

# Evaluating the Multi-Media Augmentation of LifeLog Events

**Aiden R. Doherty**

Supervisor: Prof. Alan F. Smeaton

Centre for Digital Video Processing (CDVP)

&

Adaptive Information Cluster (AIC),

Dublin City University (DCU)

# Overview

- **INTRODUCTION**
  - Introduction to CDVP, lifelogging, & the challenges
- **WORK COMPLETED**
  - Segmentation of Images into Events
  - Retrieval of Similar Events
  - Determining Important Events
  - Selecting Optimal Keyframe
  - System Screenshot
- **AUGMENTING EVENTS – OPTIMAL EVALUATION TECHNIQUE?**
  - Augmenting Lifelog images
- **CONCLUSIONS**

# Centre for Digital Video Processing

- Headed by Prof. Alan Smeaton
- 3 faculty members, 14 post-docs, 23 PhD students, 4 RAs, 3 support staff
- Focus on multimedia information retrieval
- Now investigating the area of lifelogging

# Lifelogging

- Lifelogging is about recording daily life, digitally
- Sometimes its for a reason,
  - work ... e.g. security personnel, medical staff,
  - 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 lifelogs, e.g. MyLifeBits

# Lifelogging Devices

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



# Lifelogging Devices

- Lin & Hauptmann, Carnegie Mellon, PA, USA



# SenseCam

- SenseCam is a Microsoft Research Prototype
- Multi-sensor device
  - Colour camera
  - 3 accelerometers
  - Light meter
  - Passive infrared sensor
- 1GB flash memory storage
- Smart image capture ~3 images/min
- Since April 2006 we've had two SenseCams ... recently have received 5 more



# How to Review Images?

- Make a 2 minute movie of your day!

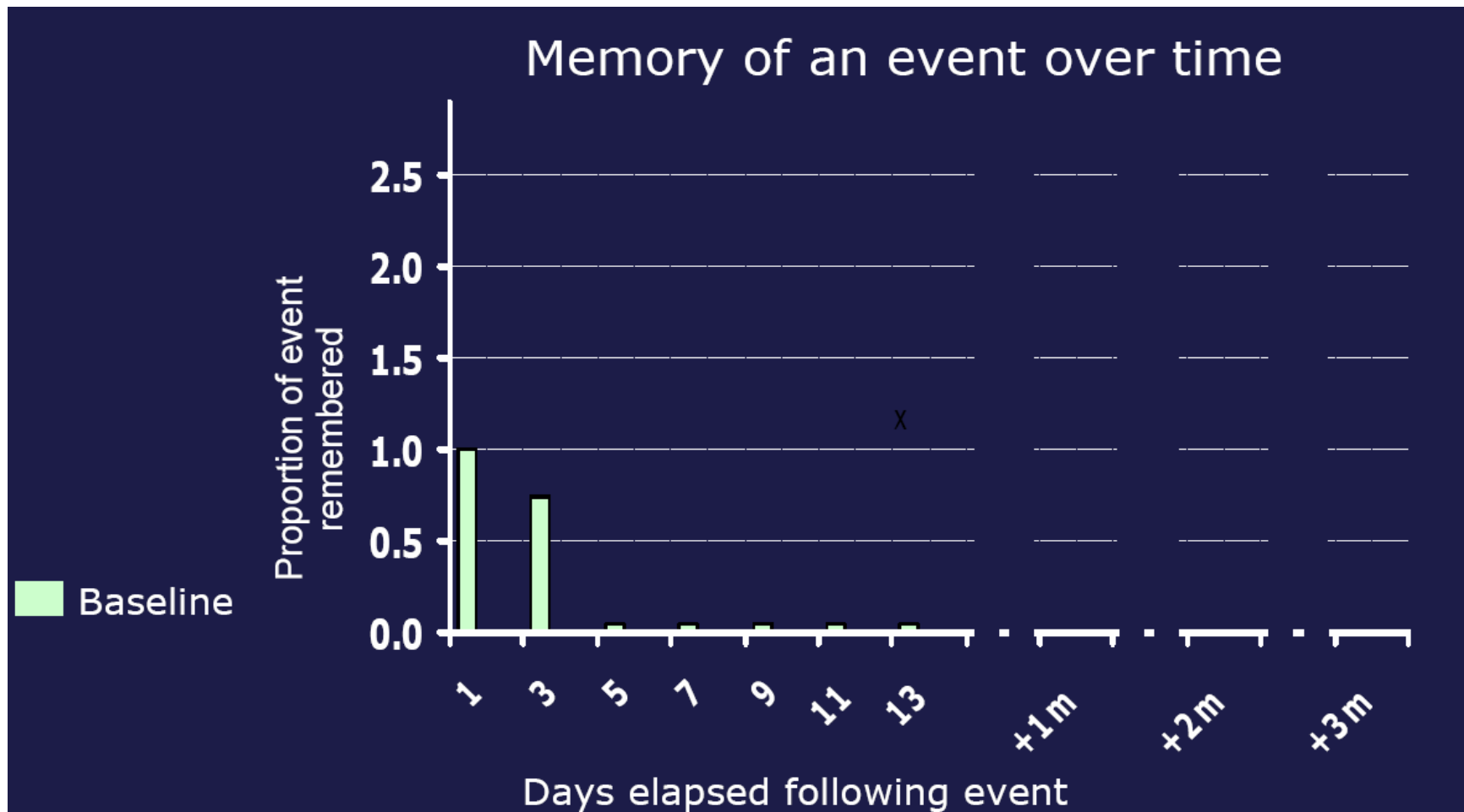




# Lifelogging Aiding Memory

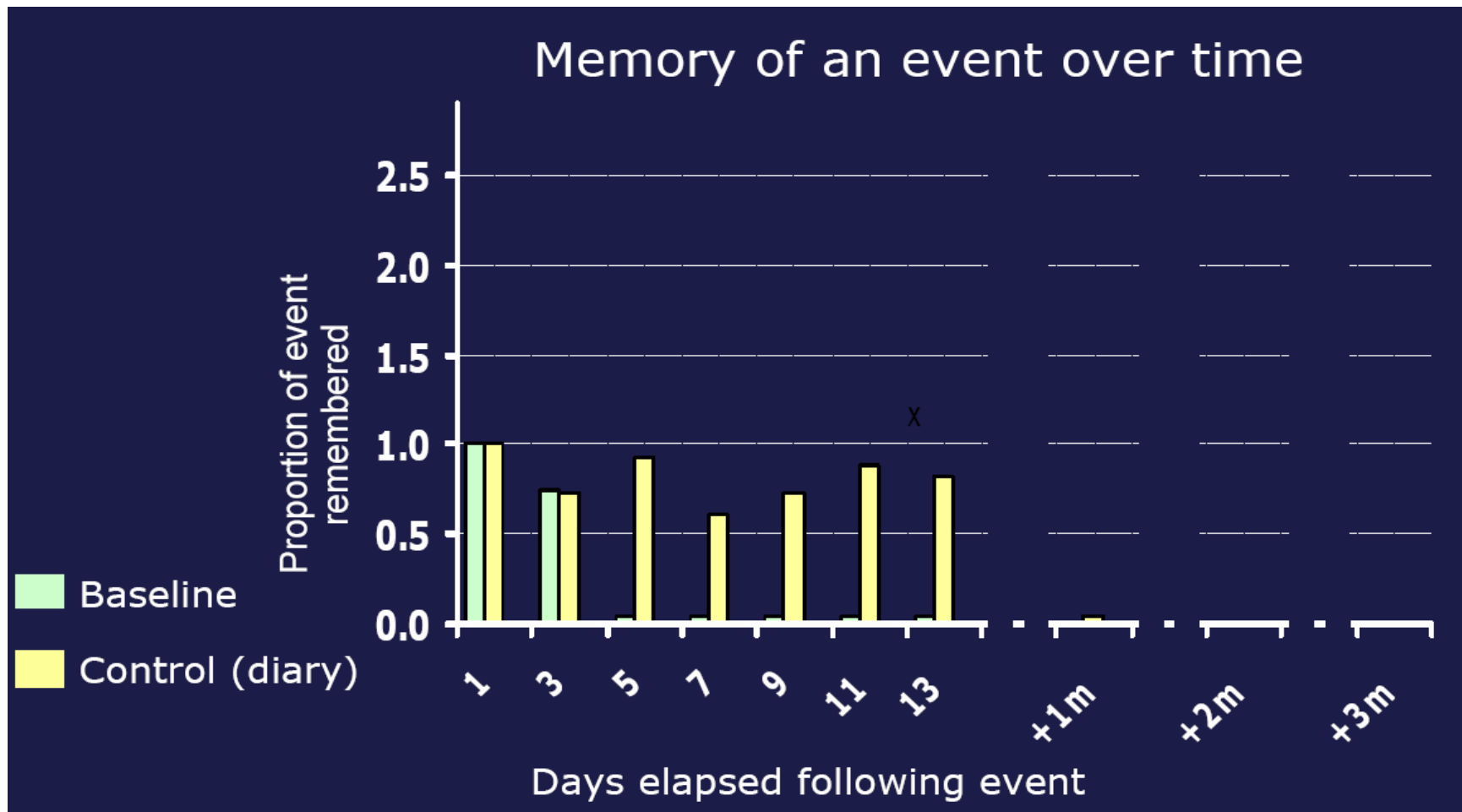
- Preliminary Study carried out by Cambridge Memory Clinic, Addenbrooke's Hospital
- 63 year old, well-educated married woman, with limbic encephalitis (usually has no memory a few days after an event)
- Each day her husband would ask her what she would remember from an event, and then talk her through it using SenseCam images afterwards
- A few days later, the same process would be repeated for that event

# SenseCam as a Memory Aid



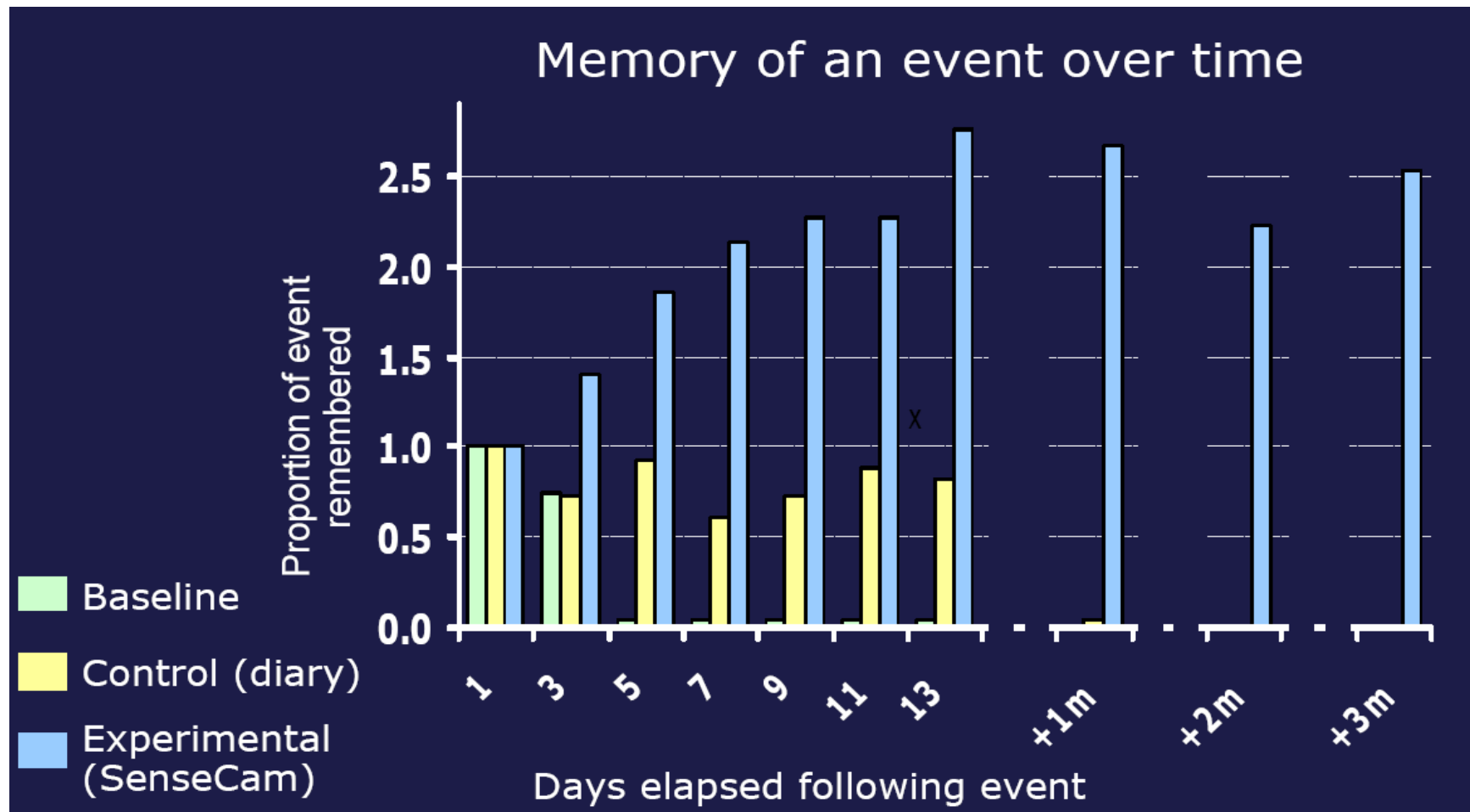
Microsoft Research Cambridge presentation: [http://research.microsoft.com/~shodges/presentations/UBICOMP\\_senseCam.pdf](http://research.microsoft.com/~shodges/presentations/UBICOMP_senseCam.pdf)

# SenseCam as a Memory Aid



Microsoft Research Cambridge presentation: [http://research.microsoft.com/~shodges/presentations/UBICOMP\\_senseCam.pdf](http://research.microsoft.com/~shodges/presentations/UBICOMP_senseCam.pdf)

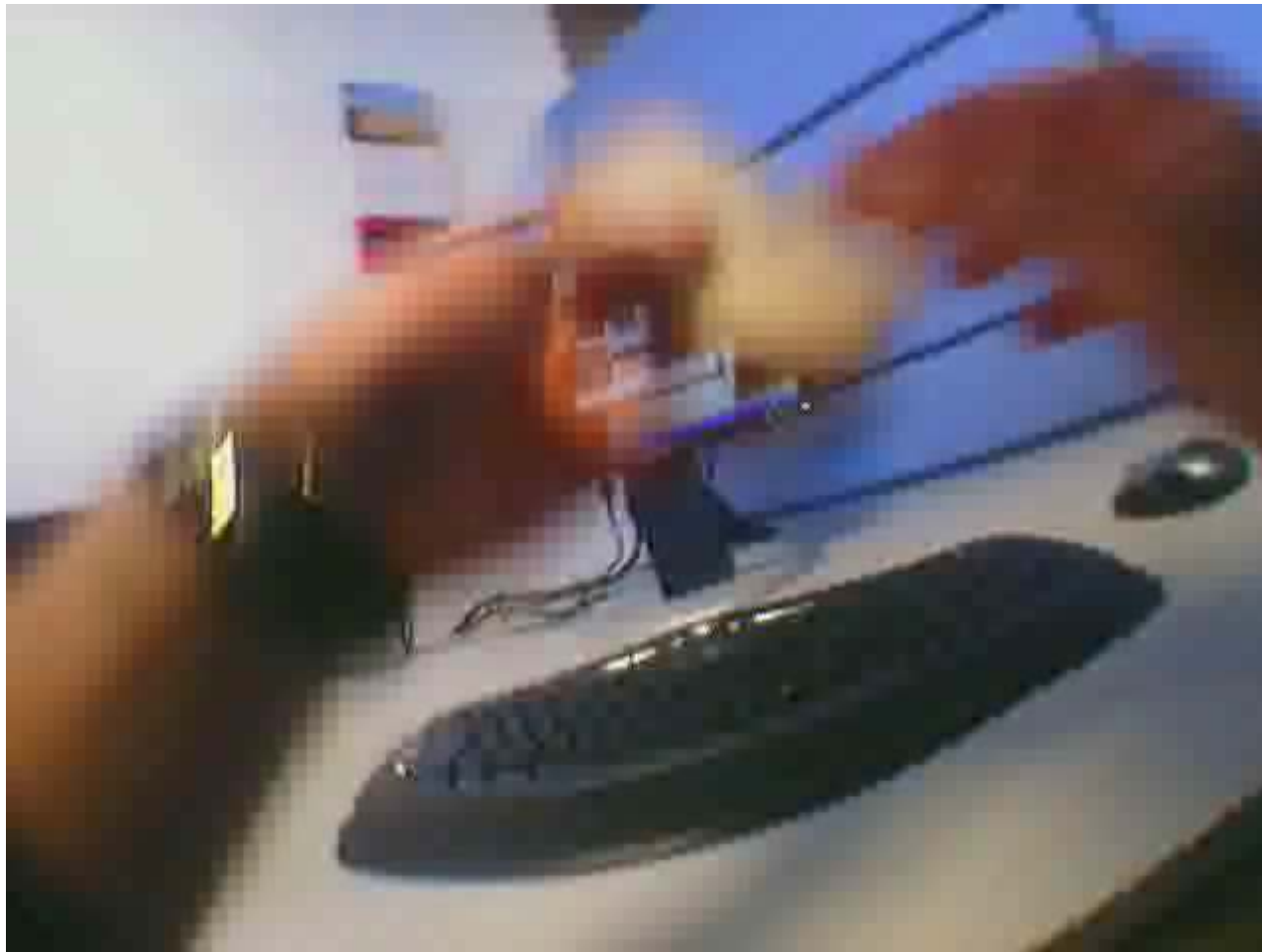
# SenseCam as a Memory Aid



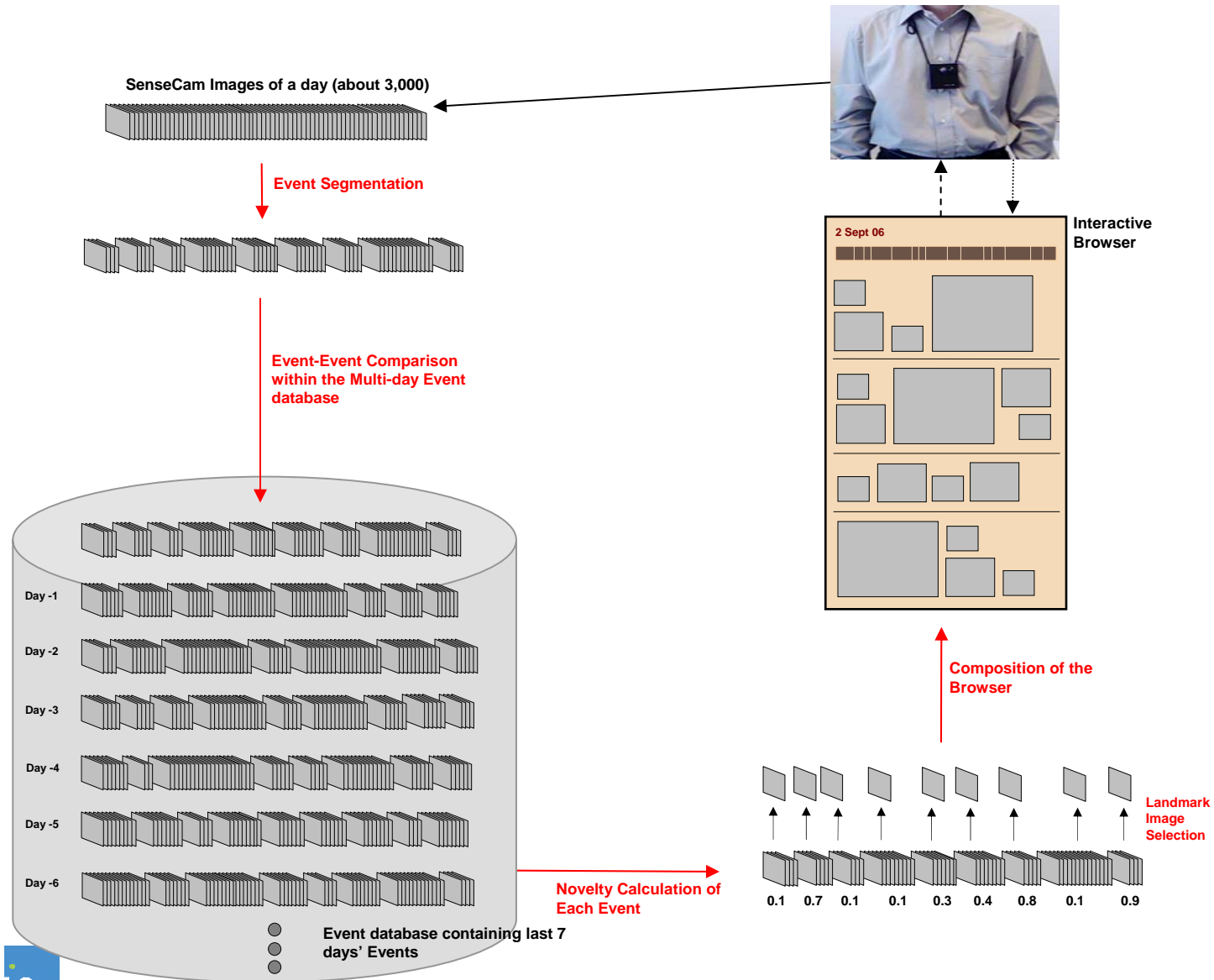
Microsoft Research Cambridge presentation: [http://research.microsoft.com/~shodges/presentations/UBICOMP\\_senseCam.pdf](http://research.microsoft.com/~shodges/presentations/UBICOMP_senseCam.pdf)

# Require Intelligent Summarisation

- Playing a movie of one's day takes too long to review



# Daily Browser Overview



# Overview

- INTRODUCTION
  - Introduction to CDVP, lifelogging, & the challenges
- **WORK COMPLETED**
  - **Segmentation of Images into Events**
  - Retrieval of Similar Events
  - Determining Important Events
  - Selecting Optimal Keyframe
  - System Screenshot
- AUGMENTING EVENTS – OPTIMAL EVALUATION TECHNIQUE?
  - Augmenting Lifelog images
- CONCLUSIONS

# Event Segmentation

Breakfast



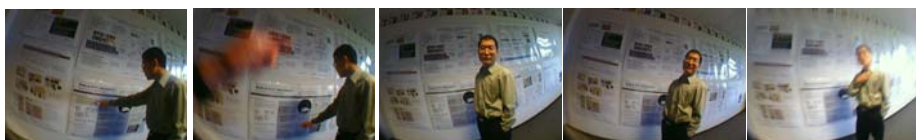
Work



Car



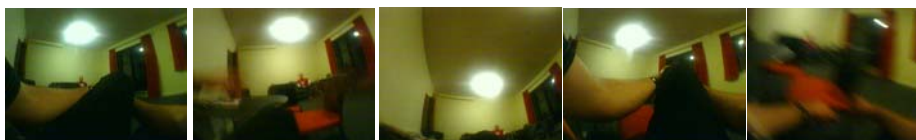
Talking to colleague



Airplane

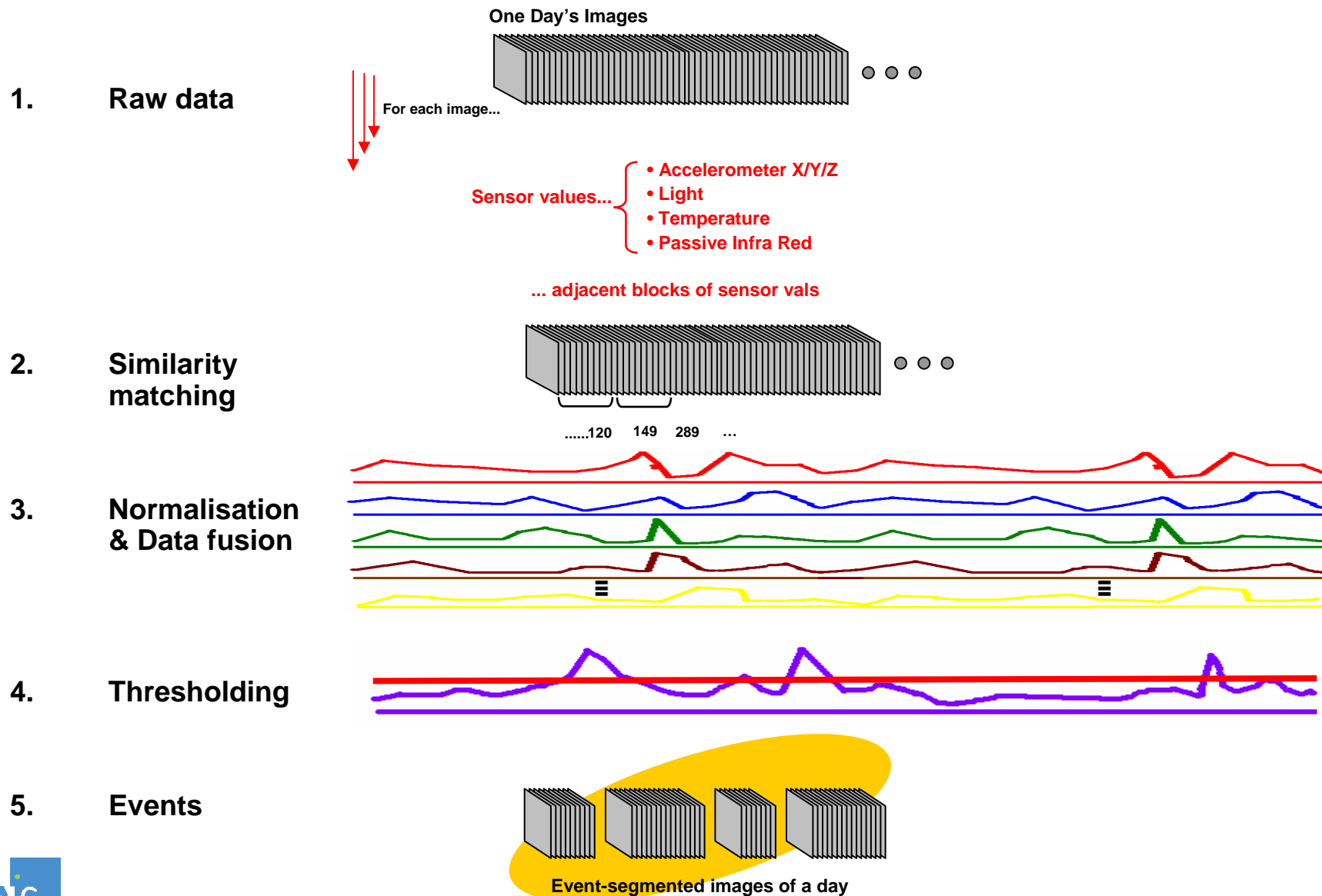


Talking to friend





# Event Segmentation



# Event Segmentation Expts.

- How well does it work?
- Recently published extensive experiments with 5 different users wearing SenseCam for 1 month each (total = 270k images)
- Each user groundtruthed their own data
- Data divided into training and test sets with thousands of different combinations evaluated

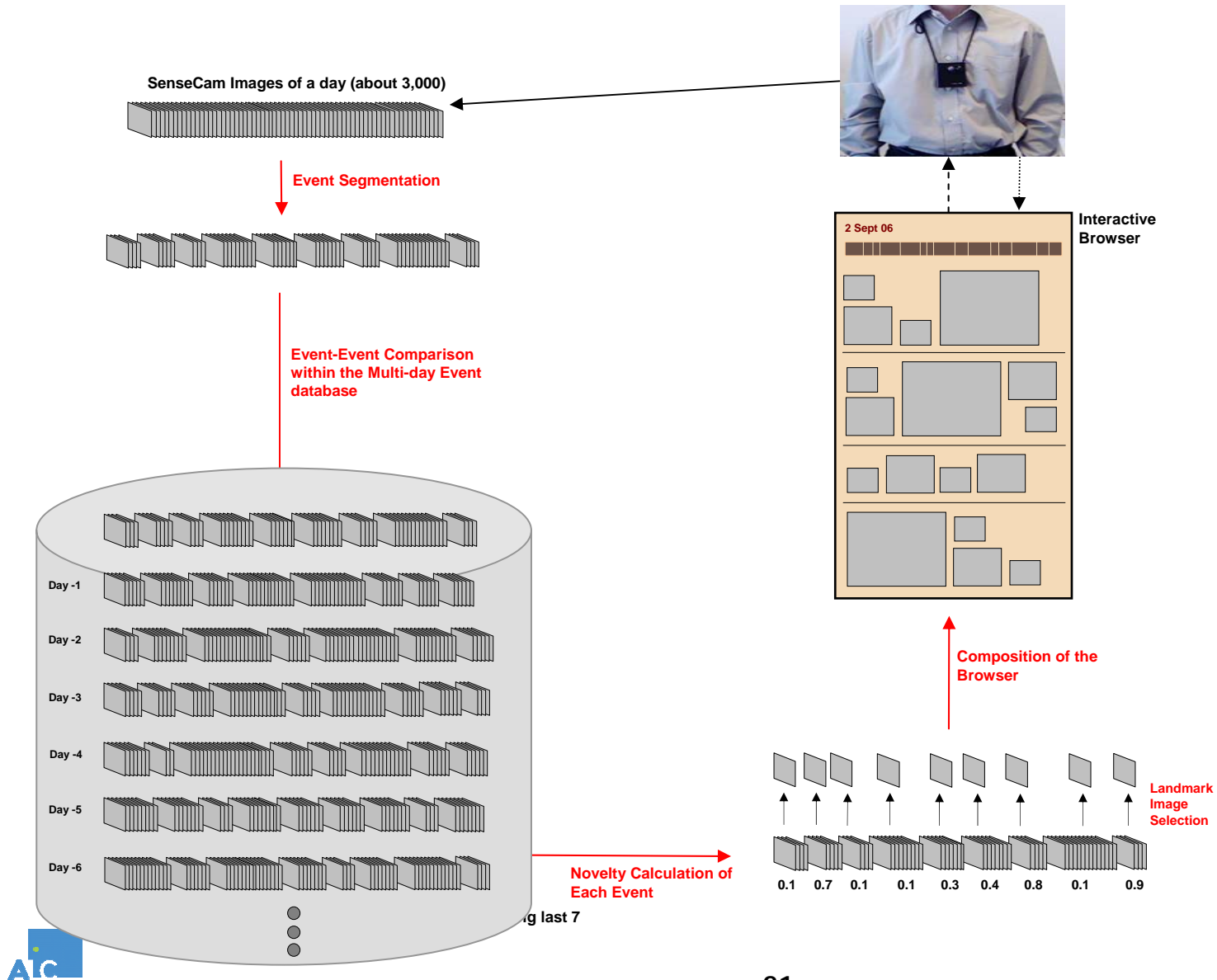
# Event Segmentation Expts.

- 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:
  - Previous publication (F1-Measure = 0.43)
  - Sensor only (F1-Measure = 0.60)
  - Image + Sensor (F1-Measure = 0.62)

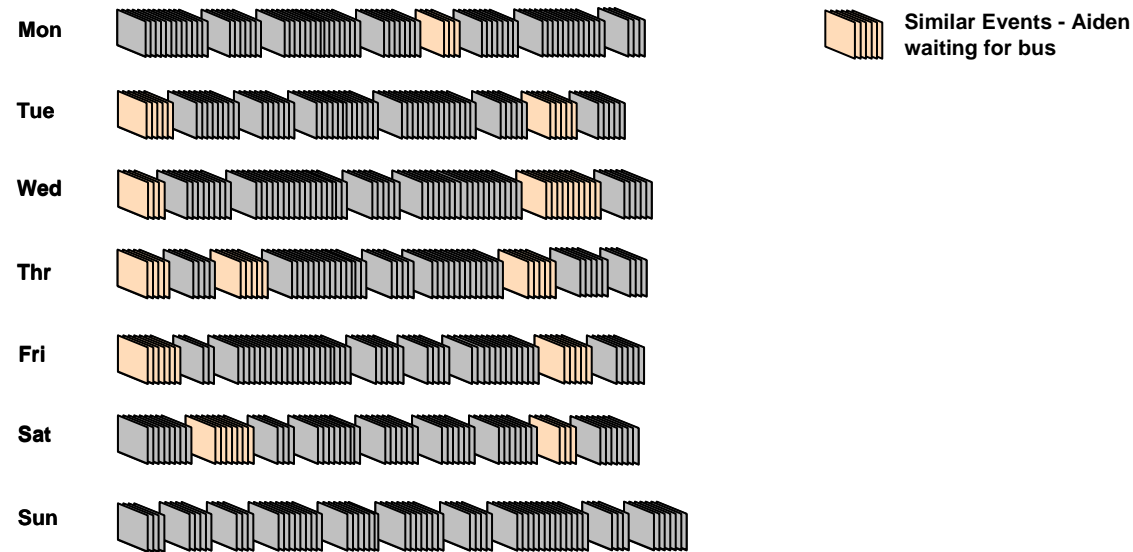
# Overview

- INTRODUCTION
  - Introduction to CDVP, lifelogging, & the challenges
- **WORK COMPLETED**
  - Segmentation of Images into Events
  - **Retrieval of Similar Events**
  - Determining Important Events
  - Selecting Optimal Keyframe
  - System Screenshot
- AUGMENTING EVENTS – OPTIMAL EVALUATION TECHNIQUE?
  - Augmenting Lifelog images
- CONCLUSIONS

# Retrieval Reminder



# Finding similar events



- Events are represented by the average values of all the images present in that event
- Investigated numerous vector distance, normalisation, & fusion metrics to match similarity of any two given events

# Event Retrieval Expts.

- How well does it work?
- Recently completed extensive experiments with 5 different users wearing SenseCam for 1 month each (total = 270k images) ... corresponds to 3,286 events
- 10 queries selected for each user e.g. driving, at work, eating, talking to friend, etc. (50 queries in total, 10x5)
- 13,399 pooled judgements made on relevance of events to query events
- Queries divided into training (60%) and test sets

# Event Retrieval Expts.

- Again thousands of combinations investigated in training phase
- Overall accuracy of top 5 returned documents is 63% ... (top 10 is 52%)
- Overall MAP score of 0.3608
- Query scores ranging from 0.0057 (Hyowon on public transport) to 0.9415 (Michael at work on his PC)



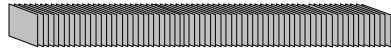
# Overview

- INTRODUCTION
  - Introduction to CDVP, lifelogging, & the challenges
- **WORK COMPLETED**
  - Segmentation of Images into Events
  - Retrieval of Similar Events
  - **Determining Important Events**
  - Selecting Optimal Keyframe
  - System Screenshot
- AUGMENTING EVENTS – OPTIMAL EVALUATION TECHNIQUE?
  - Augmenting Lifelog images
- CONCLUSIONS

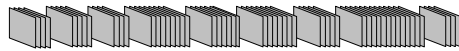
# Importance Reminder



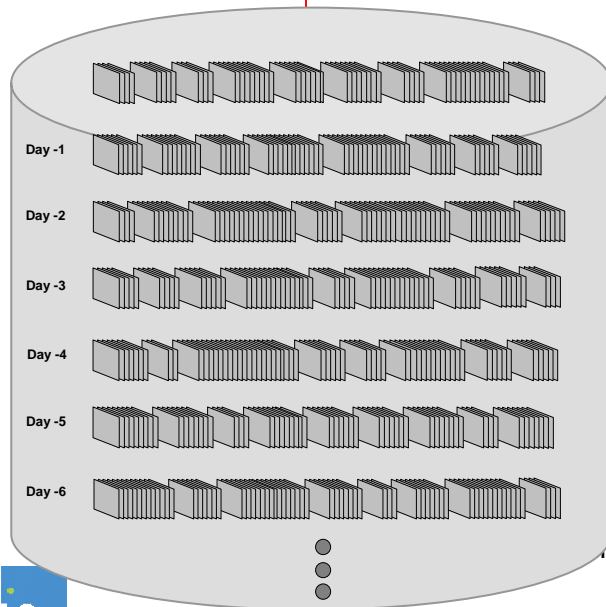
SenseCam Images of a day (about 3,000)



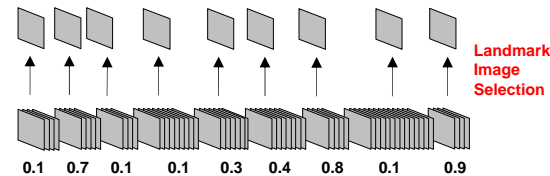
Event Segmentation



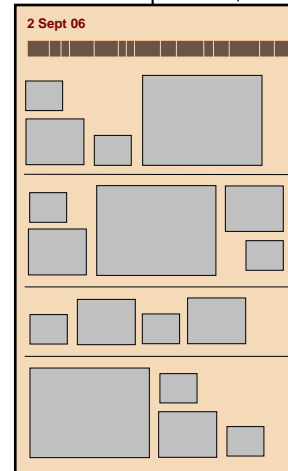
Event-Event Comparison within the Multi-day Event database



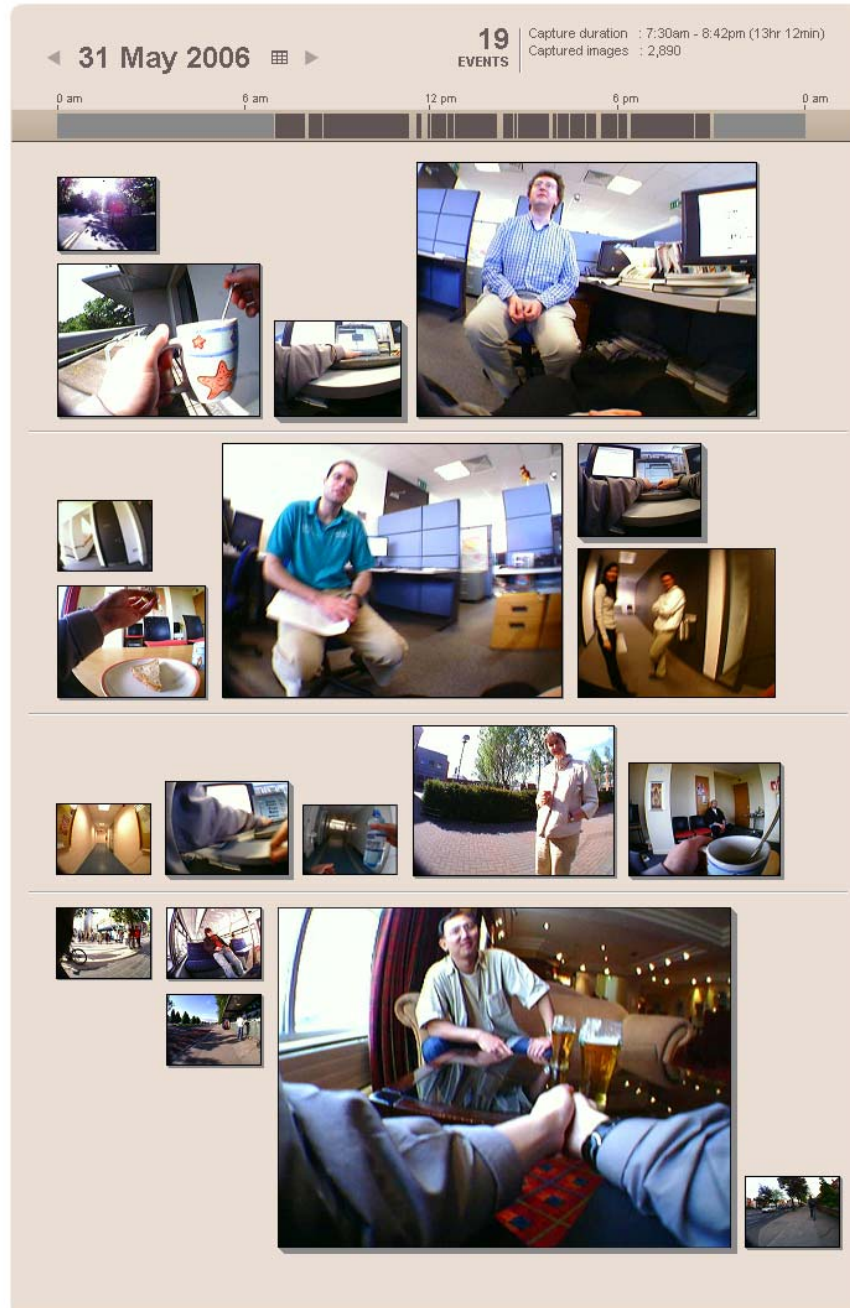
Novelty Calculation of Each Event



Composition of the Browser



Interactive Browser



# Importance

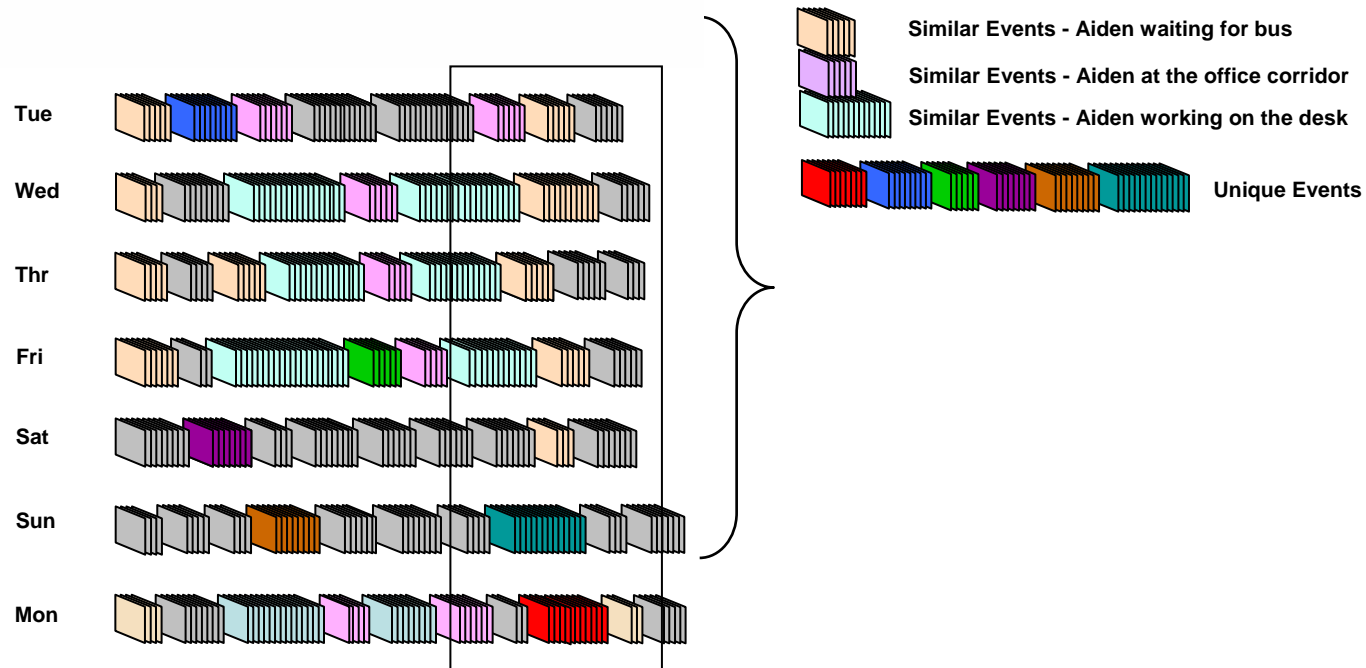
- Greater emphasis is placed on important events
- Routine/mundane events can be hidden

# Automatic Face Detection



- Trained on set of 1,758 SenseCam images
- SenseCam images are low quality
- Accuracy = 63%

# Novelty to Detect Event Importance



- Find the most dissimilar event of today by taking the previous 6 days into account.
- Also for any event, we only look at how novel it is with respect to events around the same time from other days

# Event Importance Expts.

- How well does it work?
- Recently completed extensive experiments with 3 different users wearing SenseCam for 4 weeks each (total = 176k images)
- 83 days of data collected in total, with 8 different approaches evaluated ... giving 664 judgements

< 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



# Event Importance Expts.

- 3 final approaches evaluated:
  - Face Detection Only (current state of art)
  - Novelty Only
  - Face Detection + Novelty
- Face Detection + Novelty performs at least as well as state of art 80% of the time, and 4% better overall
- Face Detection good at highlighting most important events
- Novelty good at detecting routine events



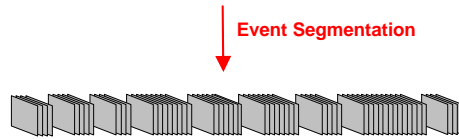
# Overview

- INTRODUCTION
  - Introduction to CDVP, lifelogging, & the challenges
- **WORK COMPLETED**
  - Segmentation of Images into Events
  - Retrieval of Similar Events
  - Determining Important Events
  - **Selecting Optimal Keyframe**
  - System Screenshot
- AUGMENTING EVENTS – OPTIMAL EVALUATION TECHNIQUE?
  - Augmenting Lifelog images
- CONCLUSIONS

# Keyframe Reminder

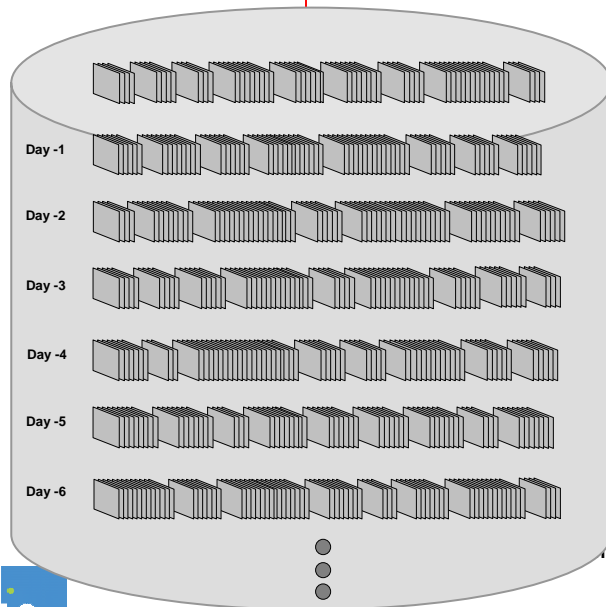


SenseCam Images of a day (about 3,000)



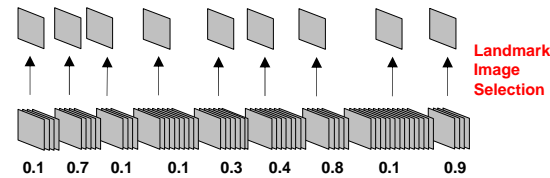
Event Segmentation

Event-Event Comparison within the Multi-day Event database

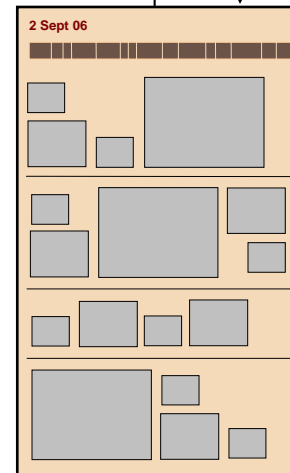


Novelty Calculation of Each Event

g last 7



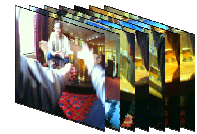
Composition of the Browser



Interactive Browser

# Keyframe Selection Techniques

- Standard Approaches
  - Middle Image
  - Within Event Image
  - Cross Event Image
- Proposed Approaches
  - Image Quality (based on colour contrast + saliency measure)
  - Quality + Within Event Image
  - Quality + Cross Event Image



# Keyframe Experiments

- How well does it work?
- Recently completed experiments with 5 different users wearing SenseCam for up to 4 weeks each (total = 194k images)
- User judgements made on 2,235 events; 6 keyframes judgements per event ... providing a groundtruth of 13,410 judgements

# Keyframe Experiments

- Image Quality + Within Event selection works best overall (Likert avg = 3.99/5)  
... middle image = 3.68/5
- Then only considering “visual high-variance” events ... Image Quality alone works best (3.92 vs 3.31 for middle image)



# Overview

- INTRODUCTION
  - Introduction to CDVP, lifelogging, & the challenges
- **WORK COMPLETED**
  - Segmentation of Images into Events
  - Retrieval of Similar Events
  - Determining Important Events
  - Selecting Optimal Keyframe
  - **System Screenshot**
- AUGMENTING EVENTS – OPTIMAL EVALUATION TECHNIQUE?
  - Augmenting Lifelog images
- CONCLUSIONS

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



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

ADD TO FAVE | FIND SIMILAR



MY ACCOUNT | SIGN OUT | ABOUT

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 ▶



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



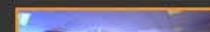
14:39 (Duration: 15m 30s)  
12 APR 2006 ▶



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



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



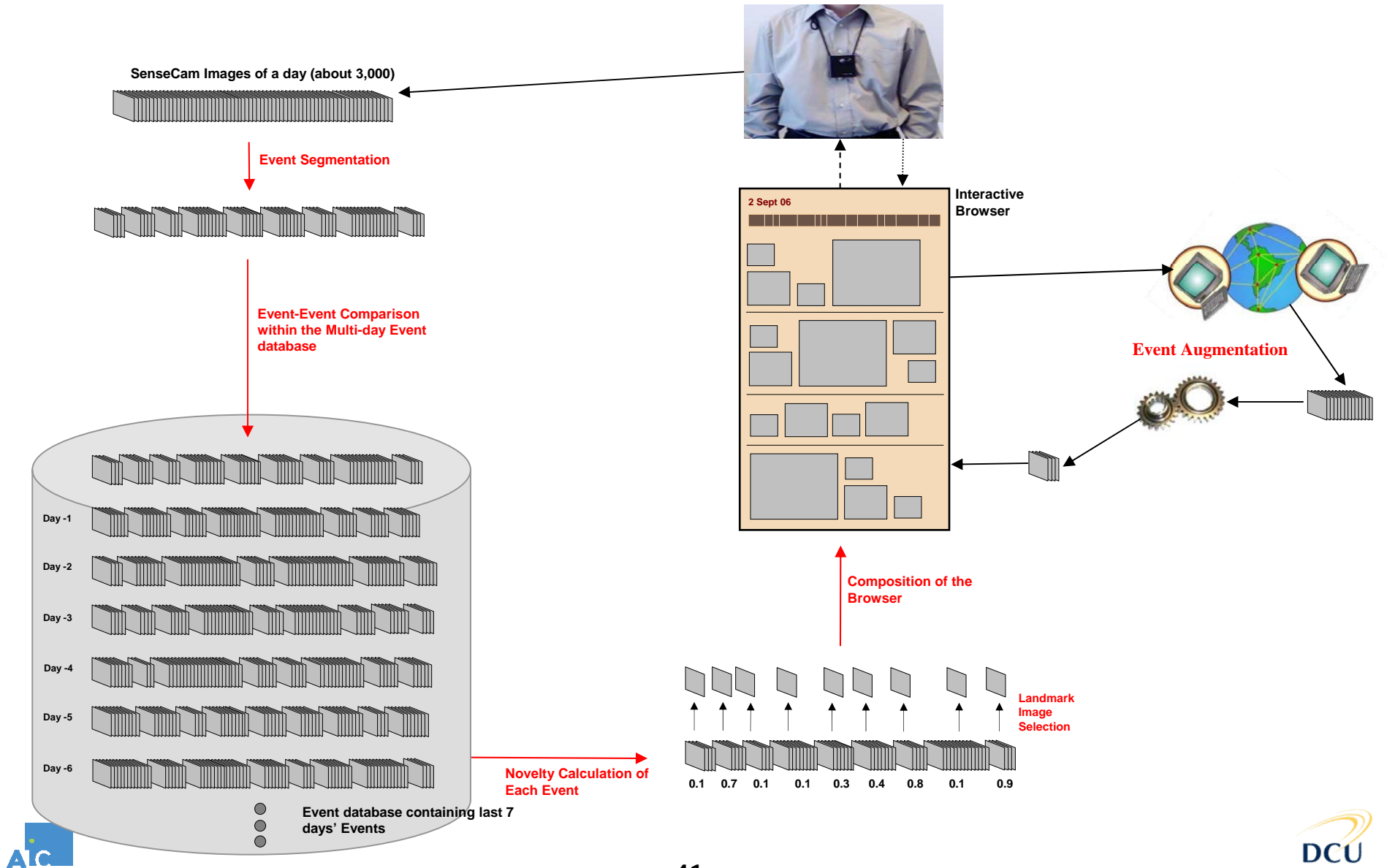
15:19 (Duration: 21m 10s)

# Overview

- INTRODUCTION
  - Introduction to CDVP, lifelogging, & the challenges
- WORK COMPLETED
  - Segmentation of Images into Events
  - Retrieval of Similar Events
  - Determining Important Events
  - Selecting Optimal Keyframe
  - System Screenshot
- **AUGMENTING EVENTS – OPTIMAL EVALUATION TECHNIQUE?**
  - **Augmenting Lifelog images**
- CONCLUSIONS



# Daily Browser Overview



# Event Augmentation – Croke Park

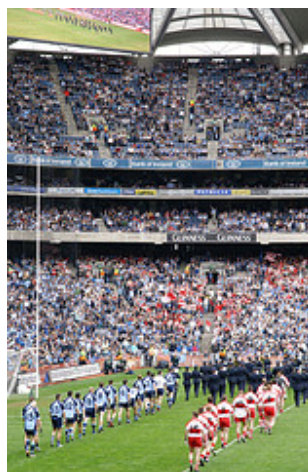
Here's an image from a SenseCam after a big match in Croke Park, Dublin. We'd really like to see other people's pictures of this match.

Let's search by location...



# Event augmentation – Croke Park

- Receive the following pictures...
- Then filter out to just those results from around the same time as the event



# Event augmentation – Santa Barbara

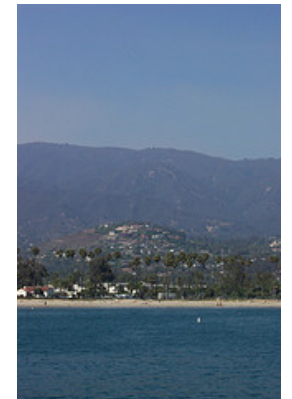
Here's a SenseCam picture of a building that I like from the pier in Santa Barbara, CA.

Again I search for other pictures in the same location...



# Event augmentation – Santa Barbara

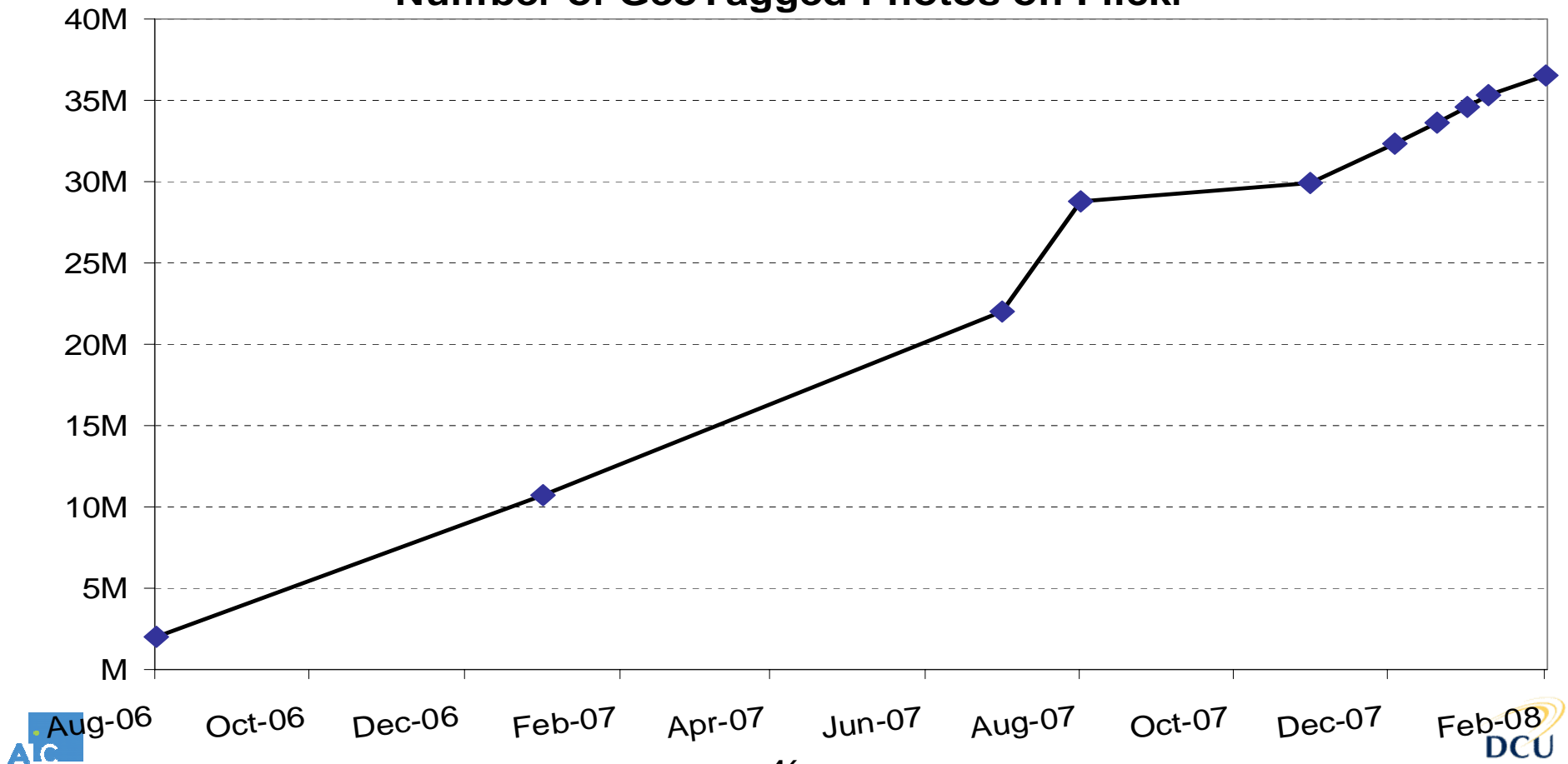
- I receive the following pictures...
- Then I filter out to just those results that are visually similar



# Event Augmentation

- Augment low-quality SenseCam images with high quality images from external sources

Number of GeoTagged Photos on Flickr



# Flickr Statistics

- Content generated by over 8 million users
- 38.3 million geo-tagged images
- 250,000 geo-tagged images uploaded per week
- Panoramio.com have 6 million geo tagged Google Earth photos (as of Nov '07)

# Types of Query

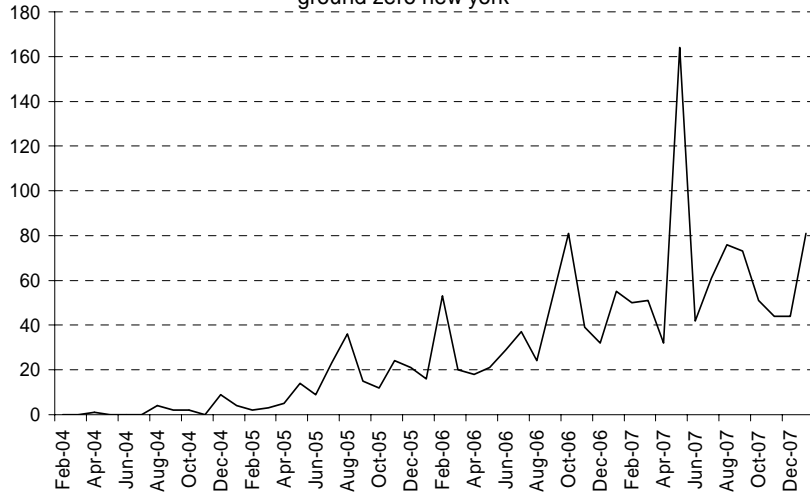
- Specific Place of Interest
  - Niagra Falls, Canada
  - Ground Zero, New York
  - Opera Houses, Sydney
- Specific Event
  - World ice hockey final, Moscow
  - U2 concert in Croke Park, Dublin
  - Monaco Grand Prix



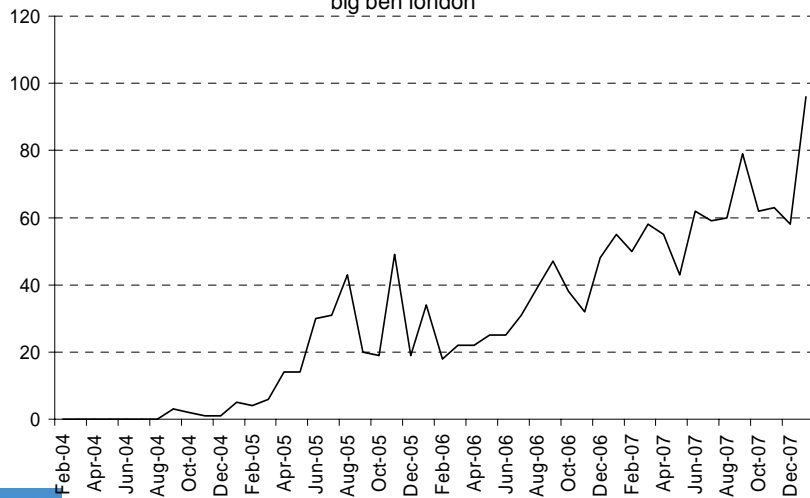
# Photo Upload Temporal Aspects

## PLACES

ground zero new york

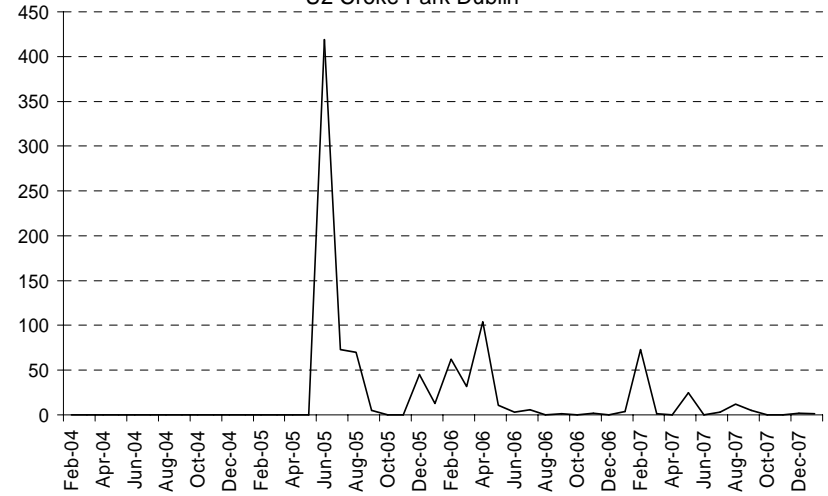


big ben london

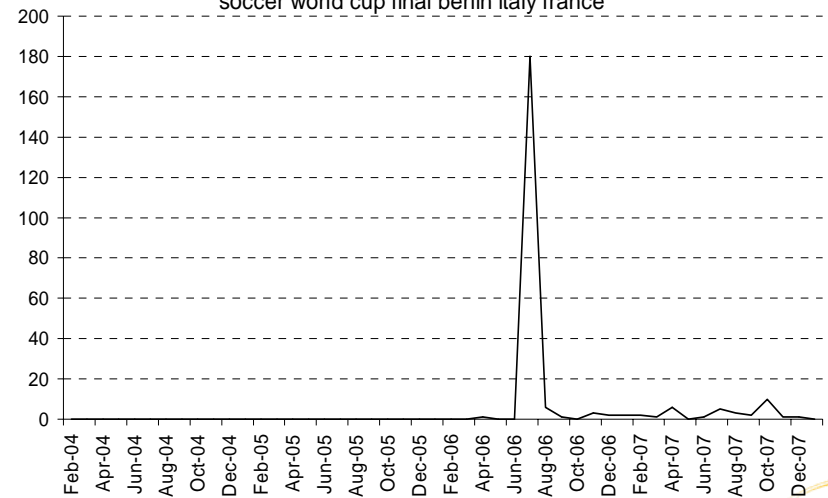


## EVENTS

U2 Croke Park Dublin



soccer world cup final berlin italy france



# Possible Search Techniques

- Search by location only
- Search by location + time (+- 5 days)
- Search by location + MPEG-7
- Search by location + time + MPEG-7

	Place P@10	Event P@10
Location only	0.4	0.2
Location + Time	0.0	1.0
Location + MPEG7	0.5	0.3
Location + Time + MPEG-7	0.0	0.8

Initial impressions on 2 query events...

# Not all images are geo-tagged!!!

- Carried out 26 textual queries on Flickr website - > 26,171 images
- 14 place specific queries; 22,486 images in total; 491 users uploading per query
- 12 event specific queries; 3,685 images in total; 32 users uploading per query
- Overall **just 22% of images are geo-tagged** (23% for places and only 14% for events)

# Tag Search Techniques

- Search by user generated tag (oracle tag)
- Search by location firstly, automatically determine common tags, then tag search (location -> tag)
- Constrain by time (location+time -> tag+time)

	Place P@10	Event P@10
Oracle tag	1.0	1.0
Location -> Tag	0.4	0.0
Location+Time -> Tag+Time	0.0	1.0

# Problem in selecting good tags

Tag	#	Tag	#
newyork	16	statue	3
unitedstates	12	thebigapple	2
nyc	11	thesphere	2
newyorkcity	9	us	2
manhattan	9	warmemorial	2
batterypark	8	worldtradecenter	2
usa	7	worldwarii	2
unitedstatesofamerica	6	2007	2
statueofliberty	5	911	2
memorial	5	eastcoastmemorial	2
park	4	geotagged	2
ny	3	gothamist	2
downtown	3	island	2
eagle	3	sculpture	2
america	3	skyscraper	2
architecture	3		

- 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?

# Tagging opens many new possibilities...

- Not only have we access to images from geotagged image sites now...
- YouTube videos (70M+ of them) are now opened up as a source of augmentation...

	Place P@10	Event P@10
Oracle tag	0.5	1.0
Location -> Tag	0.2	0.0
Location+Time -> Tag+Time	0.0	0.4

Lots of room for improvement!!!

# Future Technical Augmentation Work

- Gather more diverse queries from a range of users
- Better initial filtering of results based on event location, time, and keyframe image
- Better selection of tags
  - using WordNet to expand list of county/city names to exclude e.g. “united states”, “usa”
  - Yahoo API to expand words e.g. “statueofliberty” -> “statue of liberty”

# Evaluation Questions

- Investigating
  - Optimal initial query formation
  - Optimal selection of tags (to allow blind/pseudo relevance feedback)
  - Automatic relevance feedback vs. semi-automated (i.e. suggest query terms to the user before searching)



# Evaluation Questions

- Metrics to evaluate
  - Precision @n (n should be low to decrease burden on assessors (the owner of images) )
  - Recall not practical
  - Processing time – user wants instant feedback
  - User experience for automated queries vs. semi-automated

# Overview

- **INTRODUCTION**
  - Introduction to CDVP, lifelogging, & the challenges
- **WORK COMPLETED**
  - Segmentation of Images into Events
  - Retrieval of Similar Events
  - Determining Important Events
  - Selecting Optimal Keyframe
  - System Screenshot
- **AUGMENTING EVENTS – OPTIMAL EVALUATION TECHNIQUE?**
  - Augmenting Lifelog images
- **CONCLUSIONS**

# Conclusions

- Introduction to the concept of Lifelogging
- Past research in the field was predominantly hardware+storage based
- Extensive work complete in segmenting images into distinct events;
- Retrieval of similar events and highlighting important events (as well as keyframes) done

# Conclusions

- User generated content is expanding at *massive* scale
- An individual can be empowered by augmenting their collection with user-generated content by others
- Challenge of correctly evaluating techniques of (semi-) automatic event augmentation



# Kiitos paljon!

further information:

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

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