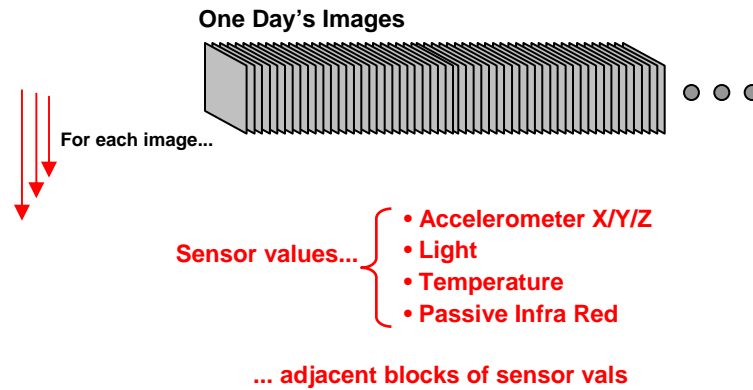


Airplane

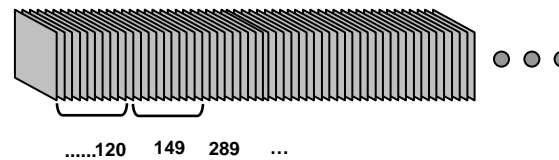


Event Segmentation

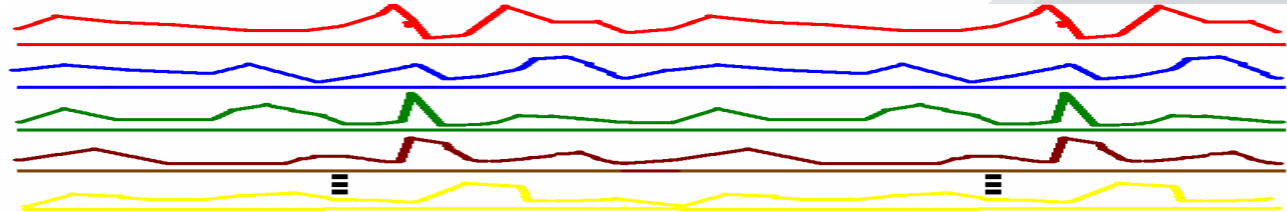
Raw data



Similarity matching



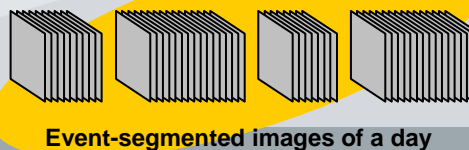
Normalisation & Data fusion



Thresholding



Events



How well does it work?

Data divided into training and test sets with thousands of different combinations evaluated

From groundtruth we noticed:

Average of 1,785 images per user per day

Average of 22 events groundtruthed per day

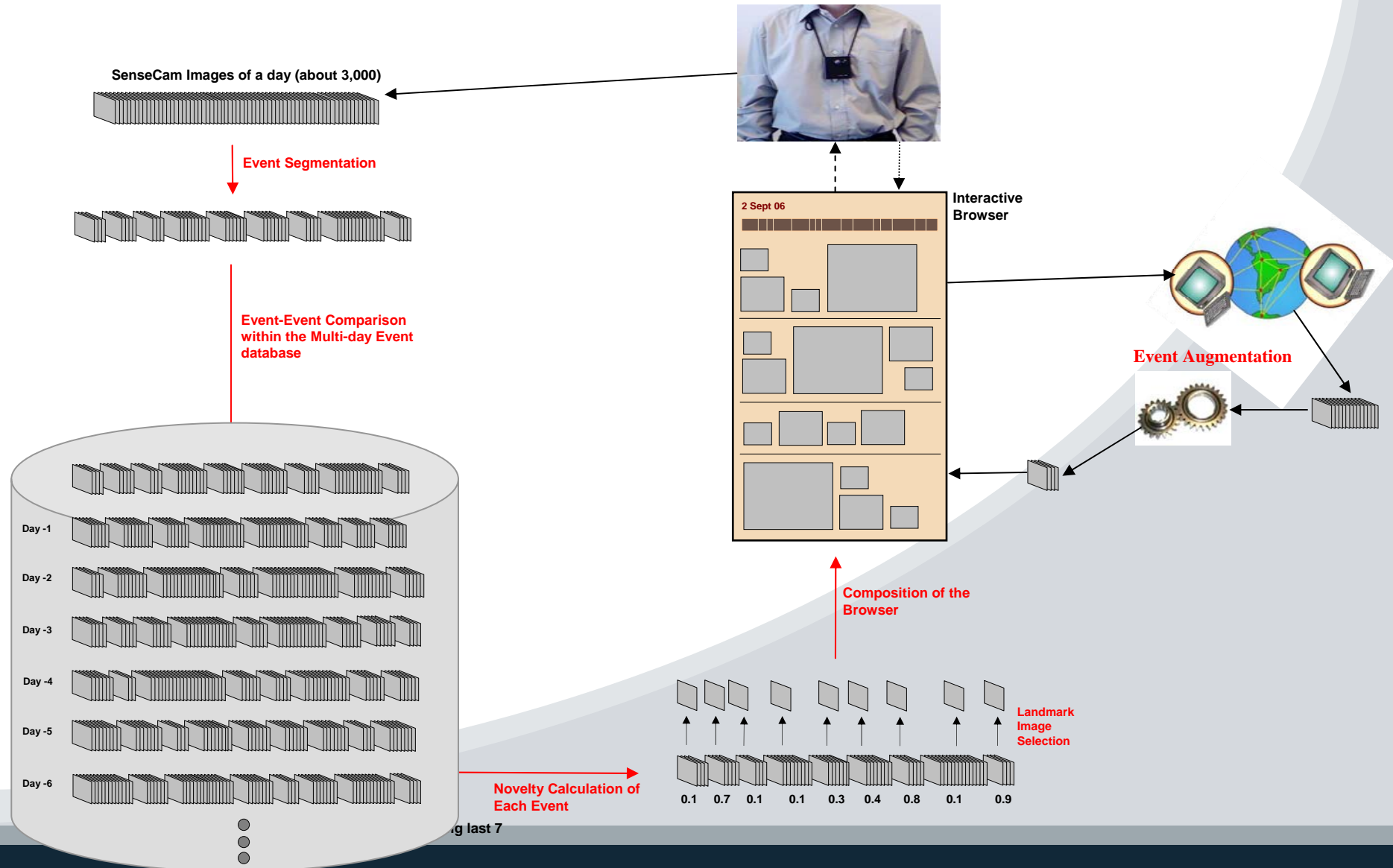
Approach Recommended:

Quick segmentation (sensor values only)

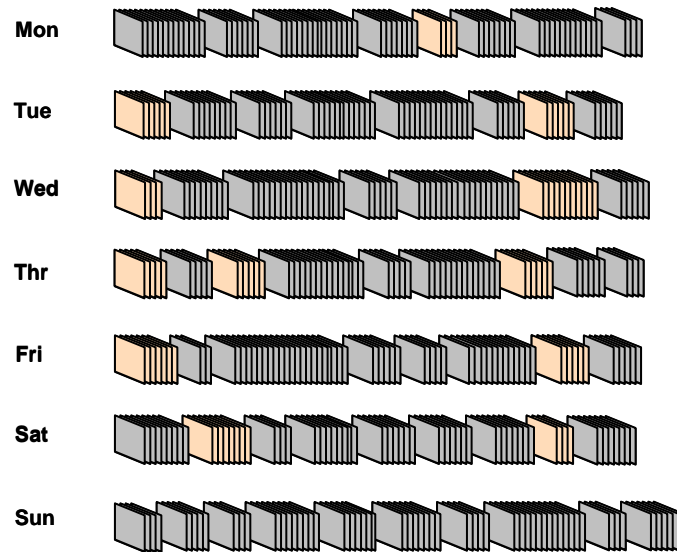
Performance:


F1 score of 60% against *users' semantic* boundaries

Retrieval Reminder



Finding similar events



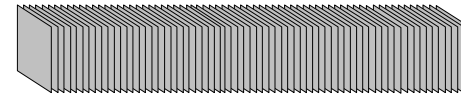
 Similar Events - Aiden waiting for bus

Storing by association

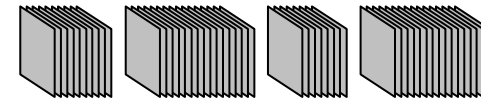
- Events are represented by the average values of all the images present in that event
- Investigated numerous computation approaches to match similarity of any two given events

Event Retrieval

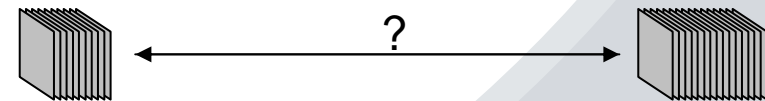
1. Day of ~2,000 SenseCam images



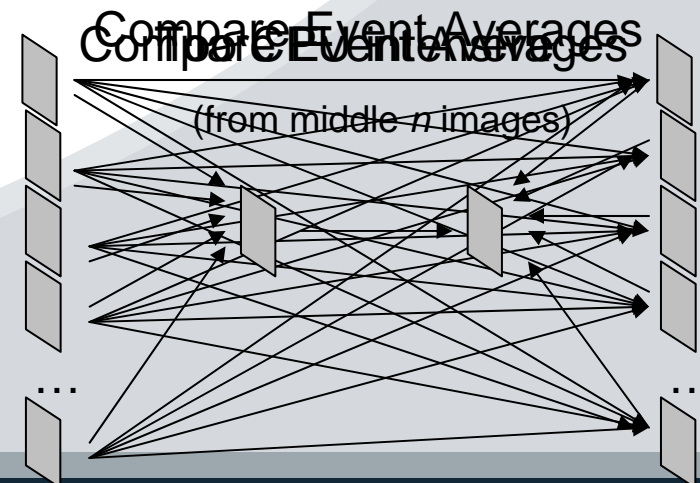
2. Segmented into ~20 events



3. How to compare events to each other?



4. How do we represent events?



SURF Approach

Use [Bay, ECCV 2006] algorithm

Hierarchical visual word vocabulary

Using 7M SURF descriptors and hierarchical K-Means clustering - vocabulary tree with 4096 leaf nodes

L1 distance between histograms of visual words

Ranking

Re-rank top 20 results based on their number of bi-directional matches

SURF Bi-directional Matching

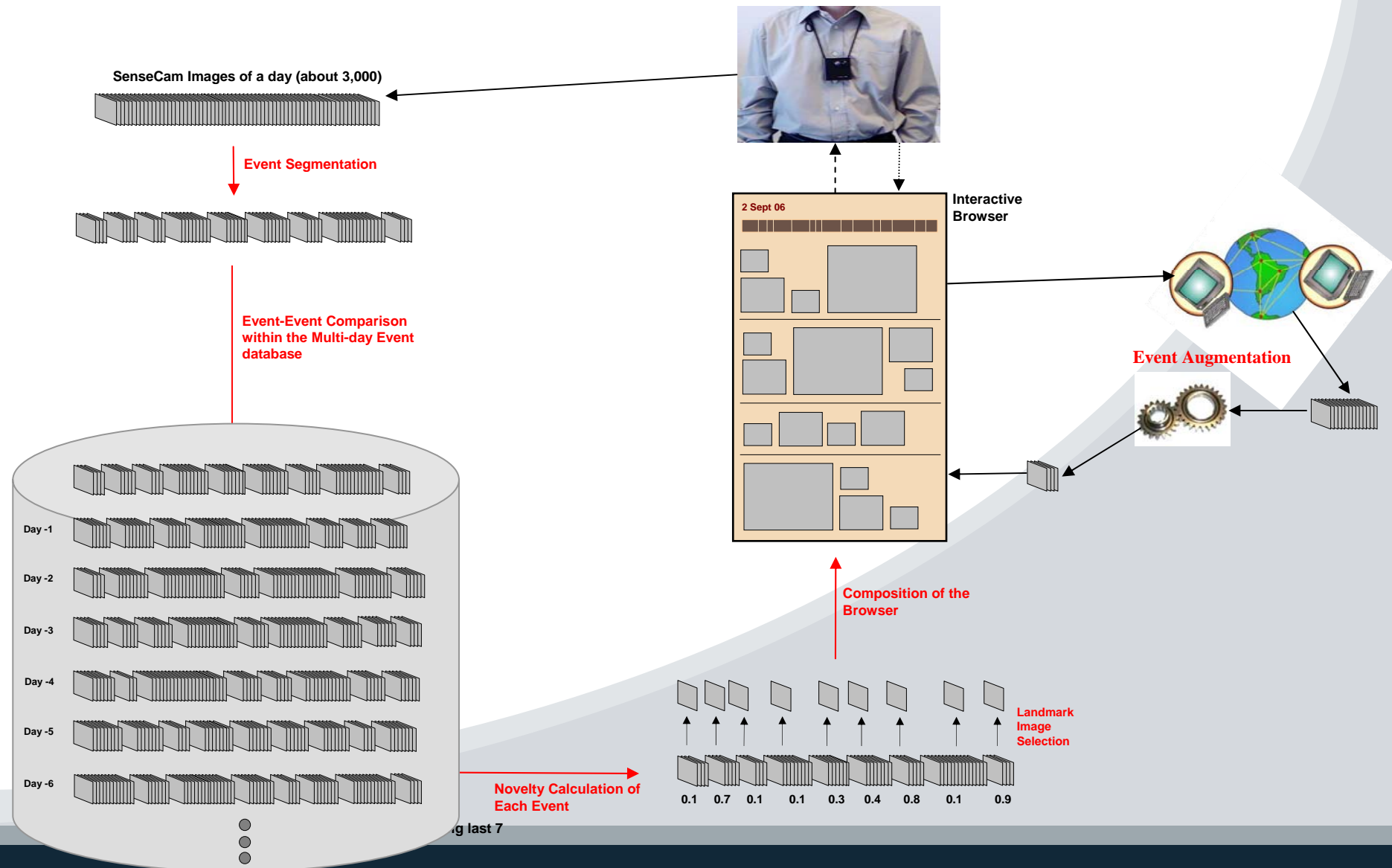


How accurate is it?

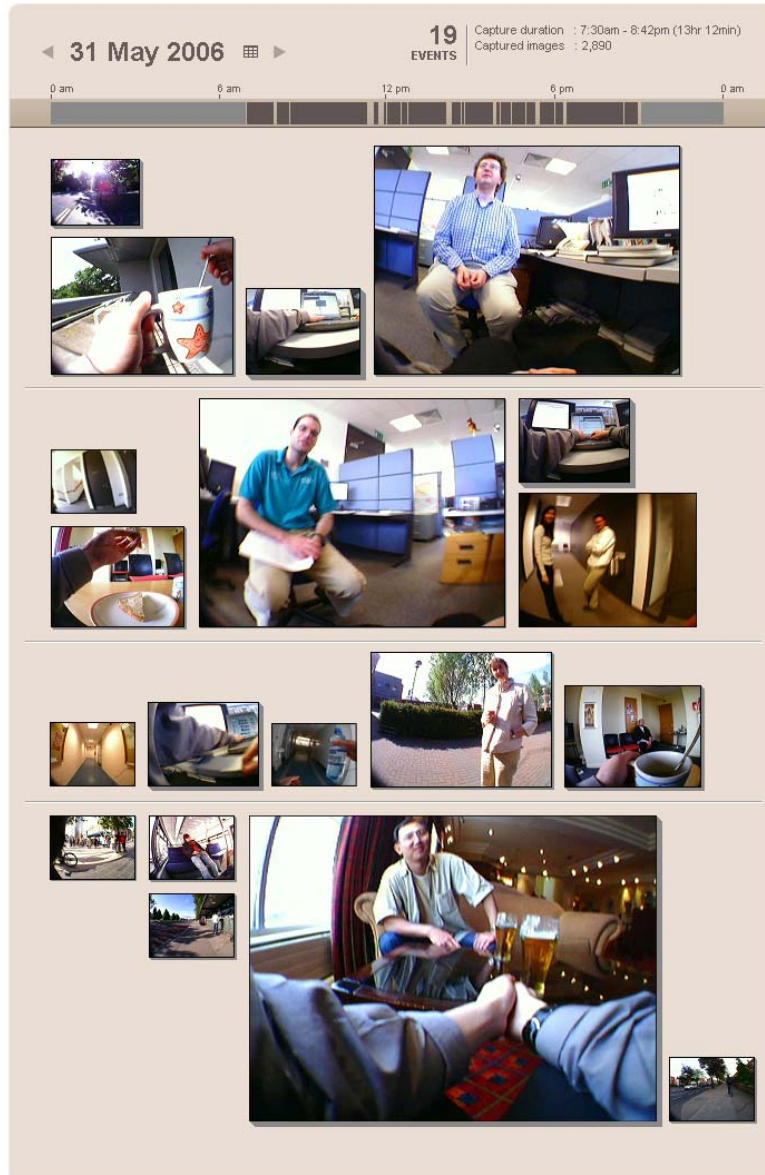
Again thousands of combinations investigated in training phase

- **Event retrieval works well for general queries (69% accuracy of top 5 results)**
 - Can help to remove “clutter” of everyday events from visual diary
- **Retrieval for specific events much more challenging (30% accuracy of top 5 results)**
 - Less events in the collection, lack of semantic meaning
- **Query MAP scores ranging from 0.6% (talking to Lynda) to 94% (Michael at work on his PC)**

Importance Reminder



Importance



- Greater emphasis is placed on important events
- Routine/mundane events can be hidden

Distinctive memories encoded strongly

Automatic Face Detection

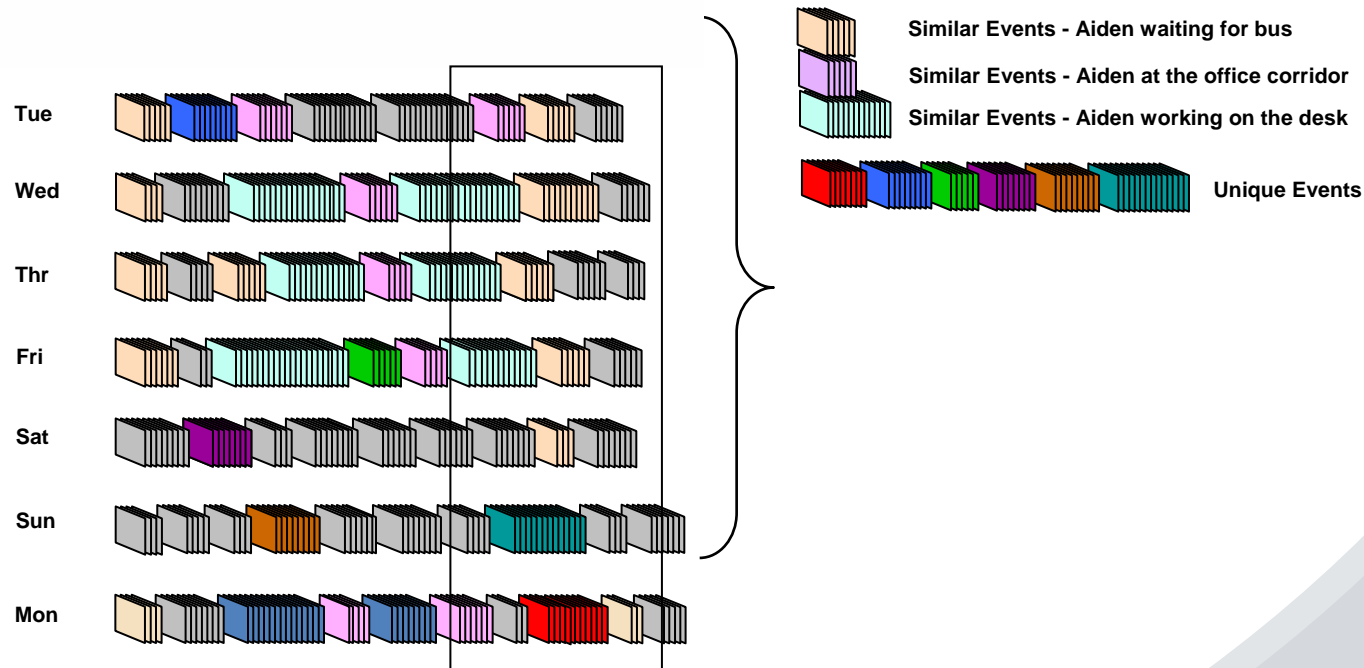


Trained on set of 1,758
SenseCam images

SenseCam images are
low quality

Accuracy = 63%

Novelty to Detect Importance



- Find the most dissimilar event of today by taking the previous 2 weeks into account.
- Also for any event, we only look at how *visually* novel it is with respect to events around the same time from other days

< October 2006 >

Mon	Tue	Wed	Thu	Fri	Sat	Sun
25	26	27	28	29	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

Would you agree that the top 2 events were among the most interesting in your day, and the bottom two were among the least interesting in your day?

- 5 - Strongly Agree
- 4 - Agree
- 3 - Neutral
- 2 - Disagree
- 1 - Strongly Disagree

99% complete!

[Log Out](#)

Most Important Event



2nd Most Important Event



2nd Least Important Event



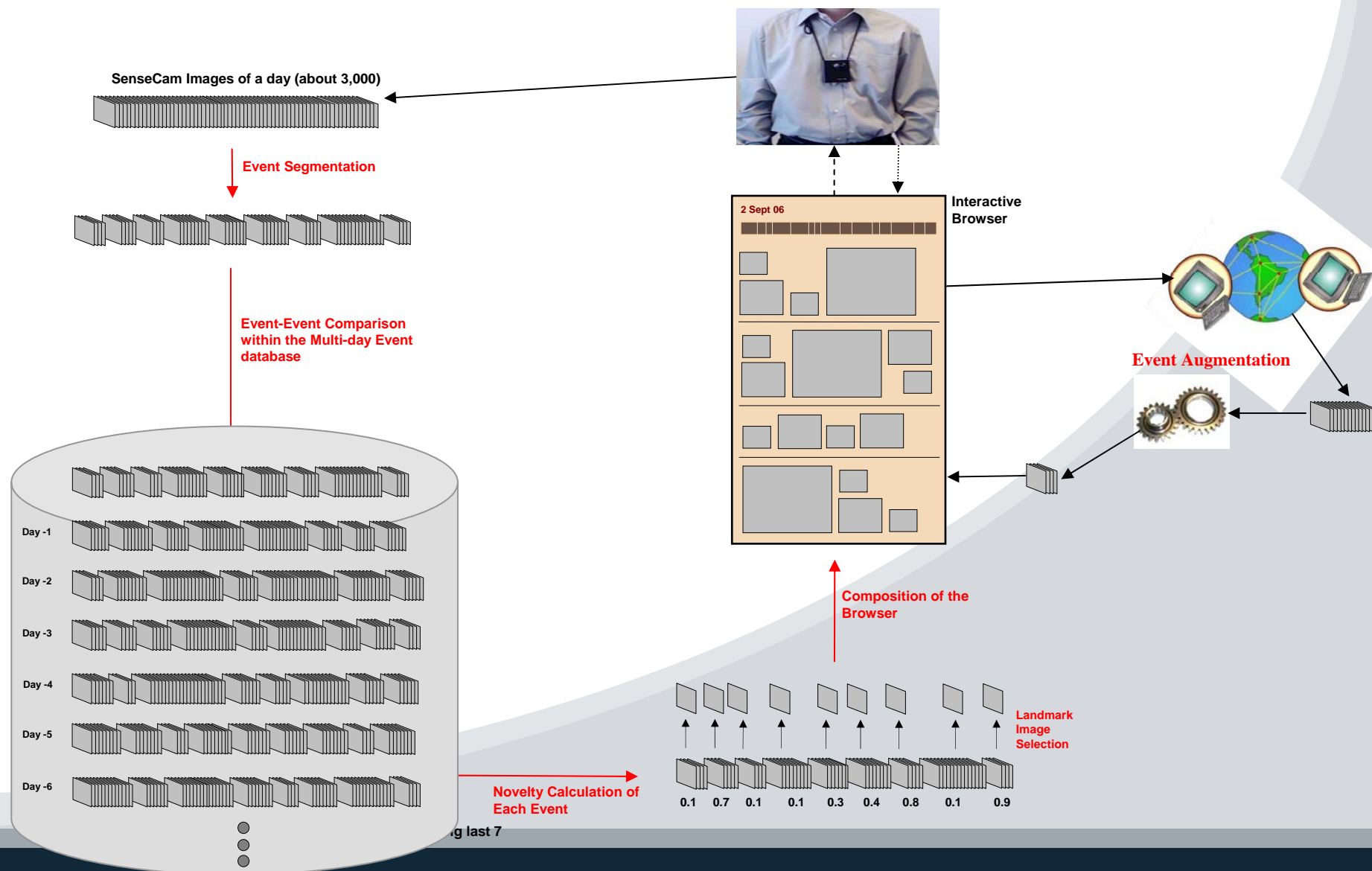
Least Important Event



How well does it work?

- Face Detection good at highlighting most important events
- Novelty good at detecting routine events
- Median Likert score of 4/5, so users generally satisfied

Keyframe Reminder



Keyframe Selection

Standard Approaches

- Middle Image
- Image Quality
- Image closest to others in same event
- Image that distinguishes event best from other events



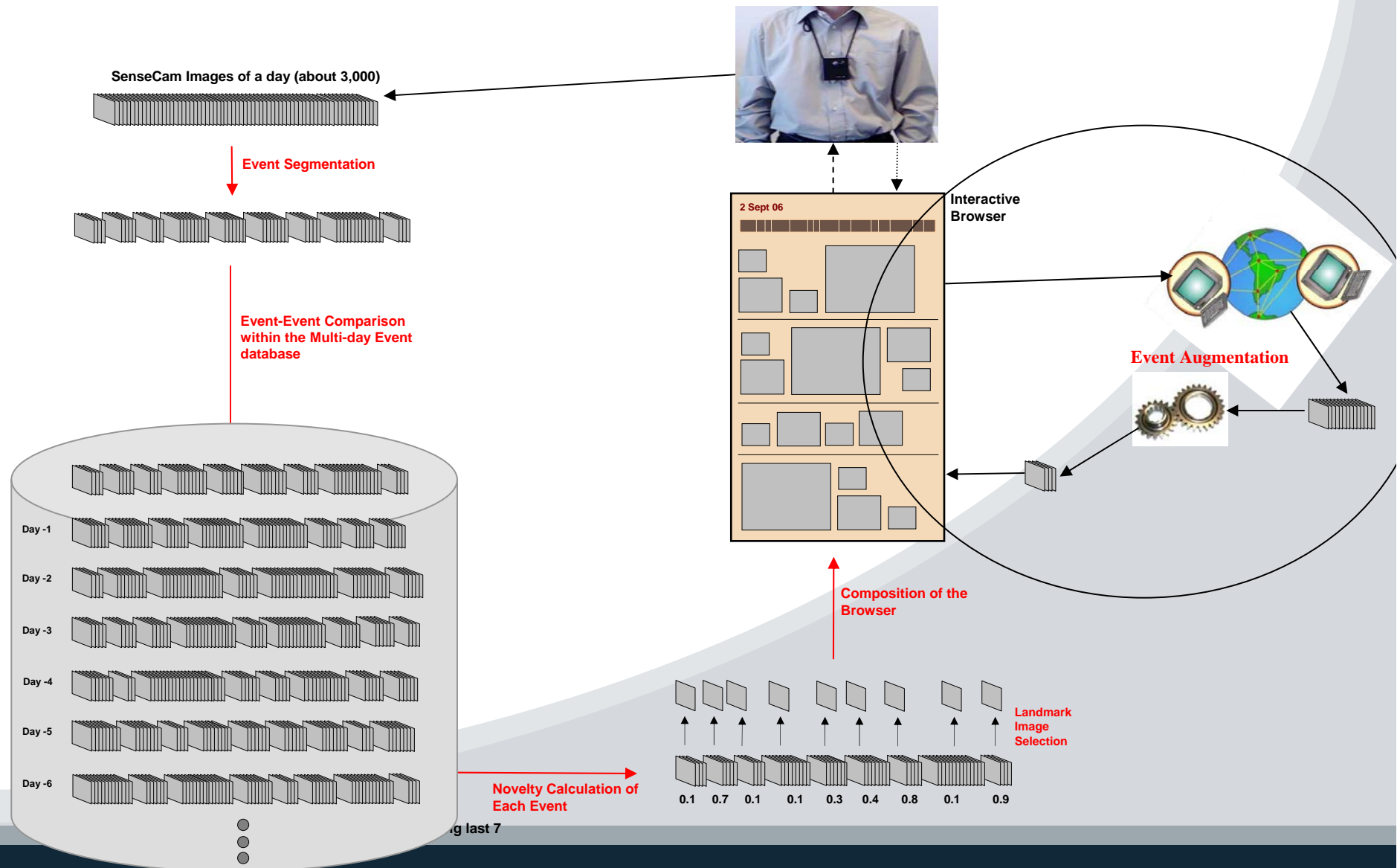
Keyframe Experiments

How well does it work?

- User judgements made on 2,235 events
 - 6 keyframes judgements per event
 - providing a groundtruth of 13,410 judgements

- Selecting highest quality image works best, although selecting middle image is also effective

Augmentation Reminder



Event augmentation

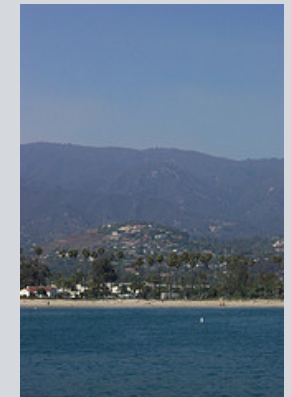
Here's a SenseCam picture of me at a pier in Santa Barbara, CA.

If I have GPS I can search for other pictures in the same location...



Event augmentation

- I receive the following “geotagged” images...
- Then after some processing on text associated with these images we get many more images, and even YouTube videos at times too!



Problem in selecting good tags

Tag	#	Tag	#
-----	---	-----	---

- Use Yahoo Spell Checker to ammend name

nyc	11	thesphere	2
newyorkcity	9	us	2
manhattan	9	warmemorial	2
batterypark	8		

- Know place from GPS
- Use Gazetteer to get placename
- Use WordNet to expand possible placenames

ny	3	geotagged	2
downtown	3	gothamist	2
eagle	3	island	2
america	3	sculpture	2

- We go with 4

No spaces in tags e.g.
“statueofliberty”

Country/region name creates a lot of
noise e.g. “newyork”,
“unitedstates”, “nyc”, etc.

How many tags to select as text for
next query?

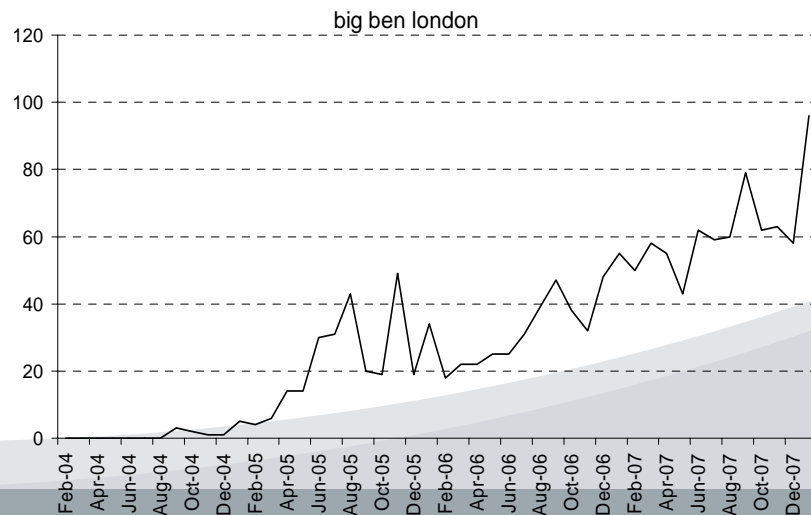
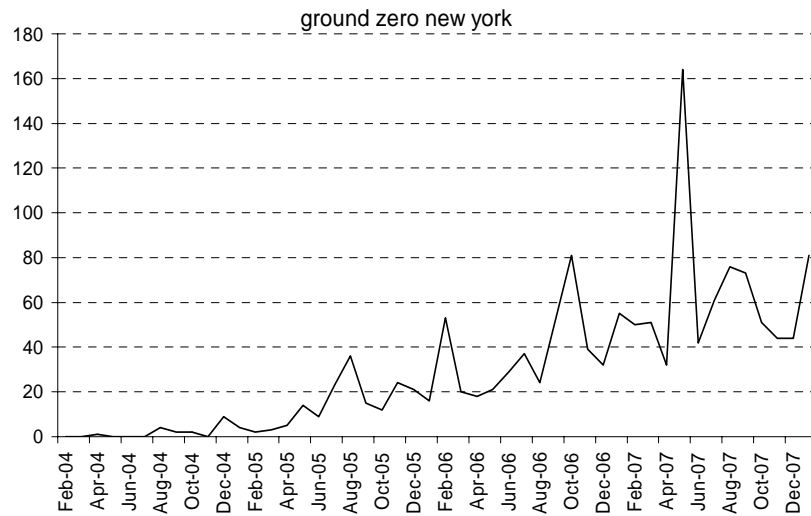
Using those tags...

We can search for material from:

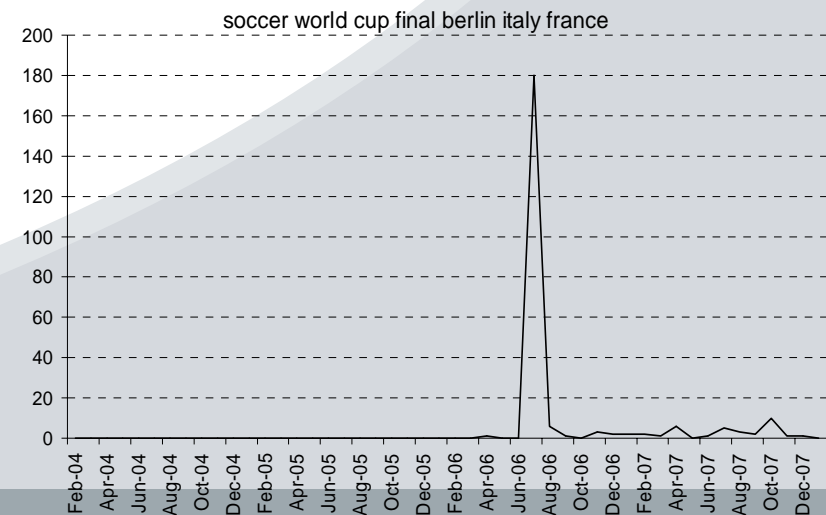
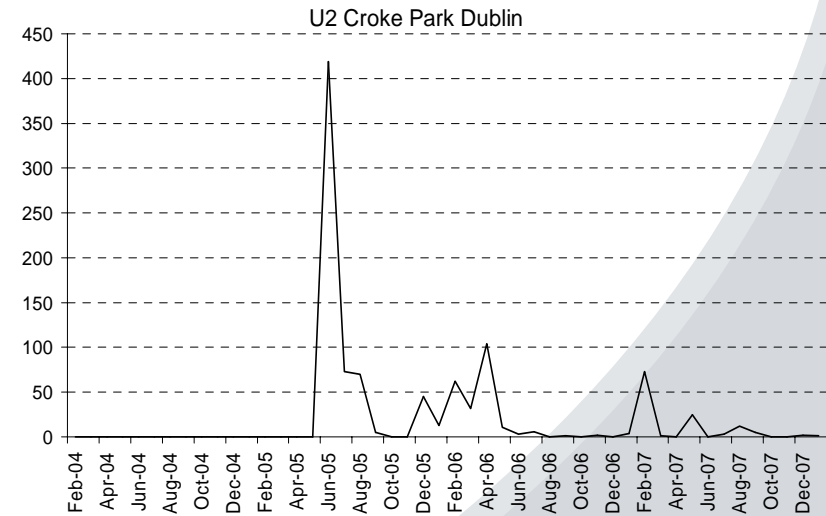
- Flickr
- Yahoo Search Engine
- YouTube
- MSN Search Engine
- Original geotagged images: Flickr & Panoramio

Photo Upload Temporal Aspects

PLACES



EVENTS



Augmentation

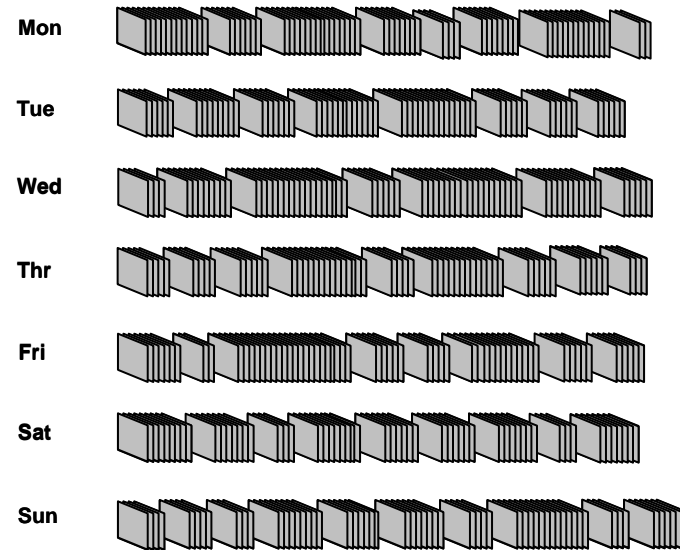
How well does it work?

- 11 users collected 1.9 million images
 - from which 67 events were selected to be augmented
- Users very satisfied with augmentation results of famous tourist locations e.g. Sagrada Familia
- Specific events still a challenge e.g. Tuesday night's Champions League match in Camp Nou

Overview

- WHAT IS LIFELOGGING?
- THE HUMAN MEMORY SYSTEM
- **AUTOMATED PROCESSING TO SUMMARISE SENSECAM DATA**
 - Segmenting Sequences of Images into Events
 - Retrieval of Similar Events
 - Determining Important Events
 - Selecting Optimal Keyframe
 - Augmenting Events
 - **Lifestyle Signatures**
- CONCLUSIONS
 - Future Work: Aggregating Data Sources

Can't "recognise" events



We can detect this event

We know when this event is

BUT

We don't RECOGNISE the event i.e. we don't know "the what" of this event



Vehicles External



Road



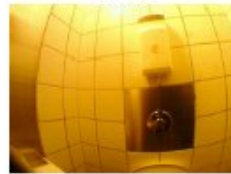
Steering wheel



Inside of vehicle



Indoors



Toilet/Bathroom



Door



Staircase



Outdoors



Buildings



Tree



View of Horizon



Grass



Sky



Vegetation



Screen



Reading



Meeting



Office



Giving Presentation



Food/eating



Hands



Holding a cup/glass



Holding a mobile



Faces

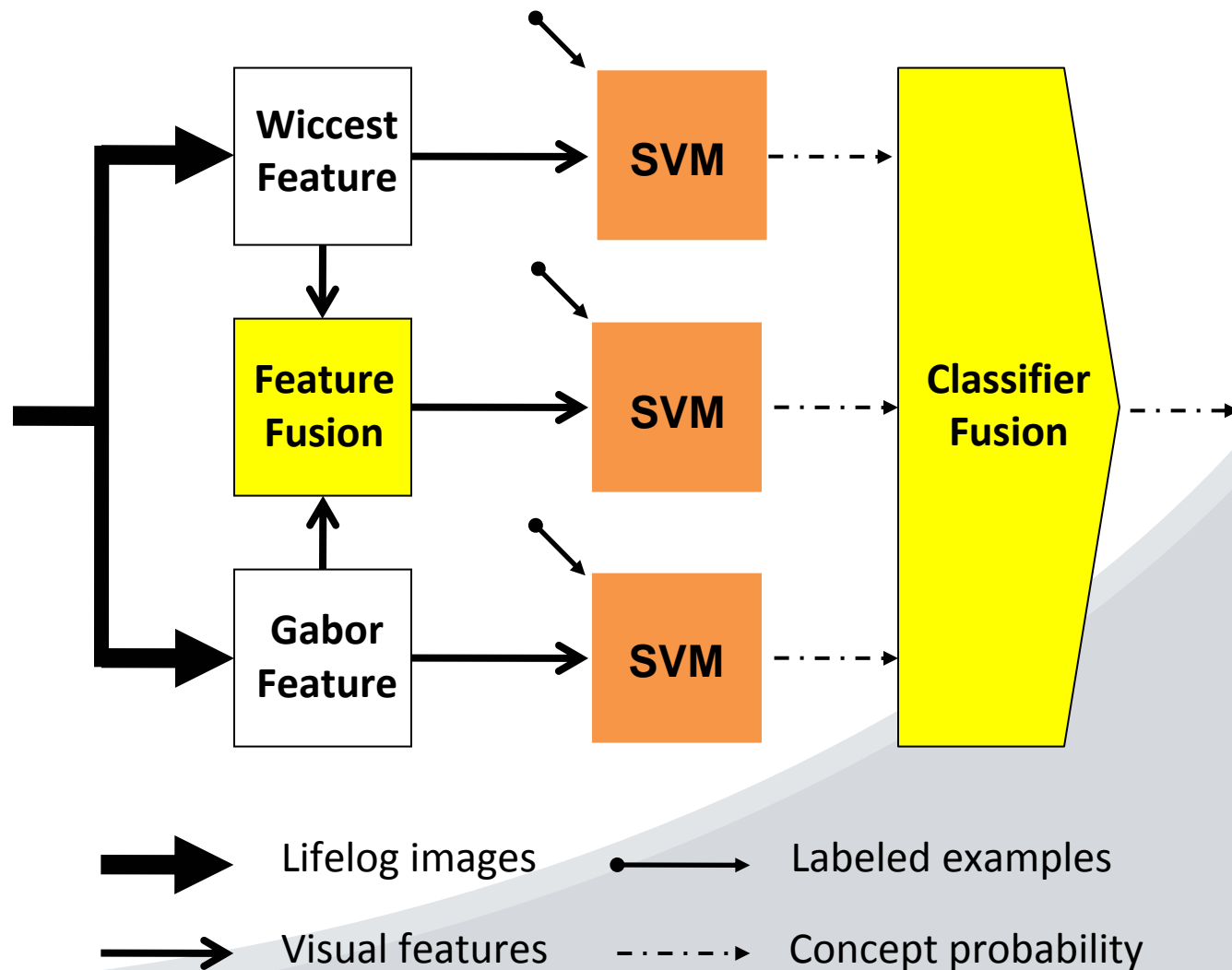
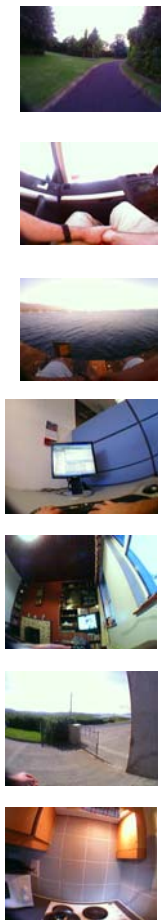


People



Shopping

Concept detection process



Concept Image Accuracy

- Precision
 - Average = 0.57
 - Median = 0.60
- Judge Agreement
 - Fleiss's Kappa = 0.68 (9 judges)
- Strong correlation of 0.75 between the number of concept training samples and test set performance



Vehicles External (46%)



Road (47%)



Steering wheel (72%)



Inside of vehicle (60%)



Indoors (82%)



Toilet/Bathroom (58%)



Door (69%)



Staircase (48%)



Outdoors (62%)



Buildings (59%)



Tree (63%)



View of Horizon (23%)



Grass (60%)



Sky (79%)



Vegetation (64%)



Screen (78%)



Reading (58%)



Meeting (34%)



Office (72%)



Presentation (29%)



Food/eating (41%)



Hands (68%)



Holding cup (35%)



Holding phone (39%)



Faces (61%)



People (45%)



Shopping (75%)

Results

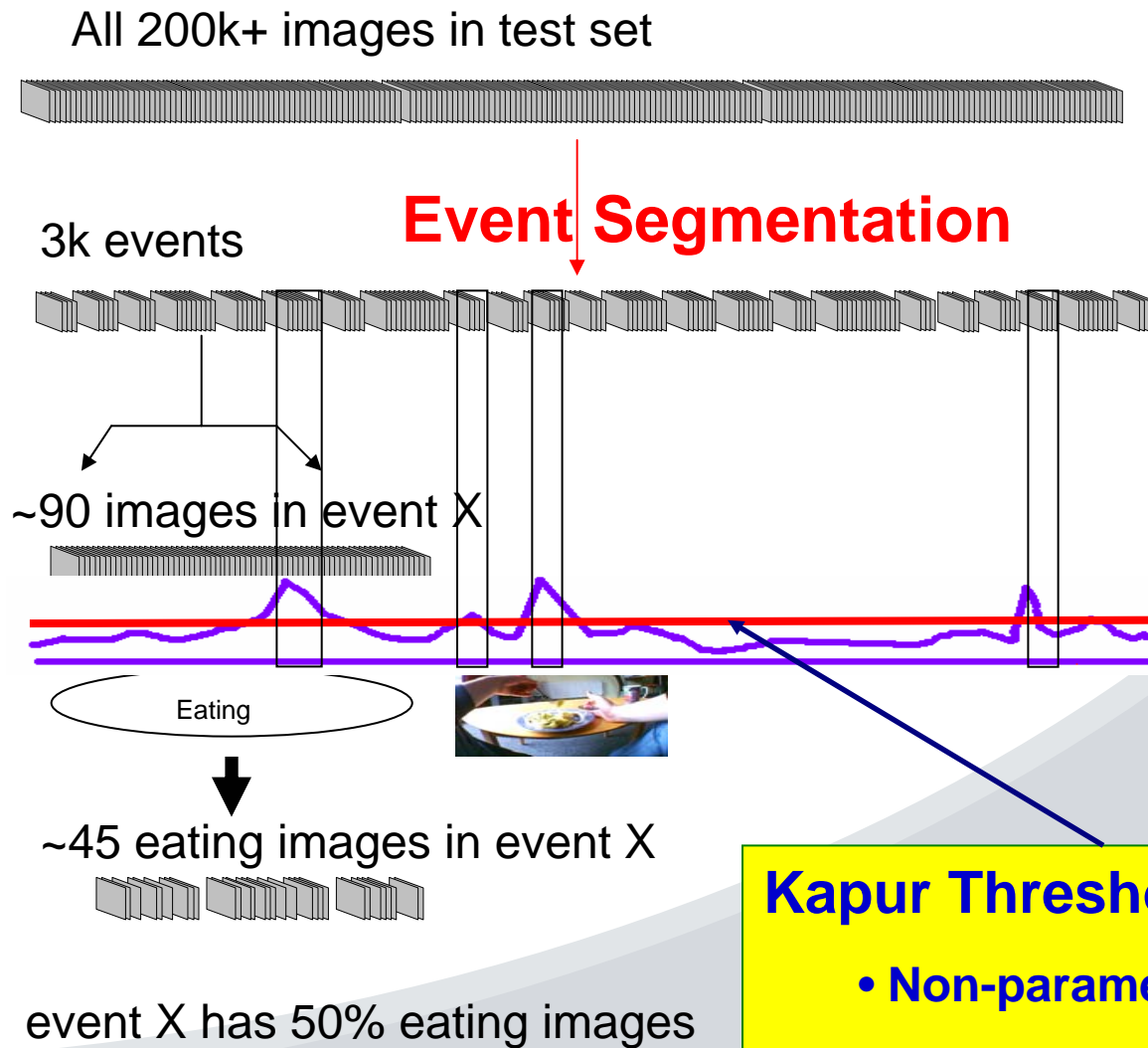
BUT applying on image level isn't so interesting

- Many SenseCam images are blurred, grainy, obscured by hands, etc.

HOWEVER

- **Considering groups of images (i.e. CONSIDERING EVENTS)**
 - Reduces inaccuracies
 - Allows us map “macro trends”

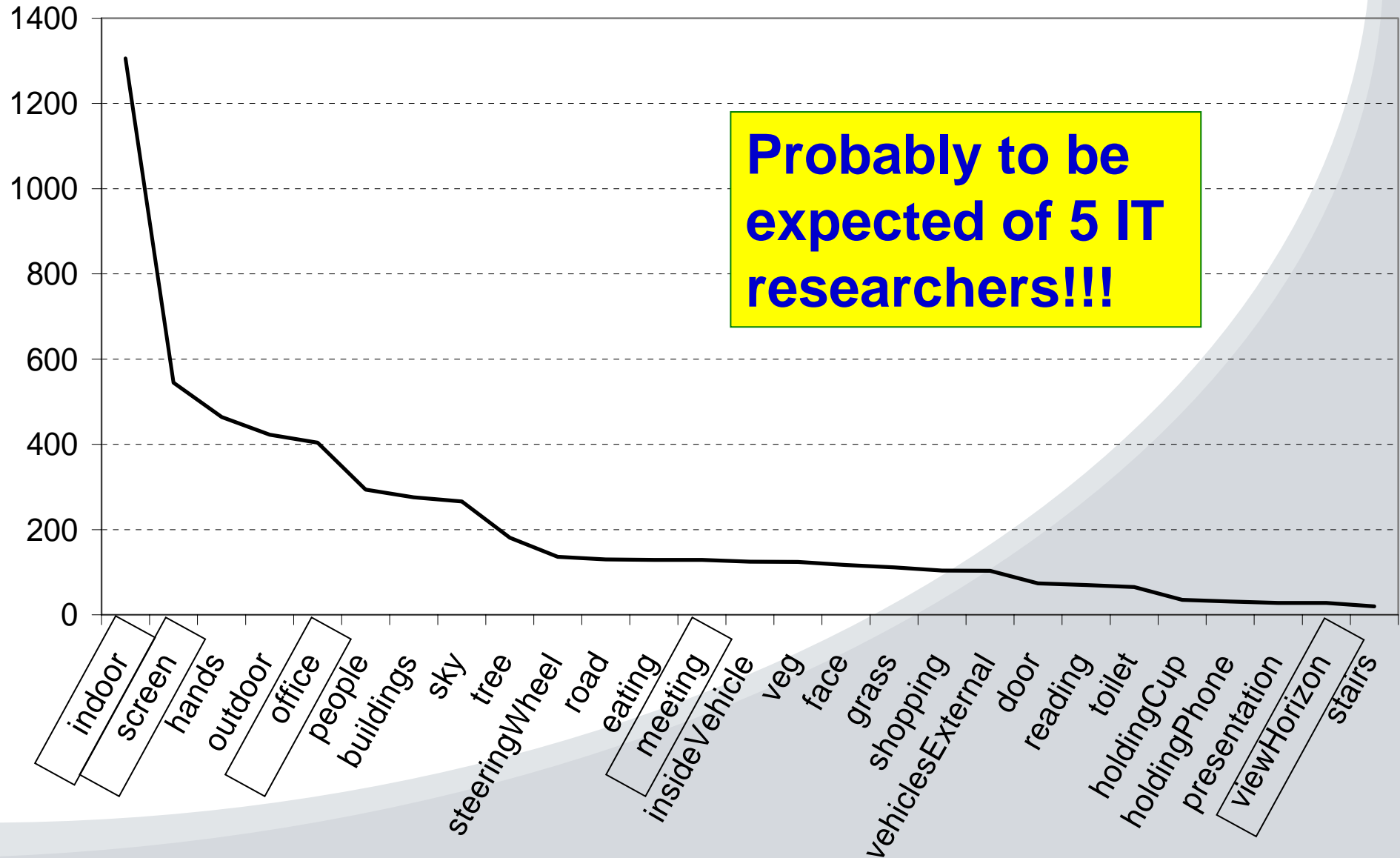
Where are the <eating> events?



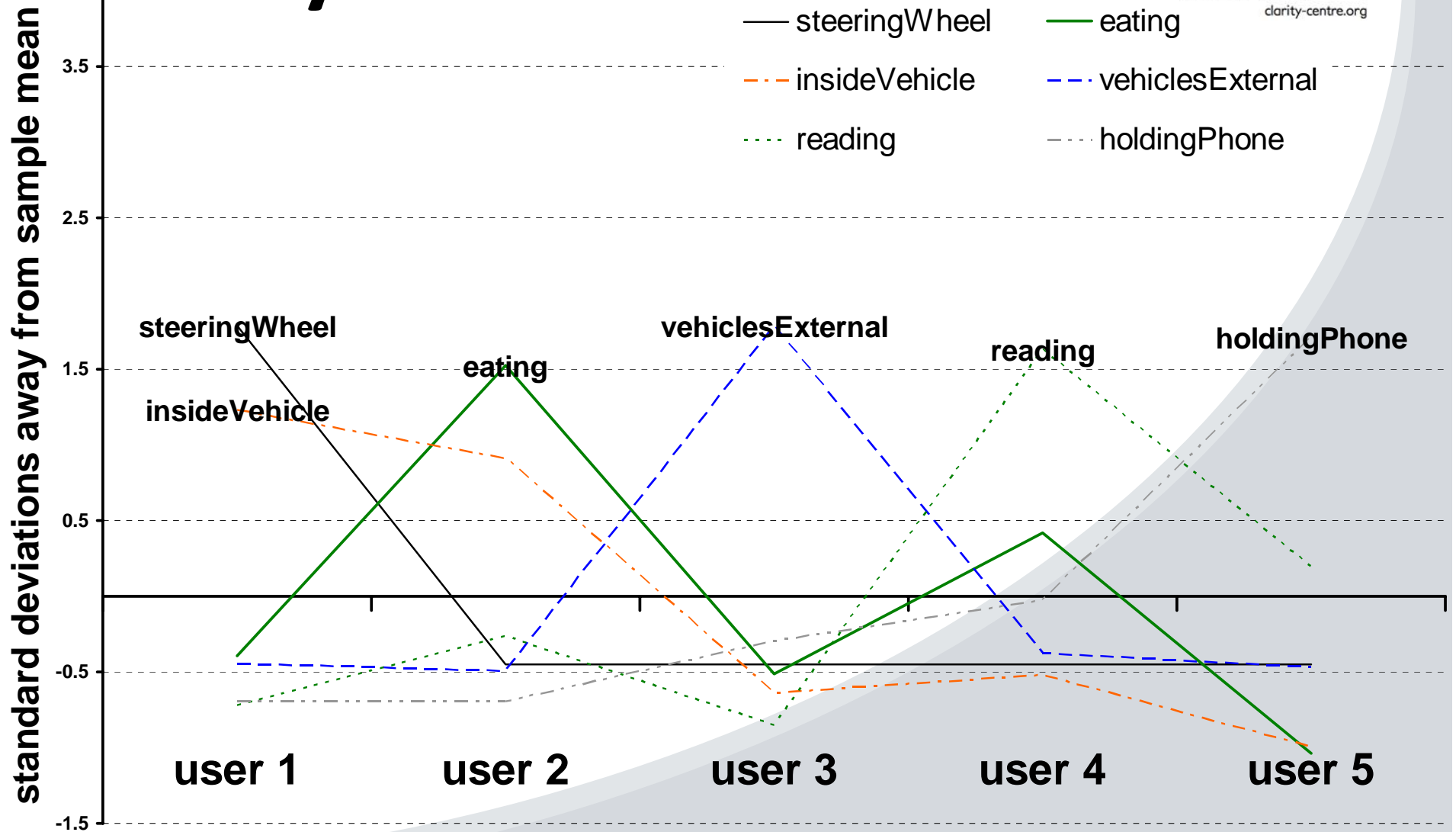
Kapur Thresholding

- Non-parametric
- Entropy based

Num Events Across 5 Users



Lifestyle Variation



Overview

- WHAT IS LIFELOGGING?
- THE HUMAN MEMORY SYSTEM
- AUTOMATED PROCESSING TO SUMMARISE SENSECAM DATA
 - Segmenting Sequences of Images into Events
 - Retrieval of Similar Events
 - Determining Important Events
 - Selecting Optimal Keyframe
 - Augmenting Events
 - Lifestyle Signatures
- **CONCLUSIONS**
 - **Future Work: Aggregating Data Sources**

Conclusions

- For a long time focus of lifelogging community was on hardware minituration and storage
- Recently focus has shifted to data management
- Potential significance of SenseCam as memory aid

Conclusions

- Event Segmentation is pretty accurate, and VERY fast
- Event Retrieval is good for most queries and helps direct the user to “associated/related” events
- Suggested Keyframes are on the whole a good approximation of nearly all events

Conclusions

- Determining event importance generally provides a good “starter cue”
- Augmentation provides many additional images and is especially useful when visiting big tourist sites
- Detecting semantic concepts begins to allow us to analyse the signature of a user’s lifestyle

Future Work

- Exploiting Augmented Images to Construct a Narrative of One's Tourist Trip
 - **Amusing/Interesting stories may provide good memory cues**
- Browsing Lifelog Content on Mobile Devices
 - **Eventually integrate into cell phones**
- Automated Blogging

Future Work

- Diet Monitoring
- Ethnographical Studies
- Leveraging Other Sources of Information

Data Sources: People Near Me

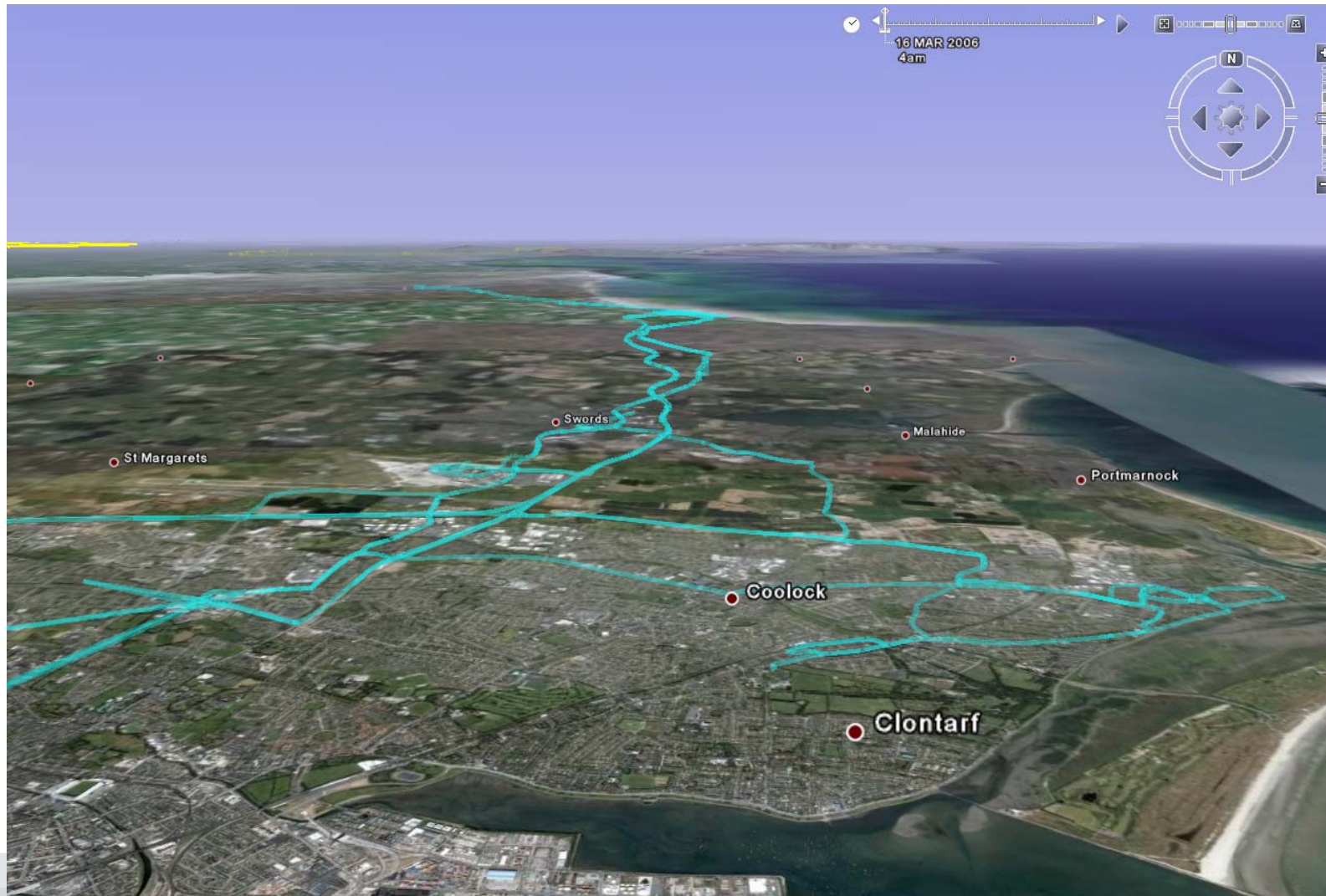
Social Network Generation

Based on real-world interactions using
Bluetooth on mobile devices

This allows us to log who is near to us at
any one time

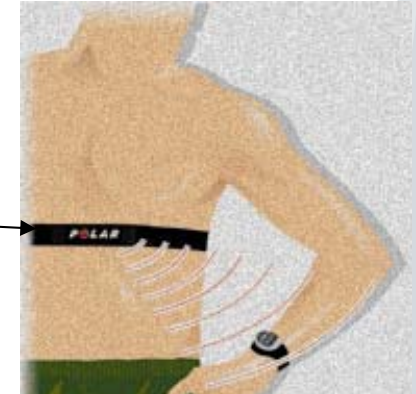


Data Sources: Logging Movement



Data Sources: Logging how we feel...

Heart Rate Monitor



BodyMedia Armband

Galvanic Skin Response (GSR)

Heat Flux

Skin Temperature

Movements



Foster Miller Vest

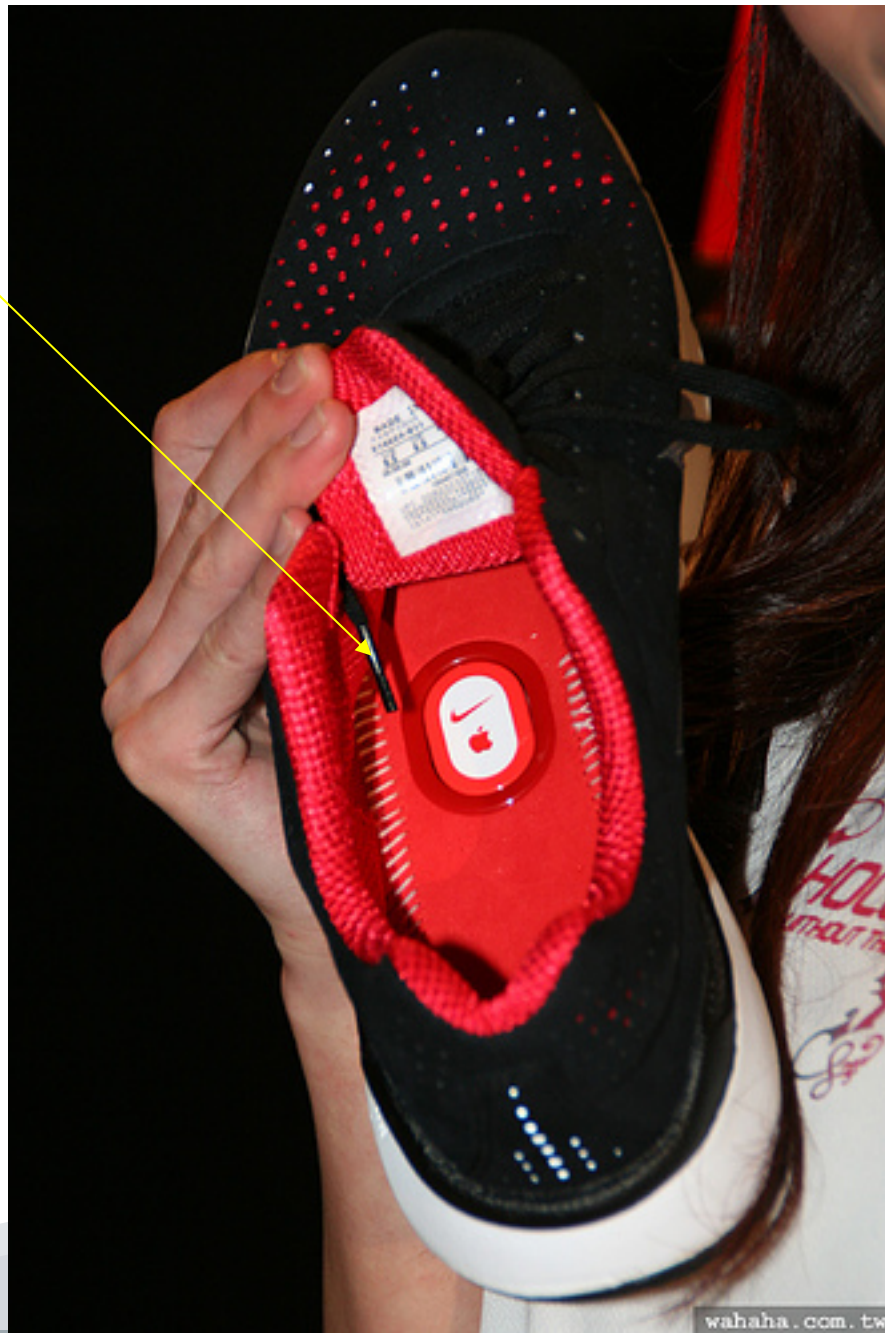
Respiration Rate

Body Temperature

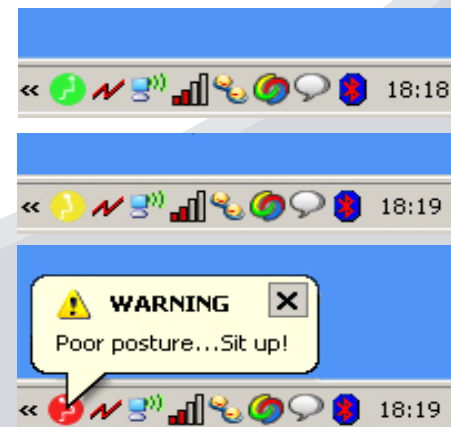
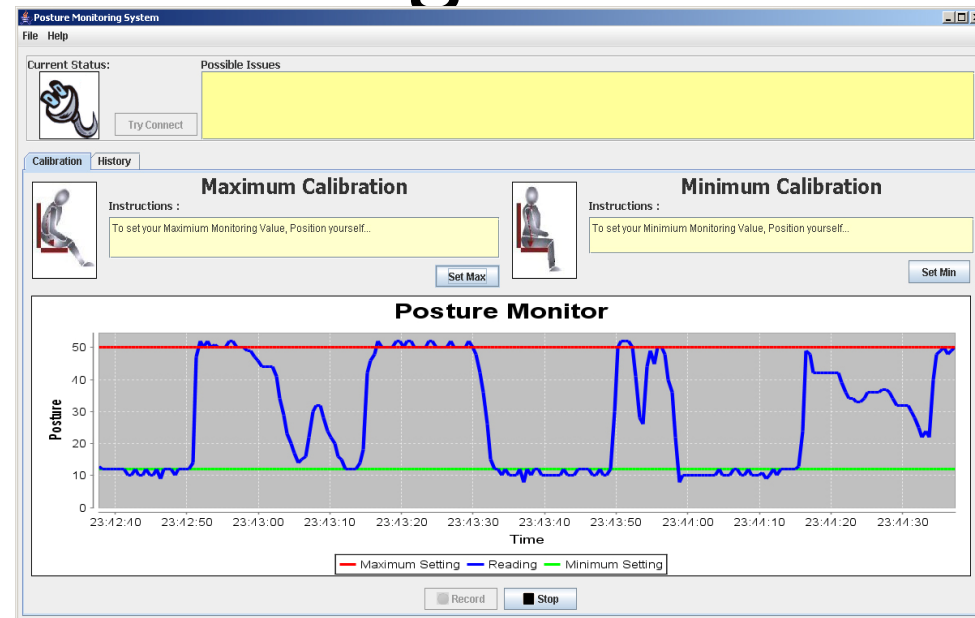
Heart Rate



Nike+



Data Sources: Posture Monitoring



Data Sources: “Web 2.0”

- MS/Google Health Vaults
- Flickr
- MySpace
- Facebook
- Twitter
- YouTube
- BedPosted
- etc.

SenseCam & Memory

- SenseCam may be a very powerful memory aid
- In autobiographical (long-term) memory
 - **“Cued Recall” better than “Free Recall”**
 - **Visual Encoding has strong effect on retrieval**
- Memory studies on-going
 - **Cambridge, U.K.**
 - **Leeds, U.K.**
 - **Toronto, Canada**
 - **Illinois, USA**
 - **etc.**

Who knows what's next...

Mr Lee, the lifelogging cat!



Thank You

further information:

<http://www.cdvp.dcu.ie/SenseCam>

<http://www.computing.dcu.ie/~adoherty>