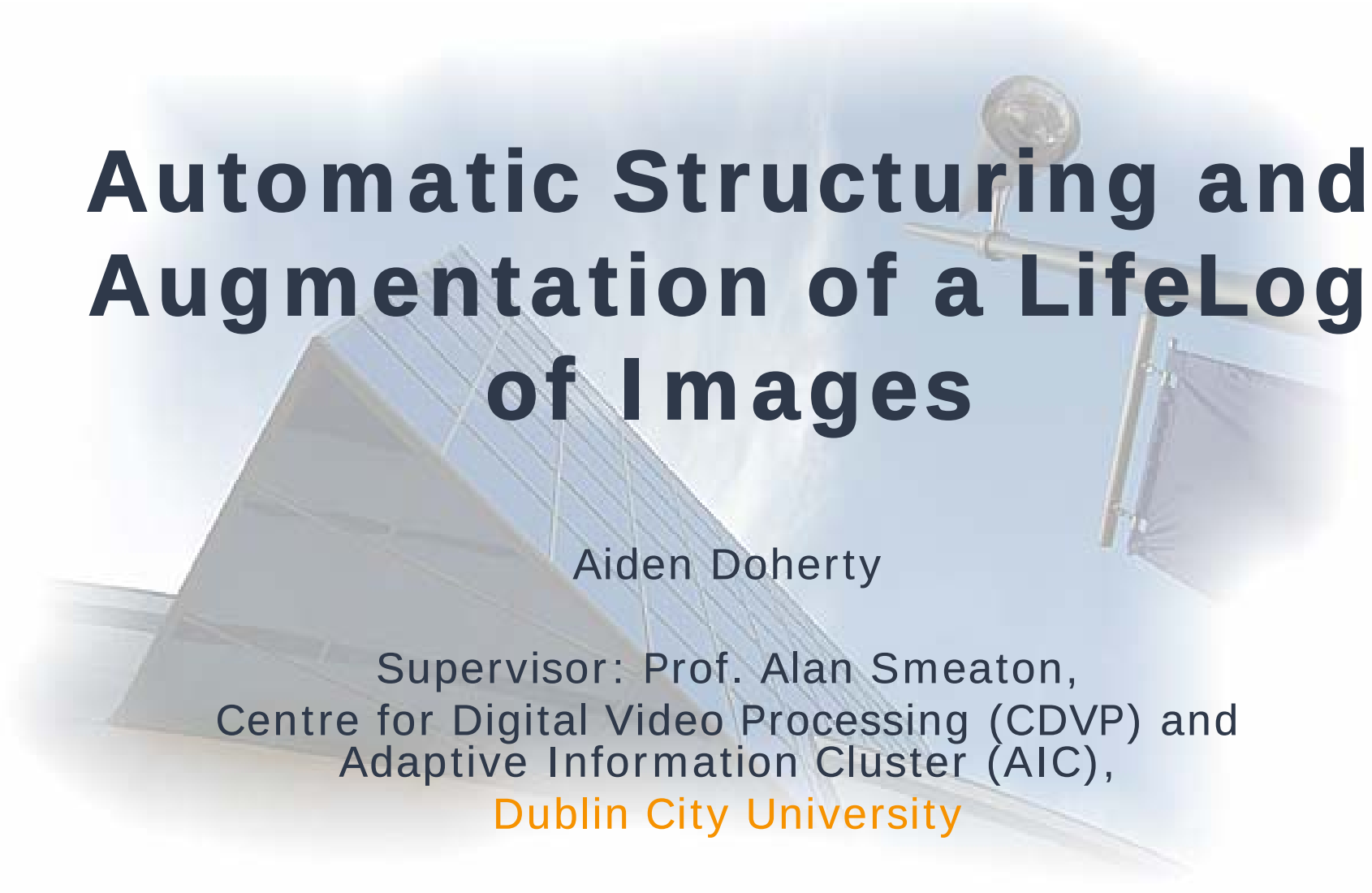


Automatic Structuring and Augmentation of a LifeLog of Images

Aiden Doherty

Supervisor: Prof. Alan Smeaton,
Centre for Digital Video Processing (CDVP) and
Adaptive Information Cluster (AIC),
Dublin City University

Outline

- **1 Introduction**
  - **1.1 Centre for Digital Video Processing**
  - **1.2 Microsoft SenseCam**
  - **1.3 Challenges**
- 2 Work completed
  - 2.1 Event Segmentation
  - 2.2 Finding Similar Events
  - 2.3 Event Importance
- 3 System Demo
- 4 Planned Work
  - 4.1 Event Augmentation

1.1 Centre for Digital Video Processing

- Headed by Prof. Alan F. Smeaton
- 45 full-time researchers
- Focus on multimedia information retrieval
- Starting to look into area of lifelogging

1.2 Lifelogging device of our group

Microsoft SenseCam

- Captures approx. 3,000 pictures/day
- Captures sensor data (light, movement, temperature, passive infra red)

1.3 How to review all these images?

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

1.3 What we would like to do – event segmentation

A day's SenseCam images (2,000 – 3,000)

Event Segmentation

Multiple Events

Finishing work in the lab

At the bus stop

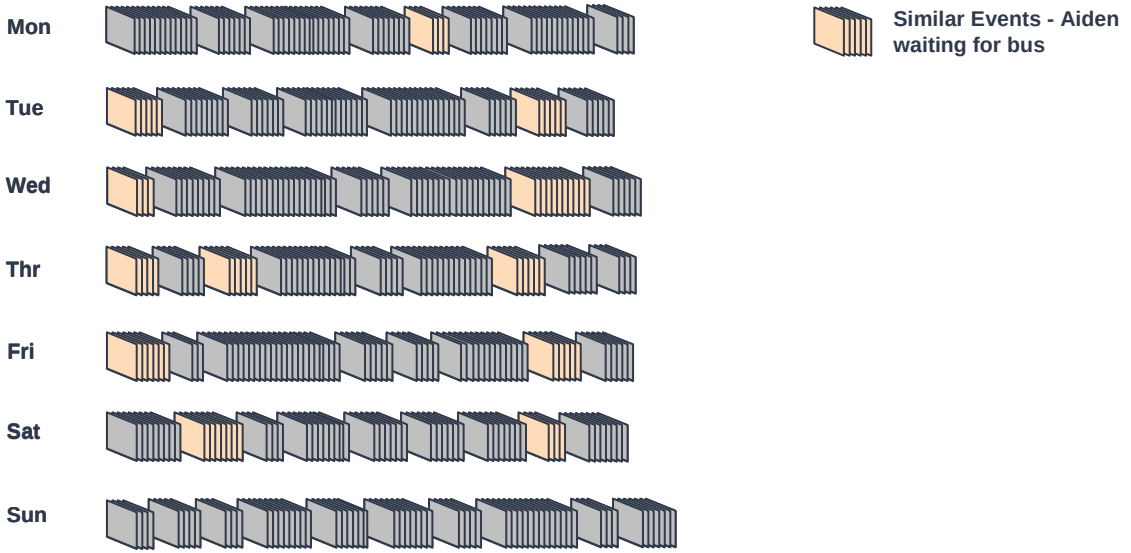
Chatting at Skylon Hotel lobby

Moving to a room

Tea time

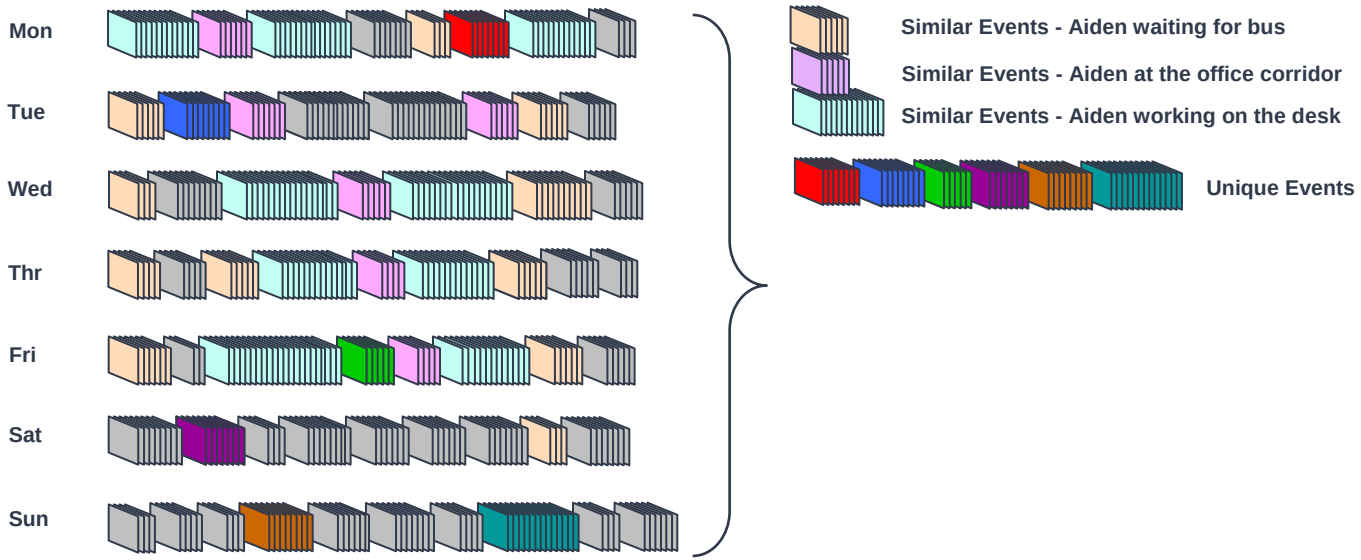
On the way back home

1.3 What we would like to do – finding similar events



- When was the last time I was doing something similar to this event of talking to my friend?

1.3 What we would like to do – most important events



- Talking to the president of the USA would be more memorable than breakfast last Tuesday morning.
- We argue that activities that only occur sporadically (talking to George Bush) are more important than those that occur frequently (having breakfast), i.e. they're more unique – novelty detection

Outline

- 1 Introduction
  - 1.1 Centre for Digital Video Processing
  - 1.2 Microsoft SenseCam
  - 1.3 Challenges
- **2 Work completed**
  - **2.1 Event Segmentation**
  - **2.2 Finding Similar Events**
  - **2.3 Event Importance**
- 3 System Demo
- 4 Planned Work
  - 4.1 Event Augmentation

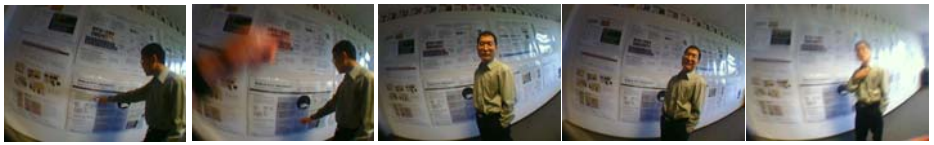
2.1 Event segmentation – sample activities

Breakfast

Work

Car

Talking to colleague



Airplane

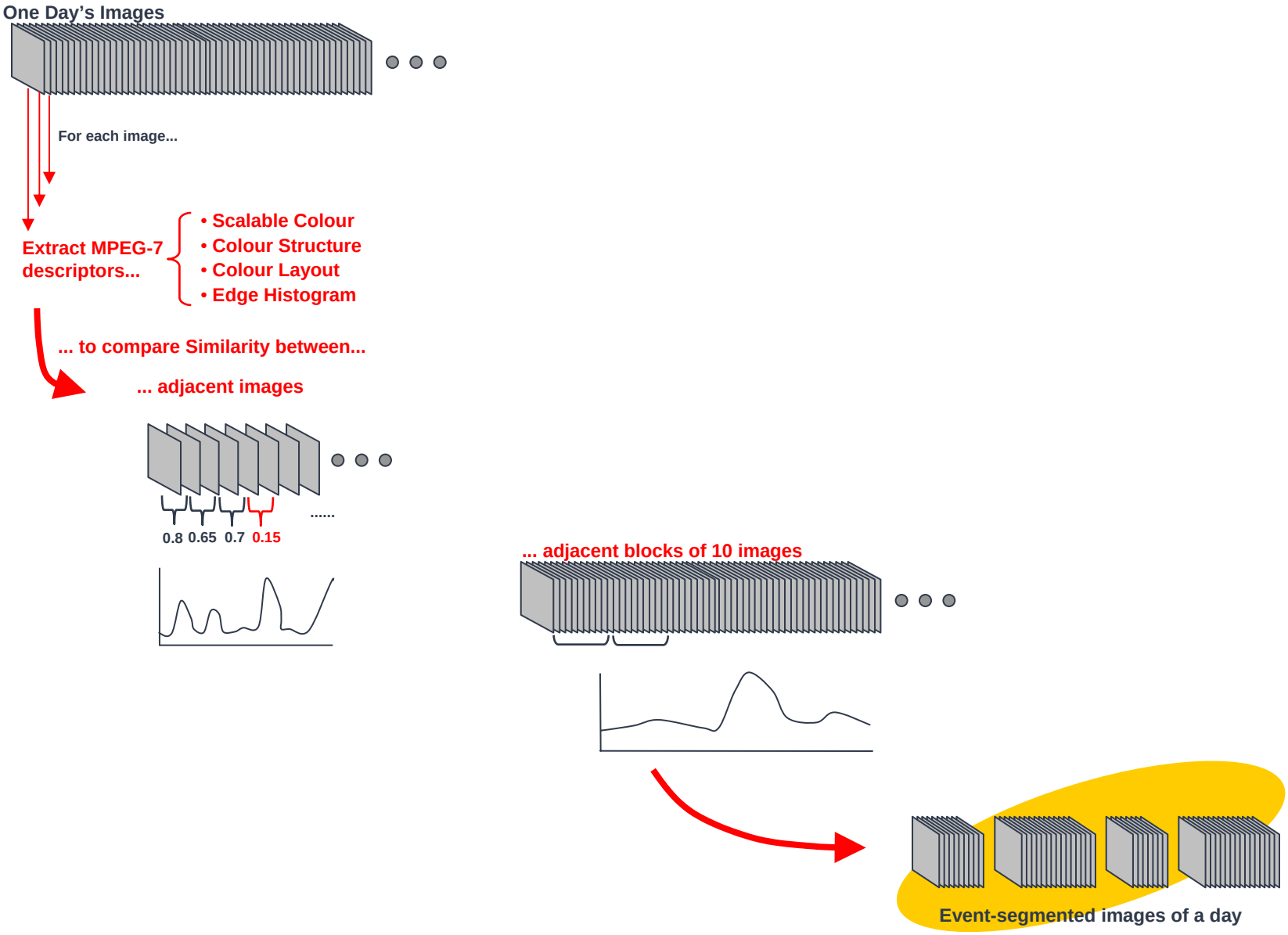
Talking to friend

2.1 Event segmentation

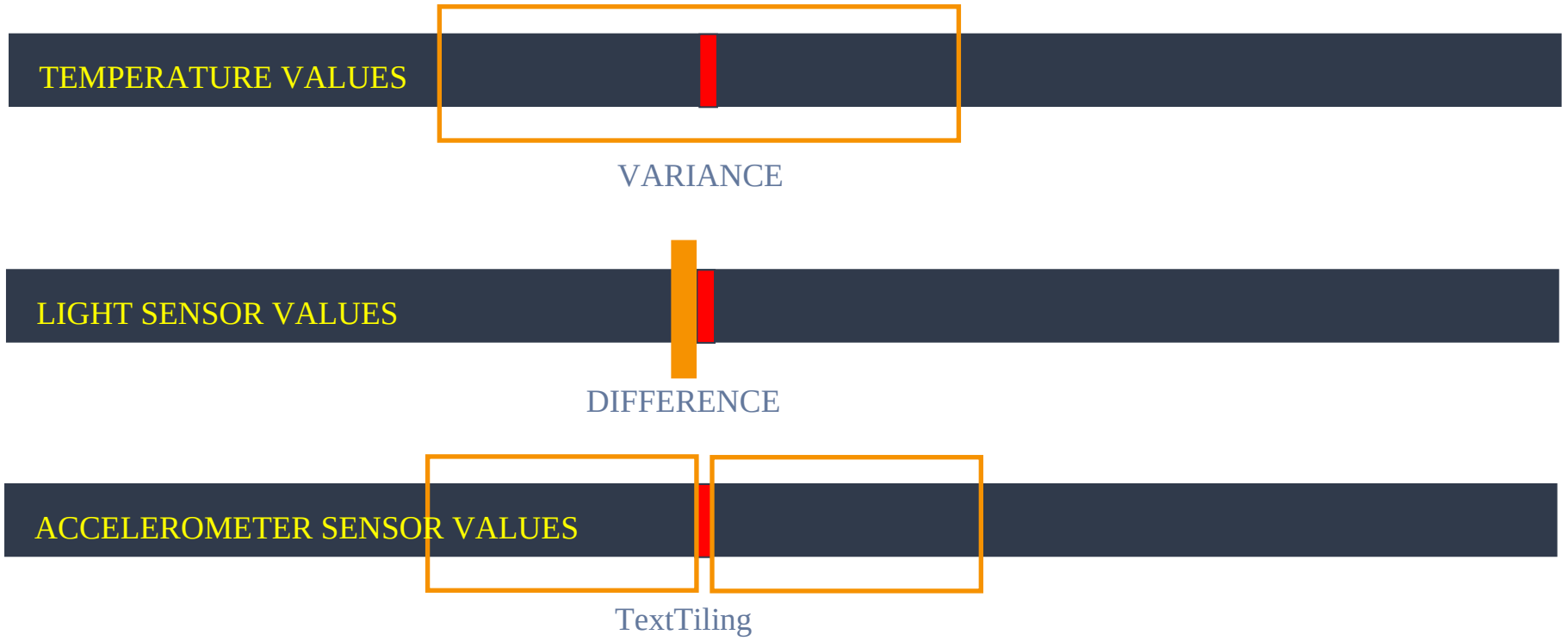
We segment images based on:

- Low-level image features
- Temperature sensor
- Light level sensor
- Accelerometer sensor

2.1 Image processing



2.1 Sensor processing



- All sources normalised and fused

2.1 Event segmentation paper

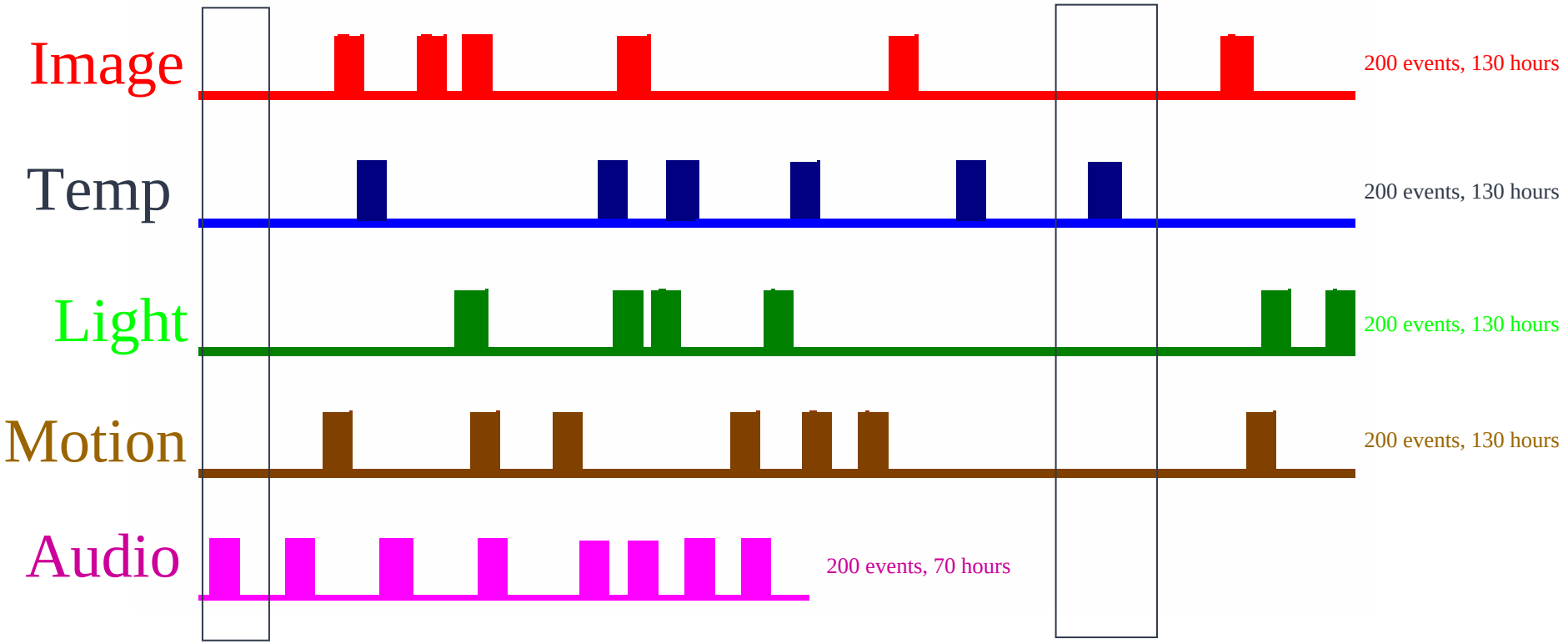
In this paper we segment images based on:

- Low-level image features
- Temperature sensor
- Light level sensor
- Accelerometer sensor

- Audio features

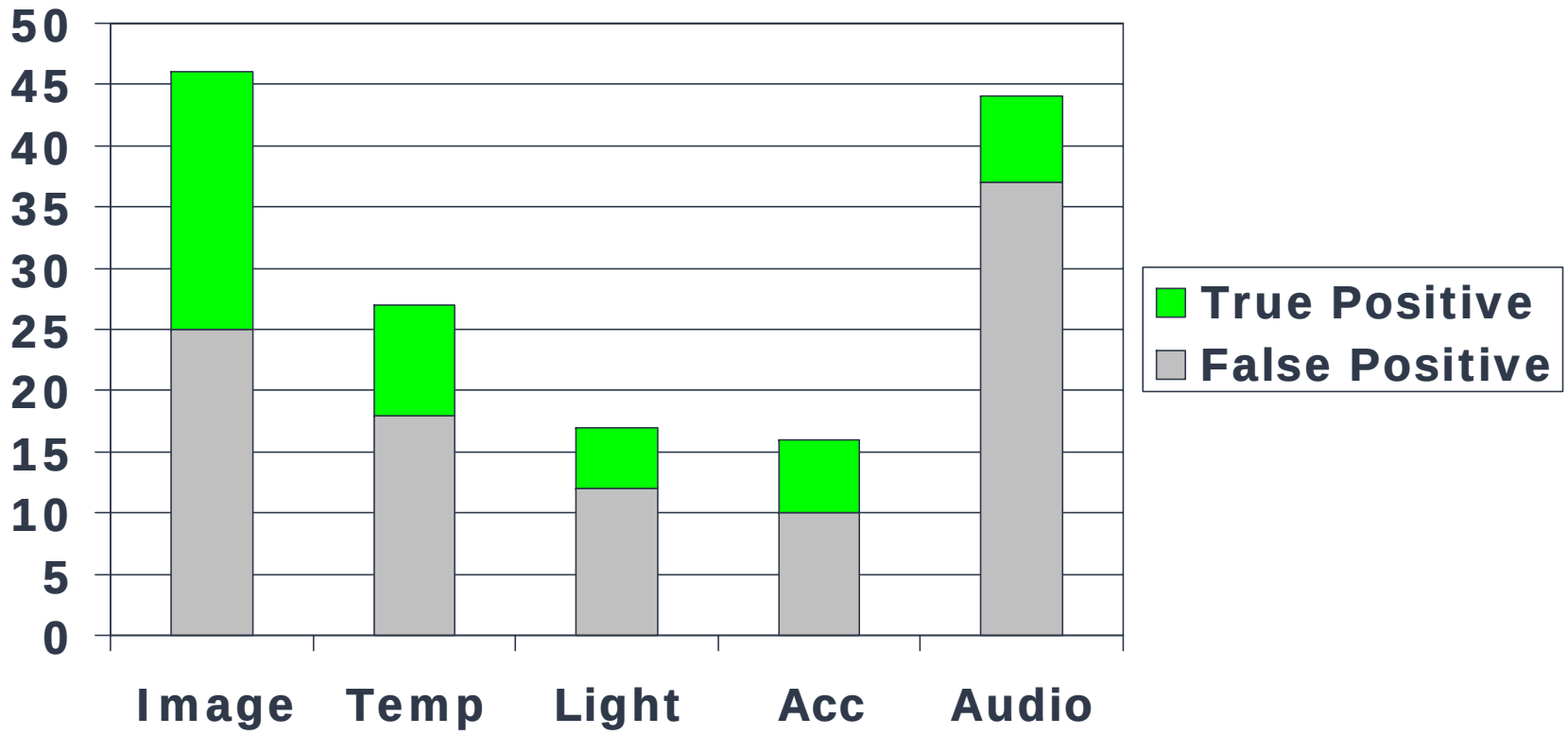
- AIM: To investigate the optimal combination of the aforementioned data sources for activity/event segmentation

2.1 Evaluation Approach



- Judge unique boundaries from each approach
- 22,173 images over 10 day period. 130 hours images, 70 hours audio

2.1 Results – individual sources



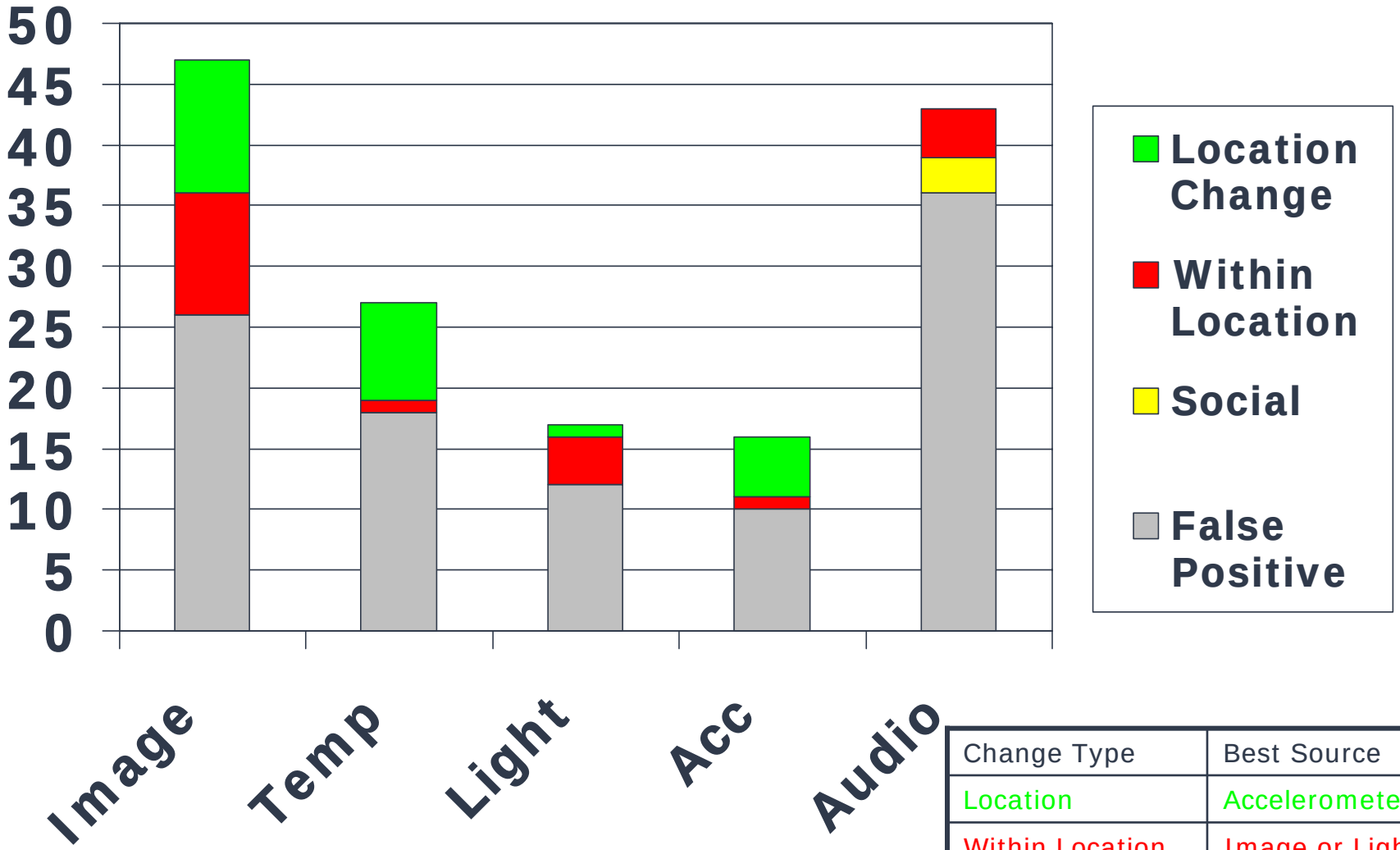
2.1 Examples of event change types

- Location changes

- Changes within the same location

- Changes due to social interaction

2.1 Results – individual sources



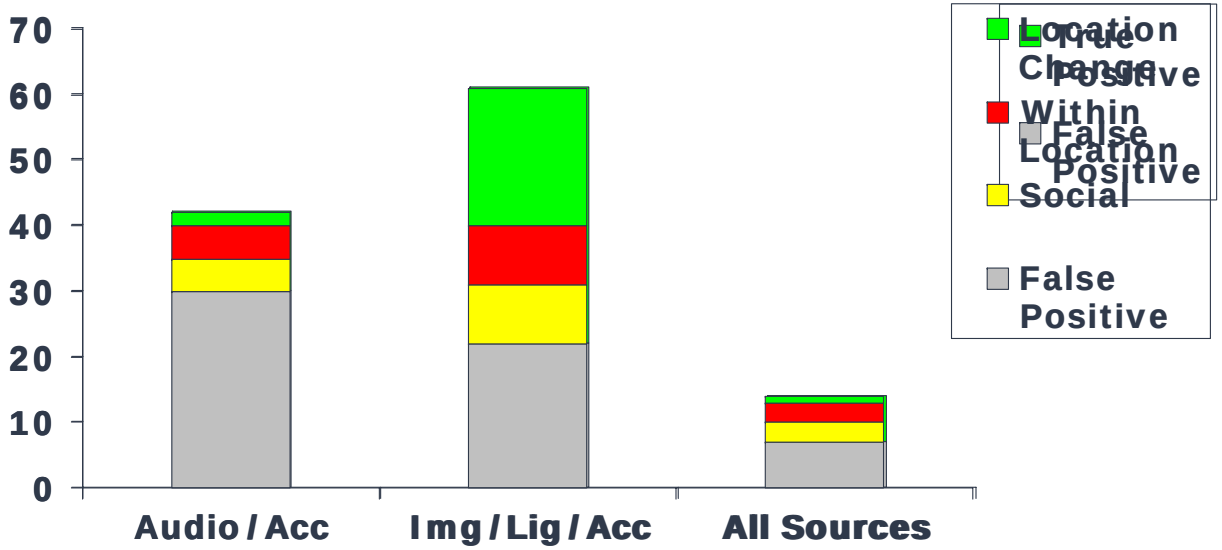
Change Type	Best Source
Location	Accelerometer
Within Location	Image or Light
Social	Audio or Img/Lig

2.1 Results – final system selection

	LOCATION	WITHIN LOCATION	SOCIAL
IMAGE			
TEMPERATURE			
LIGHT			
ACCELEROMETER			
AUDIO			

- **Audio and accelerometer**: audio for social interaction, acc for location
- **Image, light, and acc**: Image & light for within location and social, acc for location
- **All 5 sources combined**

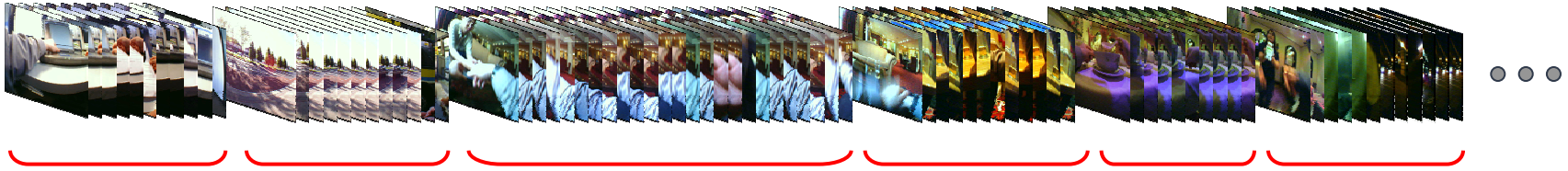
Event segmentation - results



- Combination of Image/Light/Accelerometer did best
- Produced lowest percentage of false positives and also produced largest number of unique segmentation results
- Reduced processing load, can do away with 2 sources

2.2 Event representation

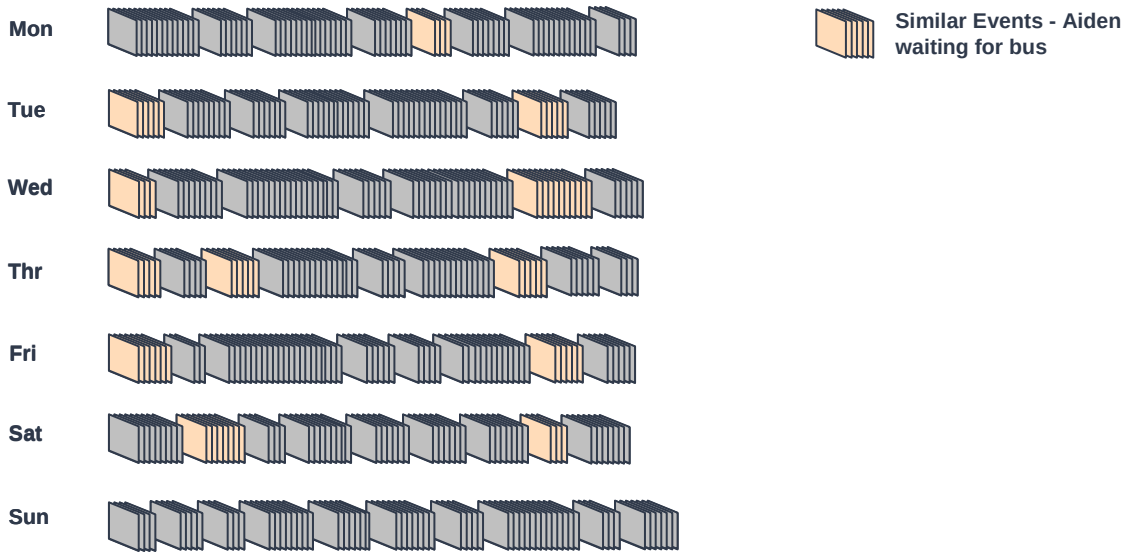
Multiple Events



- Currently to represent each event, all of the images within that event are averaged.
- It will be necessary to investigate weighting the images, with those in the middle of an event likely to be more representative of that event, than images near the start and end of the event

- Currently we represent events using low-level image descriptors only.
- It will be necessary to utilise contextual information to represent an event, e.g. location, temperature, motion, light levels.

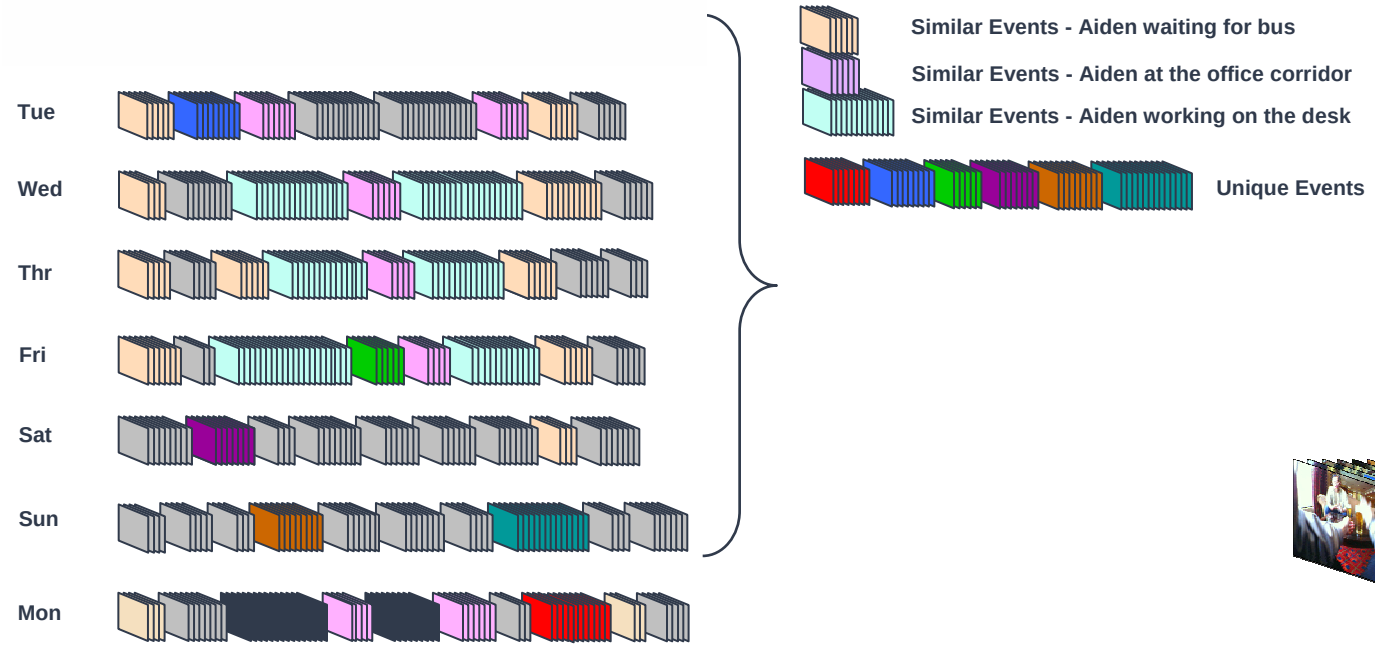
2.2 Finding similar events



- To find events that are similar to a reference event, it will be necessary to compare all events against that reference event.

- There is the research challenge of determining the optimum similarity threshold i.e. how sufficiently similar must the resultant events be?

2.3 Event uniqueness

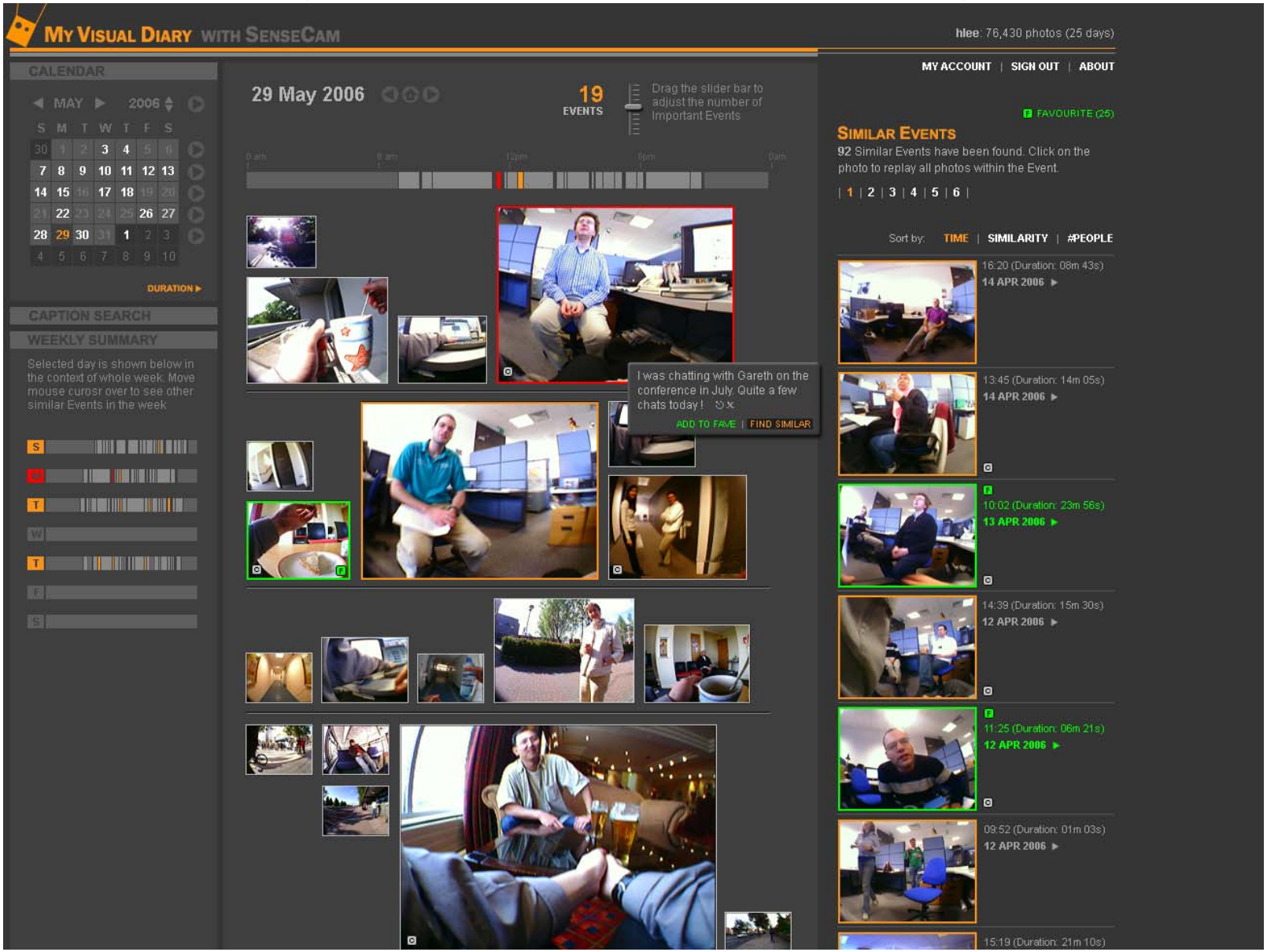


- Currently I find the most dissimilar event of today by taking the previous 6 days into account.
- What is the optimal number of previous days to consider?
- Is it better to consider the previous 6 Wednesdays as opposed to the previous 6 days?
- Optimal image to choose as event keyframe

Outline

- 1 Introduction
  - 1.1 Centre for Digital Video Processing
  - 1.2 Microsoft SenseCam
  - 1.3 Challenges
- 2 Work completed
  - 2.1 Event Segmentation
  - 2.2 Finding Similar Events
  - 2.3 Event Importance
- **3 System Demo**
- 4 Planned Work
  - 4.1 Event Augmentation

System Demo


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
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
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 !
ADD TO FAVE | FIND SIMILAR
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)

Outline

- 1 Introduction
  - 1.1 Centre for Digital Video Processing
  - 1.2 Microsoft SenseCam
  - 1.3 Challenges
- 2 Work completed
  - 2.1 Event Segmentation
  - 2.2 Finding Similar Events
  - 2.3 Event Importance
- 3 System Demo
- **4 Planned Work**
  - **4.1 Event Augmentation**

4.1 Event augmentation

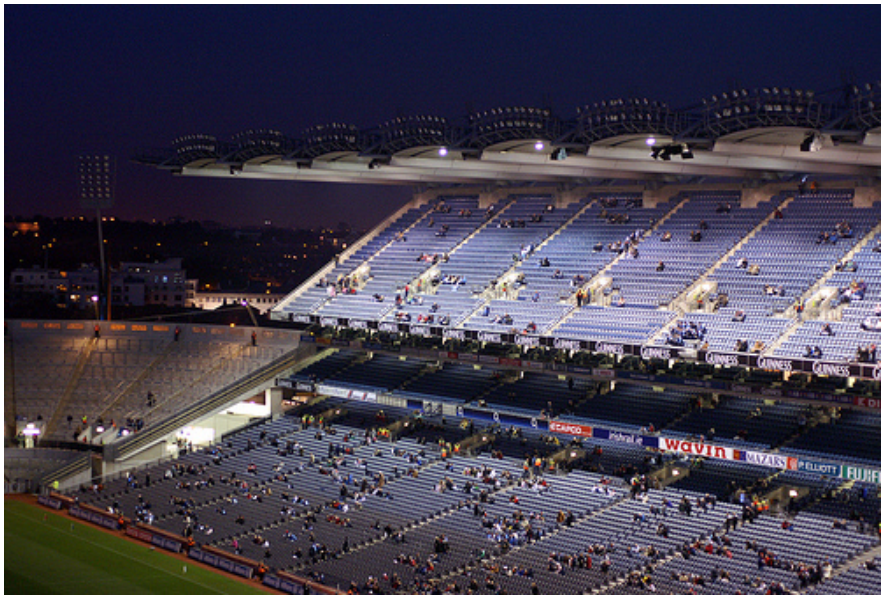
Here's an image from my SenseCam after a big match in Croke Park. I'd really like to see other people's pictures of this match.

Let's search by location…

4.1 Event augmentation

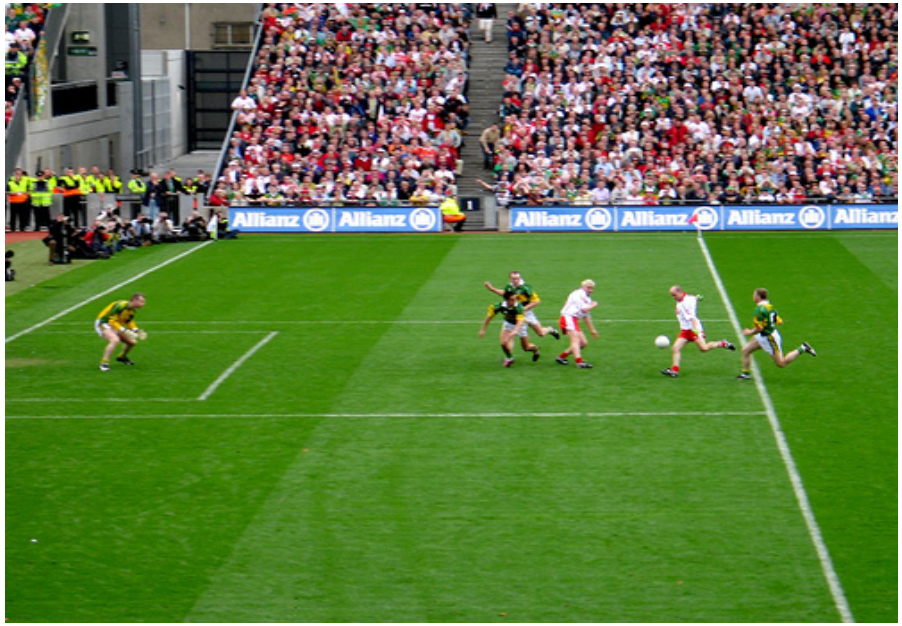
These results (based on location) are not what I wanted.

In addition I'll search by images taken at approximately same time the match was played…



4.1 Event augmentation

Excellent! These pictures are of the match I was at!



4.1 Event augmentation

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…

4.1 Event augmentation

Searching by time will not be that helpful as there was no specific event occurring at the time I was there.

Therefore let's try filtering the results to only include those that are visually most similar to the reference SenseCam image…

4.1 Event augmentation

These results are much better!

Thank You

further information:

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