Computing Moving Images: beyond the pixel

Dr Andrew Salway (a.salway@surrey.ac.uk)
Dept. Computing, University of Surrey

Cross-overs in Audiovisual Arts and Interactive Media

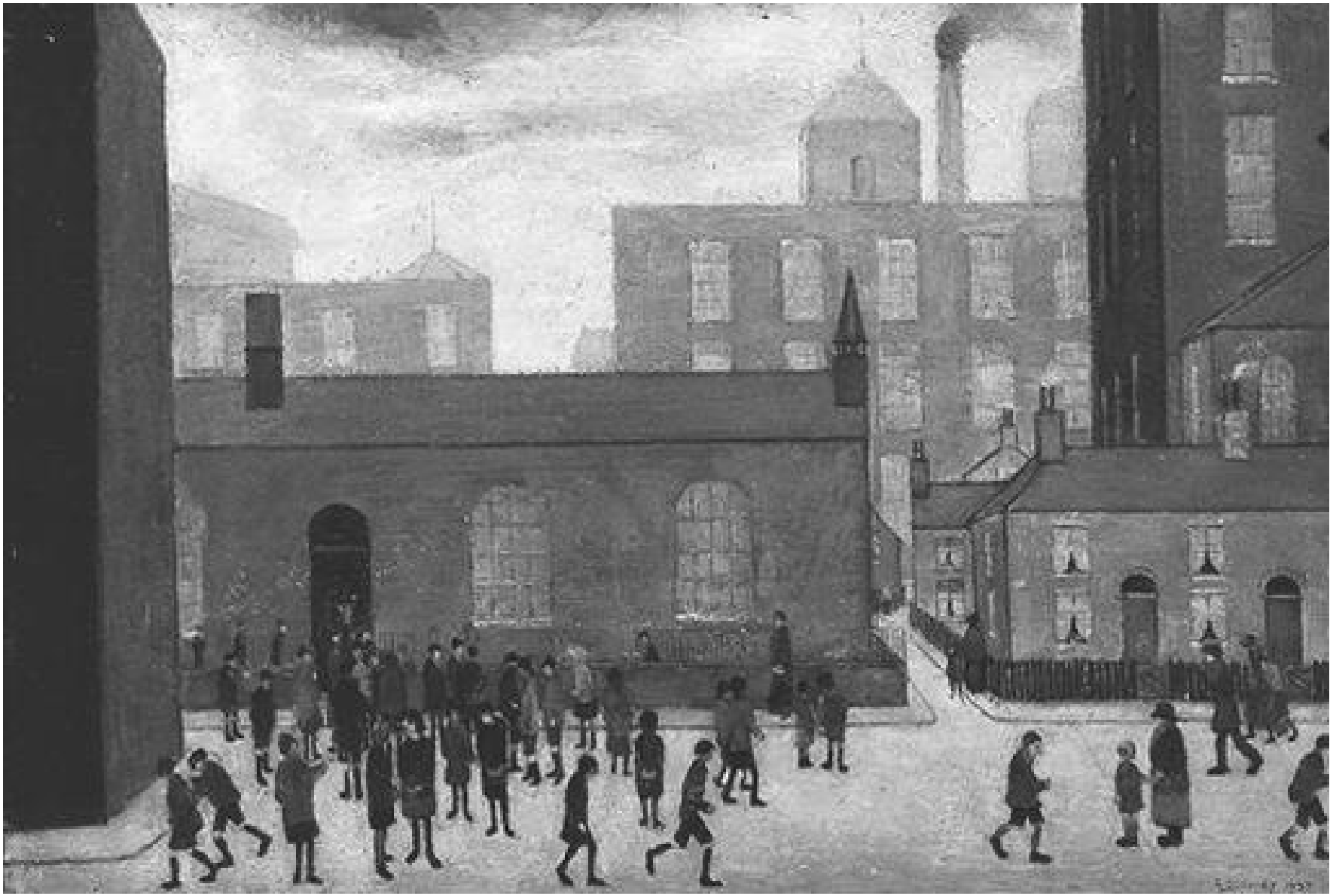
Media Centre Lume, 8 June 2004

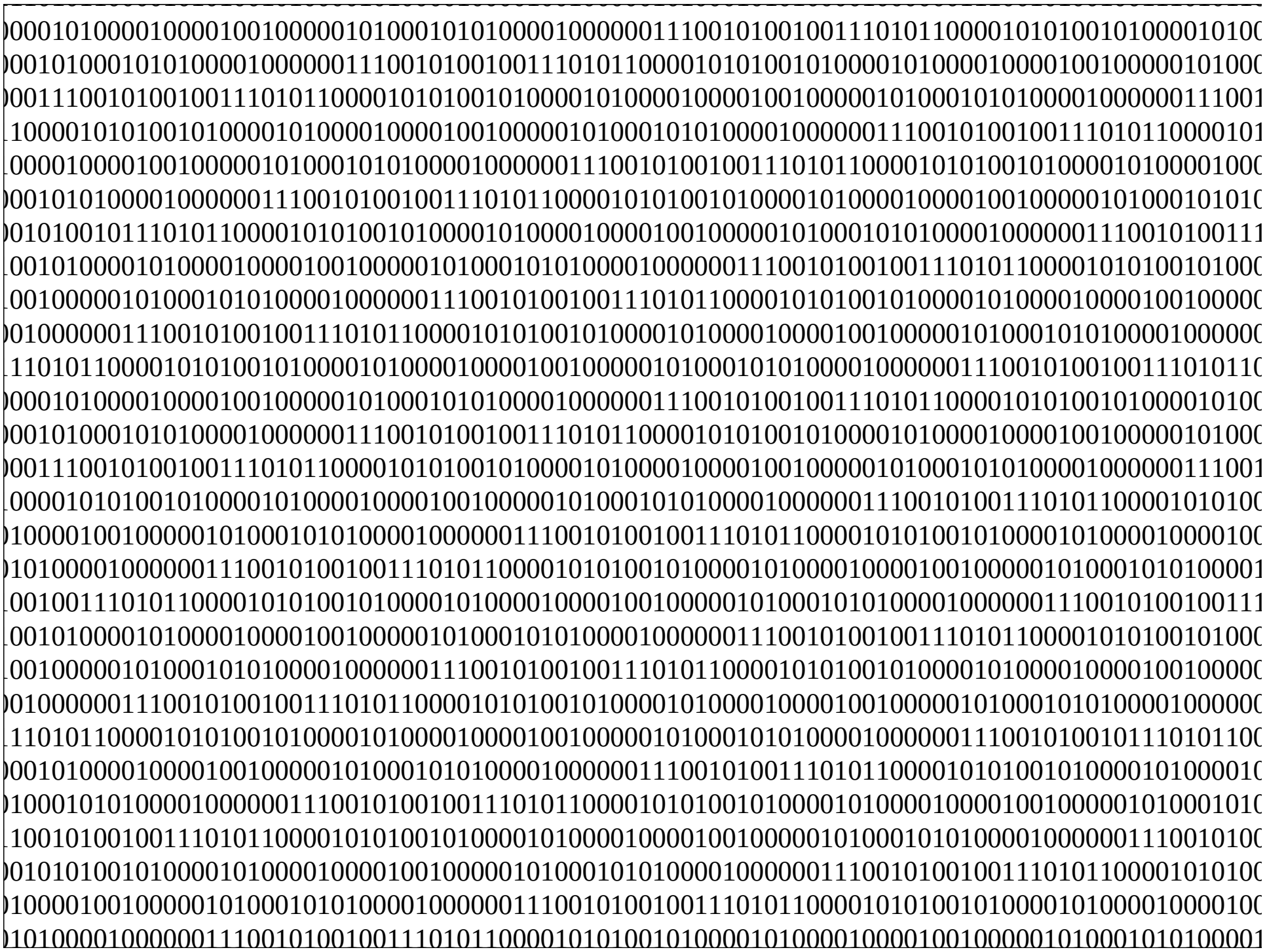
UNIVERSITY OF ART AND DESIGN HELSINKI

Computing…

- **Computer.** A device or system that is capable of carrying out a **sequence of operations** in a distinctly and **explicitly defined** manner. The operations are frequently numerical computations or **data manipulations** but also include **input/output**; the operations within the sequence **may depend on particular data values**.

Oxford Dictionary of Computing, 4th Edition, 1996.





Computing…

- **Representation.** Storage and data values used to carry information.

  Oxford Dictionary of Computing, 4th Edition, 1996.

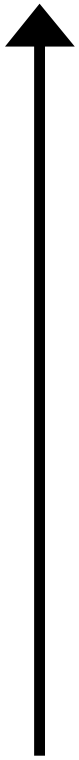
- cf. “Metadata – data about data”

Interactivity in new media is (and will be) **facilitated** (and constrained) **by machine-executable representations of media items**

Machine-executable Representations of Moving Images…

- **For a Video Rental Store Database:**
  - Film Title
  - Date
  - Director
  - Actors
  - Genre
- **For Video-on Demand** - maybe add scene-related information, e.g. actors and location
- **What about for your latest new media project?**

Levels of abstraction for moving images



Meanings…

Shots and Scenes

Actions

Movements

People / Objects

Regions (colour, shape, texture, motion)

Pixels

Electronic bits

OVERVIEW

Interactivity in new media is (and will be) **facilitated** (and constrained) **by machine-executable representations of media items**

Two issues to consider…

- The **design** of machine-executable representations: must make a commitment to one (or more) formal descriptions of media items for a particular system
- The **instantiation** of machine-executable representations: to what extent can machines generate them automatically?

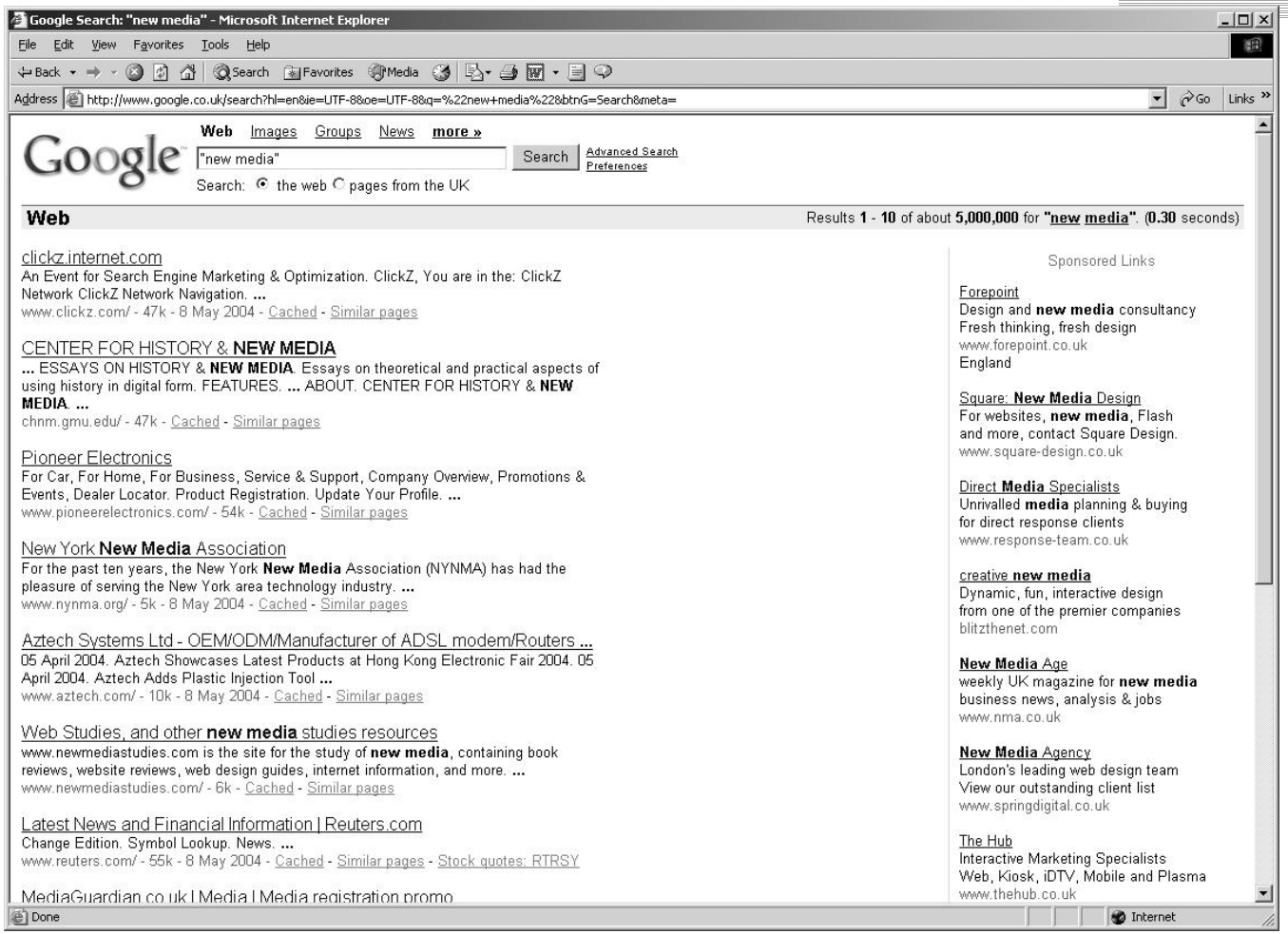
Content Technologies: **do what?**

- **Generate machine-executable representations to facilitate interactions with media**

{Retrieve, Extract, Segment, Convert, Generate} Information

Information = {text, image, video, audio} data

Information Retrieval



Information Extraction

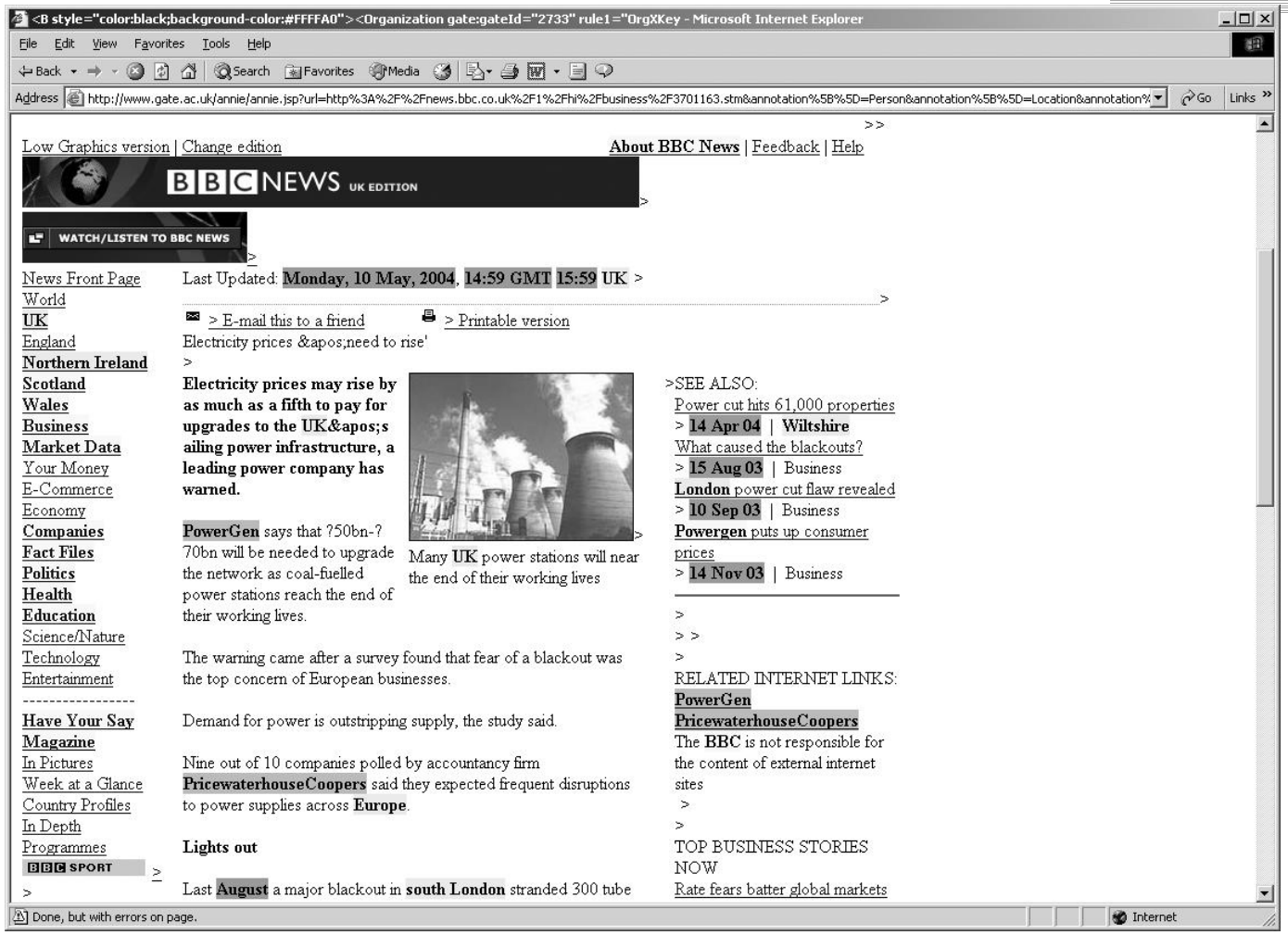
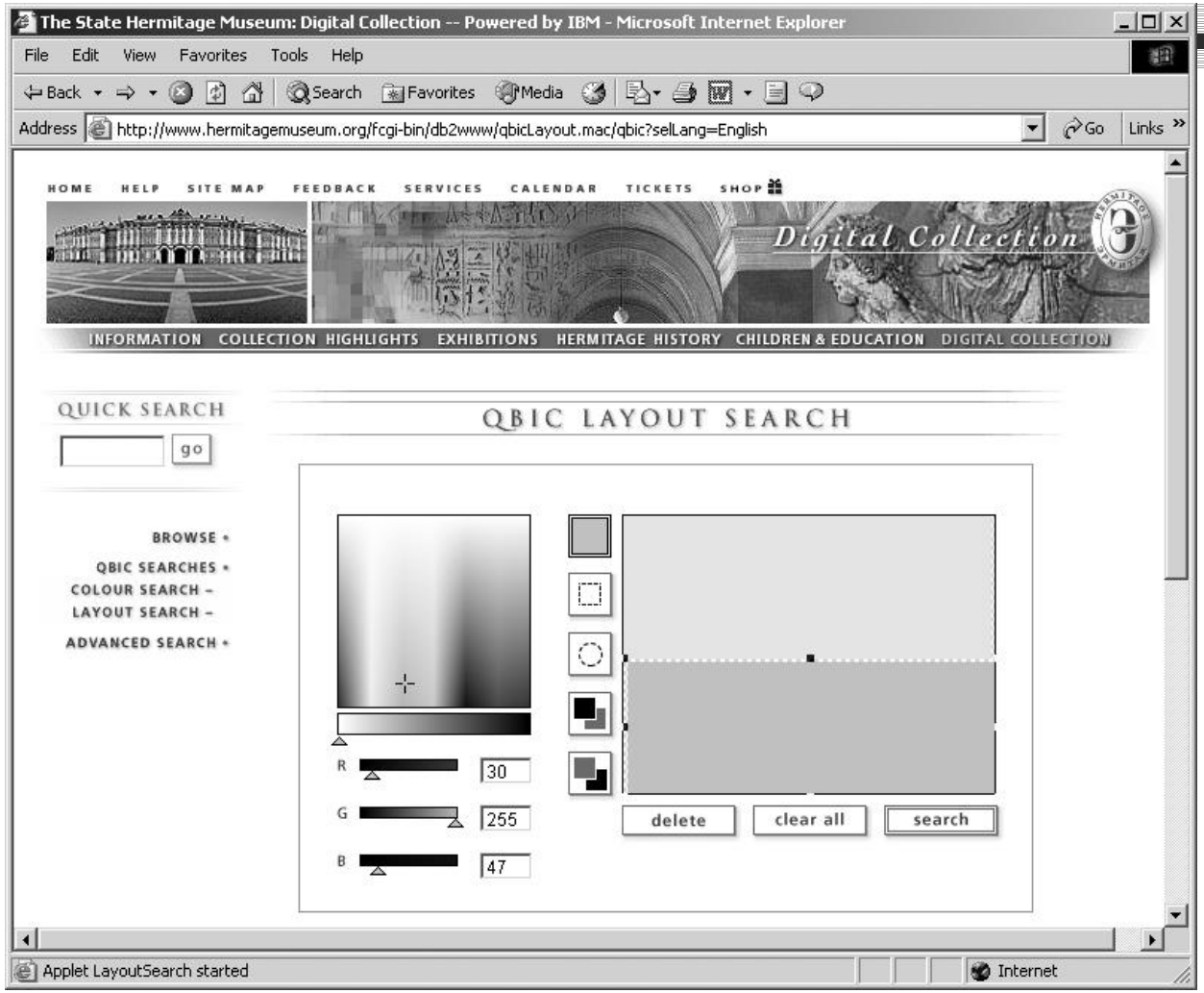


Image Retrieval:visual similarity



The State Hermitage Museum: Digital Collection -- Powered by IBM - Microsof...

File Edit View Favorites Tools Help

Back Search Favorites Media

Address www.h Go Links

1) View of the Summer Palace of Empress Elizabeth Petrovna

Grekov, Andrei Angliyevich 1753

2) Ruins of Ancient Salt Works on the Island of Salina

Houel, Jean-Pierre-Laurent Between 1776 and 1779

3) View of the English Embankment from Vasilyevsky Island (2nd section)

Paterssen, Benjamin 1799

4) View of the English Embankment and Galerny (Gallery) Dvor from Vasilyevsky Island (3rd section)

Paterssen, Benjamin 1799

5) The Gates of Narva in St Petersburg. Main Facade

Quarenghi, Giacomo 1814

6) View of the Large Palace in Peterhof from the Gulf of Finland

Artemyev, Prokofy; Chelnakov, Nikita 1761

7) Church with the Shrine in Stolnoye, the Estate of Prince

8) View of the Winter Palace from Vasilyevsky Island

The State Hermitage Museum: Digital Coll

File Edit View Favorites Tools Help

Back Search

Address www.h

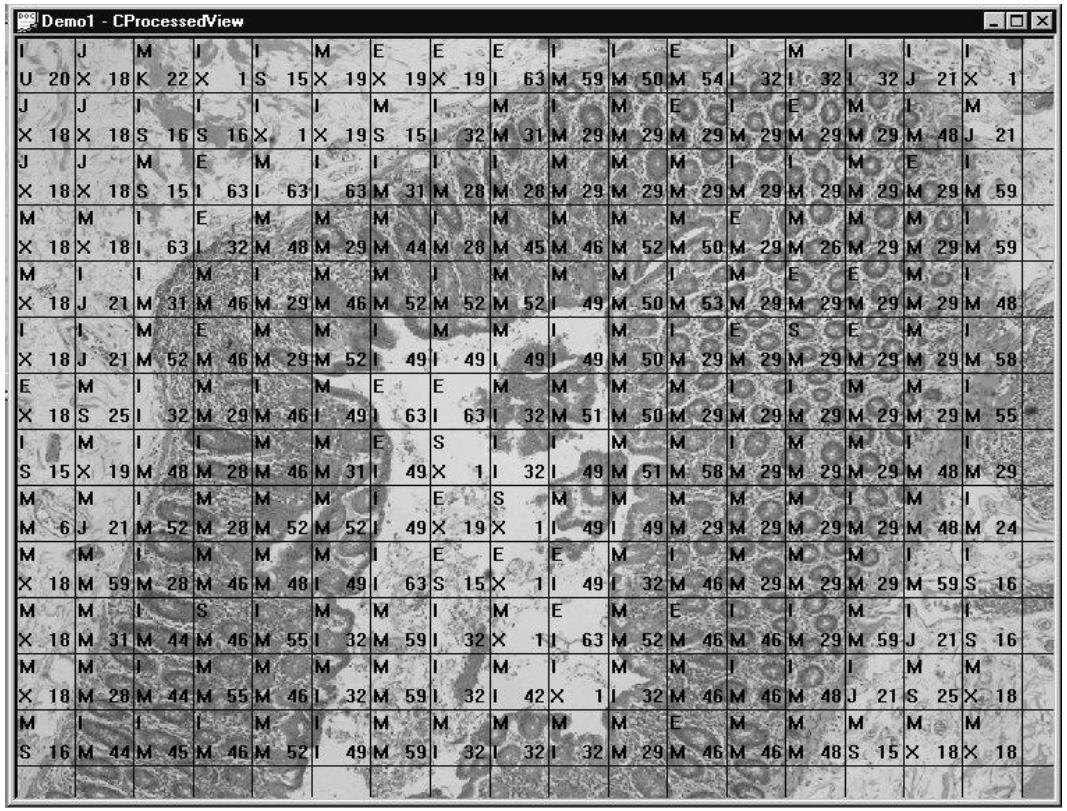
1) View of the Summer Palace of Empress Elizabeth Petrovna

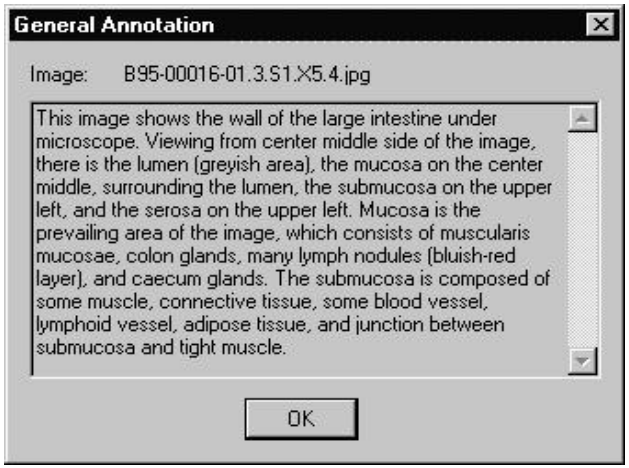
Grekov, Andrei Angliyevich 1753

3) View of the English Embankment from Vasilyevsky Island (2nd section)

Paterssen, Benjamin 1799

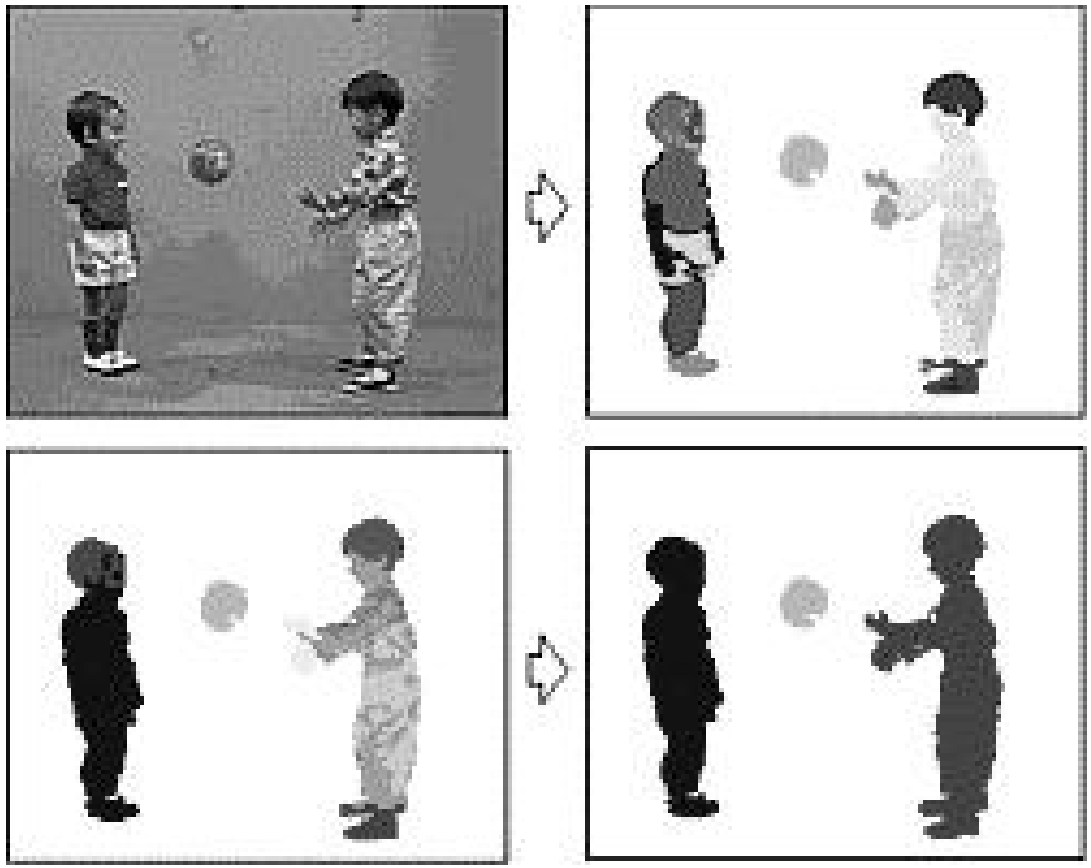
Information Conversion: image to text

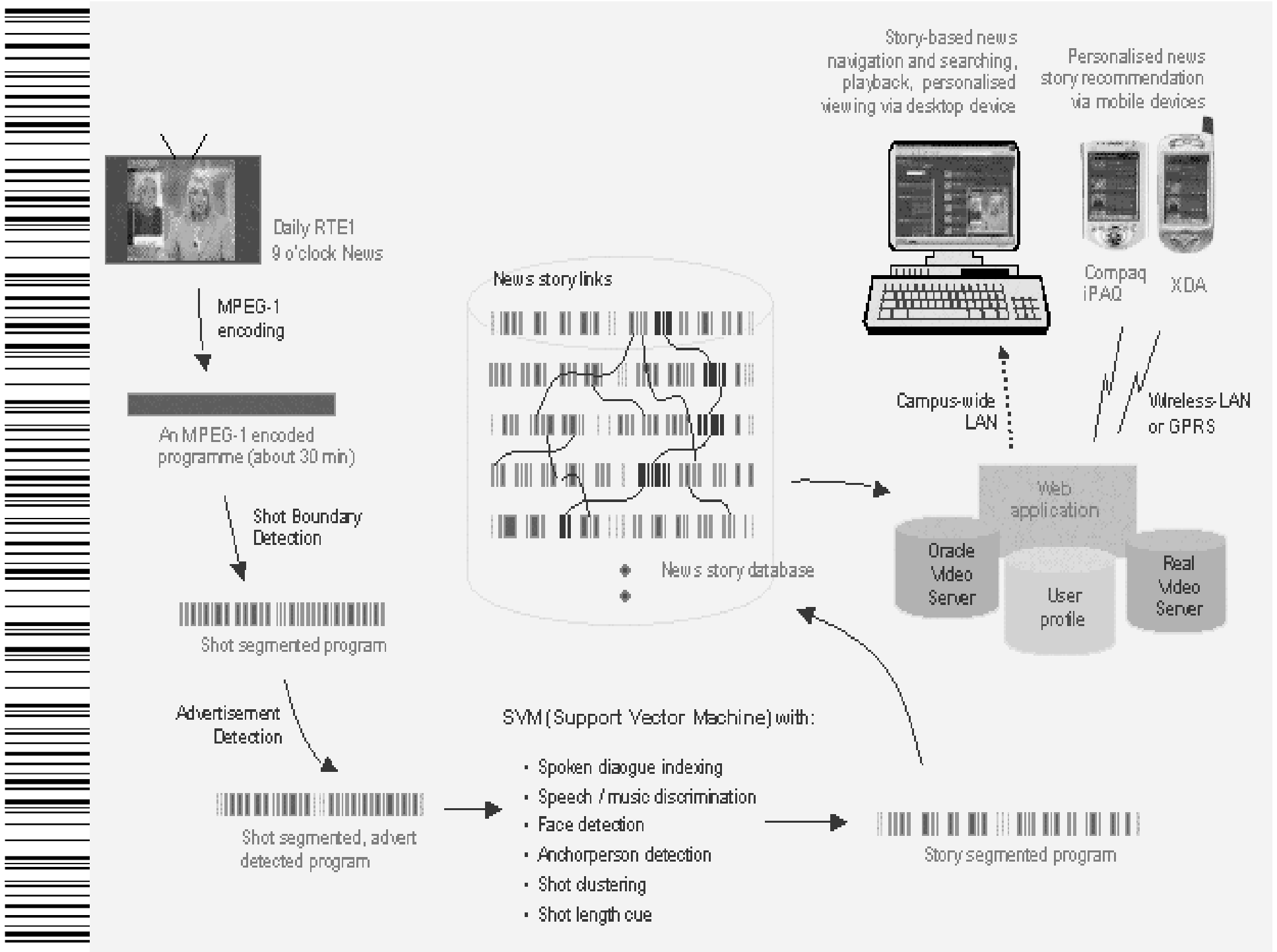




Information Conversion: text to image

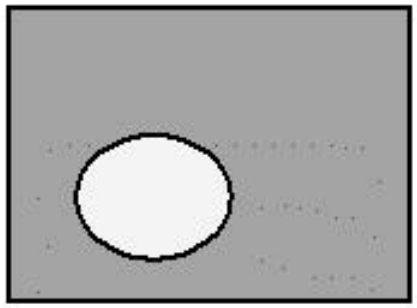
Video Segmentation




Story-based news navigation and searching, playback, personalised viewing via desktop device
Personalised news story recommendation via mobile devices
Daily RTE1 9 o'clock News
MPEG-1 encoding
An MPEG-1 encoded programme (about 30 min)
Shot Boundary Detection
Shot segmented program
Advertisement Detection
Shot segmented, advert detected program
News story links
News story database
Compaq iPAQ
XDA
Campus-wide LAN
Wireless LAN or GPRS
Web application
Oracle Video Server
User profile
Real Video Server
SVM (Support Vector Machine) with:
Spoken dialogue indexing
Speech / music discrimination
Face detection
Anchorperson detection
Shot clustering
Shot length cue
Story segmented program

Video Retrieval by Visual Similarity

Sunrise

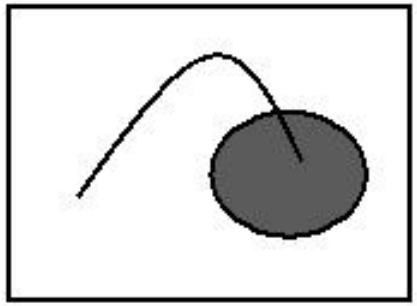


Color and size

Sky and ocean

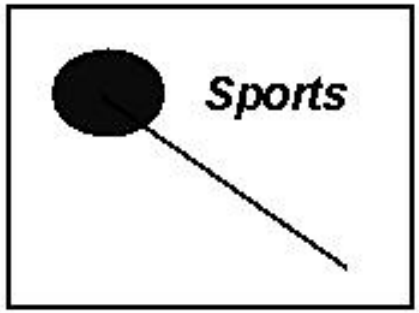
Texture and size

High jumper



Motion and size

Downhill skier



Motion and keyword

Video Retrieval with Collateral Text

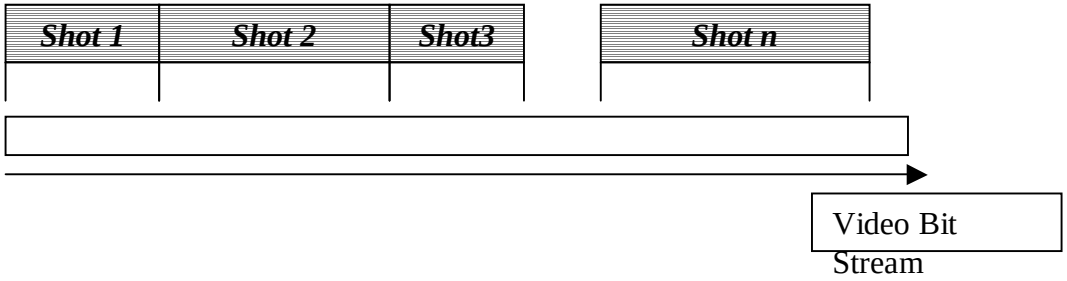
Automatic Video Editing

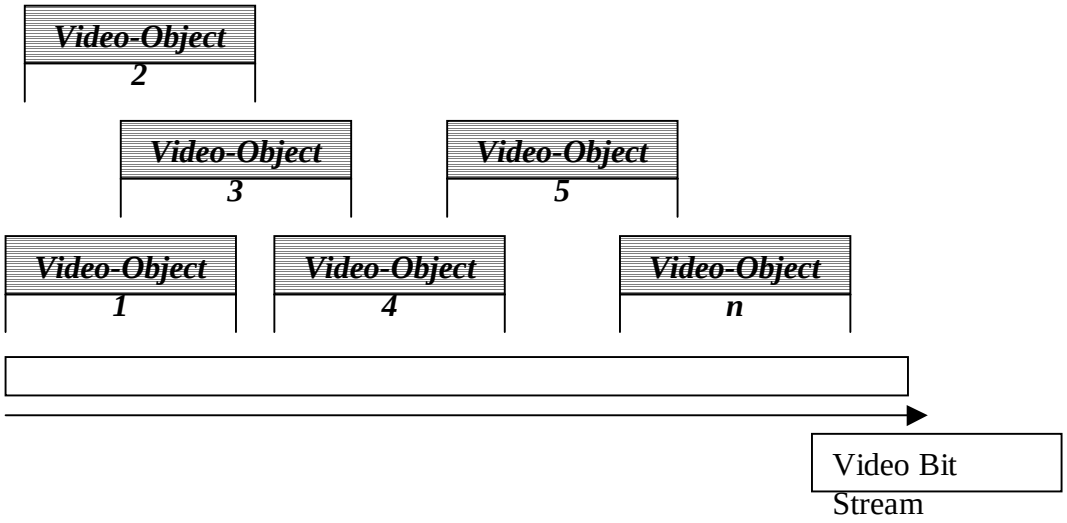
Video Data: the MPEG story

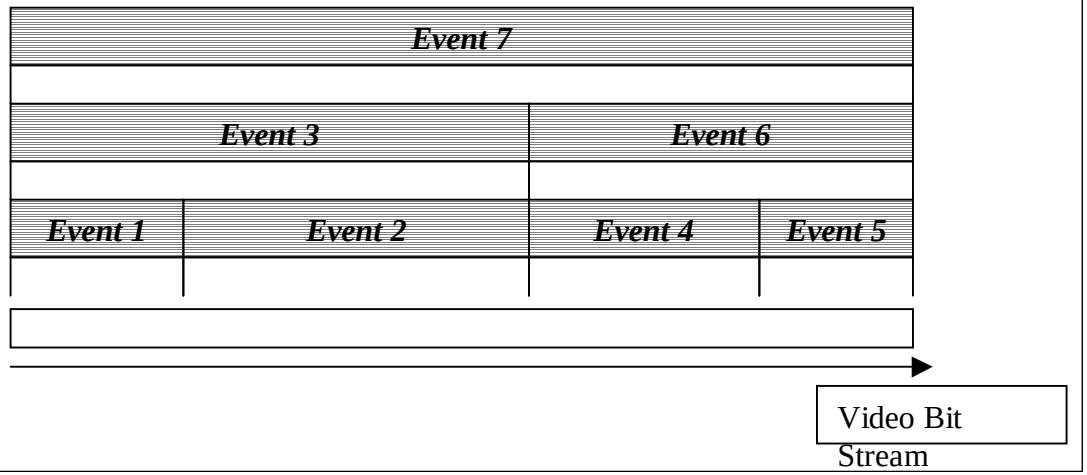
MPEG = Moving Pictures Expert Group

- Compression of Moving Images treated as sequences of frames **(MPEG-1, 1992; MPEG-2, 1996)**
- Enhanced manipulation of Moving Images treated as audio-visual objects **(MPEG-4, 2000)**
- Descriptions of content **(MPEG-7, 2001)**

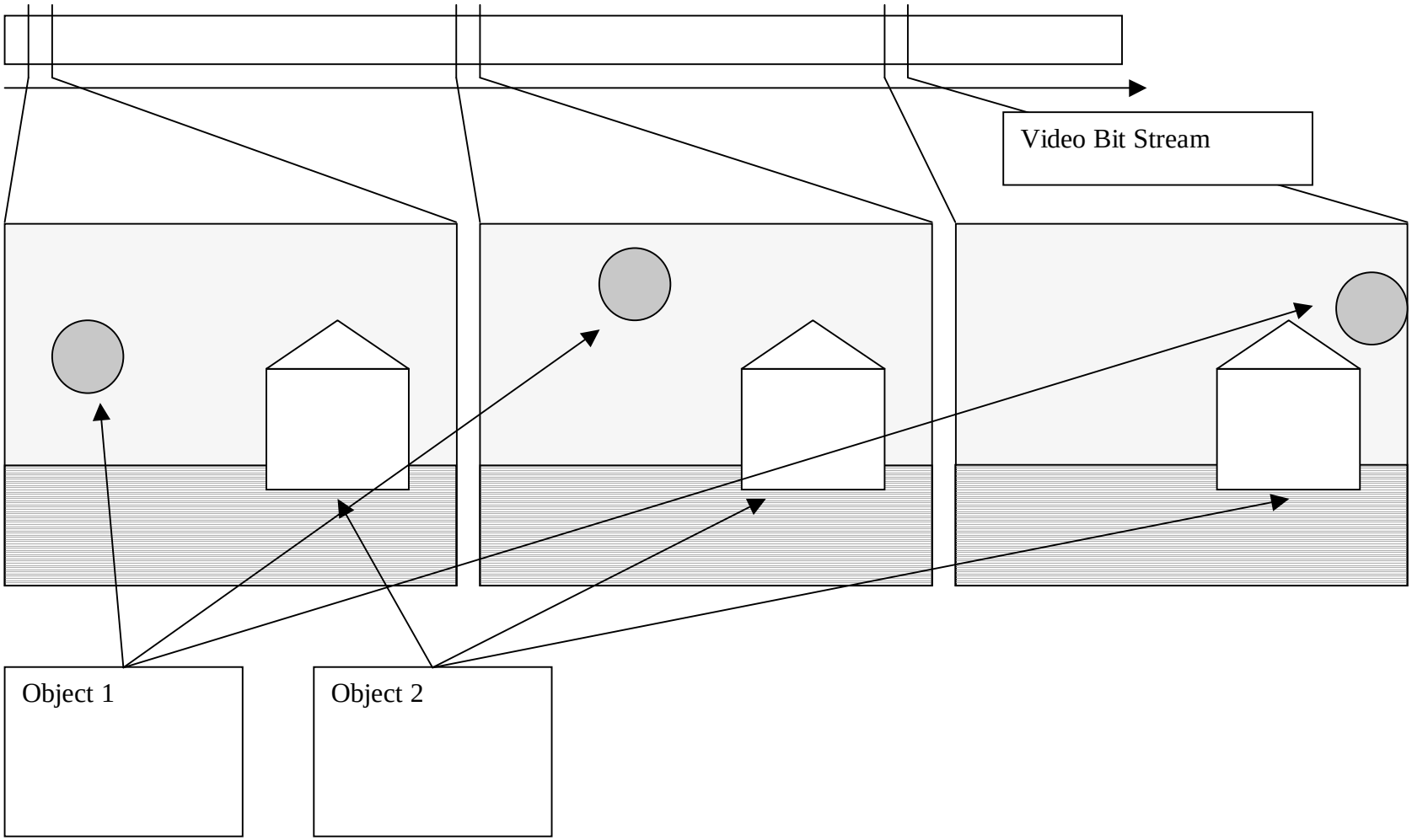
Video Data Models: interval-based







Video Data Models:object-based



Over to you…

What machine-executable representations of video data do you want?

What operations do you want the machine to perform?

Narrative: Does it compute?

- "Narrative is a sequence of connected events, organised in time and space…"
  - cf. video data models reported in IEEE Trans K&D Eng.
- BUT, agents of cause and effect are characters with goals, beliefs and emotions
- **BUT**, "More than reconstructed timelines and inventories of existents, … interpreters are called upon to live out complex blends of cognitive and imaginative response, encompassing sympathy, the drawing of causal inference, identification, evaluation, suspense, and so on"

(Herman 2002)

Computing Narrative

Narratology tells us something about…

- What are the essential features of a story? What do different manifestations of the same story have in common?
- How do different combinations of media artefacts convey stories differently?
- What are the cognitive processes involved in telling and understanding stories? What is the role of narrative in human intelligence?

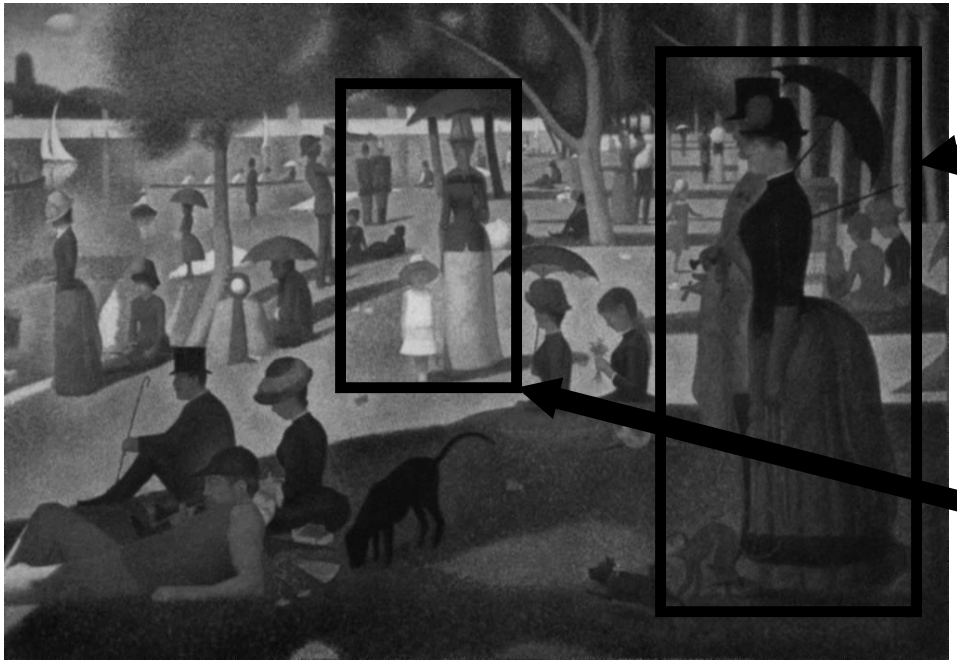
➔ Machine-executable representations of media artefacts that capture something about narrative?

Collateral Text

- Visual information often has associated textual information which is more or less informative about it; and / or text can be elicited (cf. more than one way to express a narrative)
- Content technologies can often 'understand' text data better than image/video data

➔ **Exploit collateral text to generate machine-executable representations of visual information**

What happens when you put pictures into words?



"at the extreme right, appears a scandalously hieratic-looking couple … a young mother is strolling with her little girl dressed in white with a salmon-colored sash"

I can see what appears to be a male laying in the prone position on the floor.

He is wearing a maroon striped shirt with white collar and cuffs, blue jeans, and has a pair of left and right training shoes which have become slightly dis-extended from the foot.

There appears to be a green tie down by his right hand and I can see a possible footwear impression in blood on his right hand.

Surrounding the body there are droplets of blood, footwear impression in blood and several pieces of broken glass and bottles.

5 : Describe : Swan Lake - Matthew Bourne

[0:04] lots of people on the stage
[0:08] one character is standing
[0:10] and the other characters are in a stooped, leaned over position
[0:16] very specific gestural movement of hands which brings the other characters to standing
[0:21] hands crossed behind their back
[0:23] they lean back looking at the ceiling
[0:25] the movement is slow and sustained
[0:29] the main character in the middle turns slowly and looks around
[0:35] the chorus of characters continue slow, gestural, sustained movements

15 : Describe : Swan Lake - Matthew Bourne

[0:05] a single man walks across the stage area
[0:10] his back is to the audience
[0:11] he hugs himself
[0:13] he is surrounded by a group of dancers who are bent over from the waist
[0:18] extending their arms outwards and behind their backs into a cross shape
[0:22] so that their hands cross
[0:25] they are looking upwards
[0:27] the central character who is a male who has walked across the stage wander
[0:36] meanwhile the male group of dancers of about twelve are continuing to sprea
[0:55] they wear feathered style leggings
[0:58] there is a group of four of them in the background
[1:00] the central character is kneeling
[1:05] there is now another character entered

7 : Describe : Swan Lake - Matthew Bourne

[1:08] - - - - - - attitude -
[1:14] - - arabesque -
[1:16] - - - - - - - - - - - - -
[1:30] - - adagio - - - -
[1:35] - - - - -
[1:37] - - - - - - - - - - - - - - -
[1:44] - - - - - - - - - - -
[1:47] - - - - - - - - - arabesque -
[1:56] - développés

15 : Interpret : Swan Lake - Matthew Bourne

[0:01] ok a man is surrounded by other men
[0:06] with their almost feathery-style leggings ***suggesting birds well
[0:17] their arms undulating ***like wings
[0:19] the stretching of the neck ***like a swan also ***suggesting from
[0:28] continued with the arms undulating ***like wings
[0:33] leaning over almost ***as if heads are underwater and stretching
[0:38] we have a solo figure here
[0:40] references to swan lake with the prince in relation to a group of sv
[0:44] certain balletic references
[0:47] out at night with a full-moon certainly has got some of the fantasy
[0:55] out in the park at night from the iron railings
[1:00] we have a single character who enters who we will call the lead s
[1:04] and he has been recognised ***as a solo figure
[1:08] undulating movements

Interpretation cues

Highlight Key Terms | Filter In | Filter Out | Generate Annotations

Go to Selected

Close

TIWO: Television in Words

- **AIM:** generate machine-executable representations of films from audio description (and other collateral texts), in order to support novel ways to interact with video data

- 3-year EPSRC research project, ending 2005

Audio Description Script

[11.43] Hanna passes Jan some banknotes.

[11.55] Laughing, Jan falls back into her seat as the jeep overtakes the lorries.

[12.01] An explosion on the road ahead.

[12.08] The jeep has hit a mine.

[12.09] Hanna jumps from the lorry.

[12.20] Desperately she runs towards the mangled jeep.

[12.27] Soldiers try to stop her.

[12.31] She struggles with the soldier who grabs hold of her firmly.

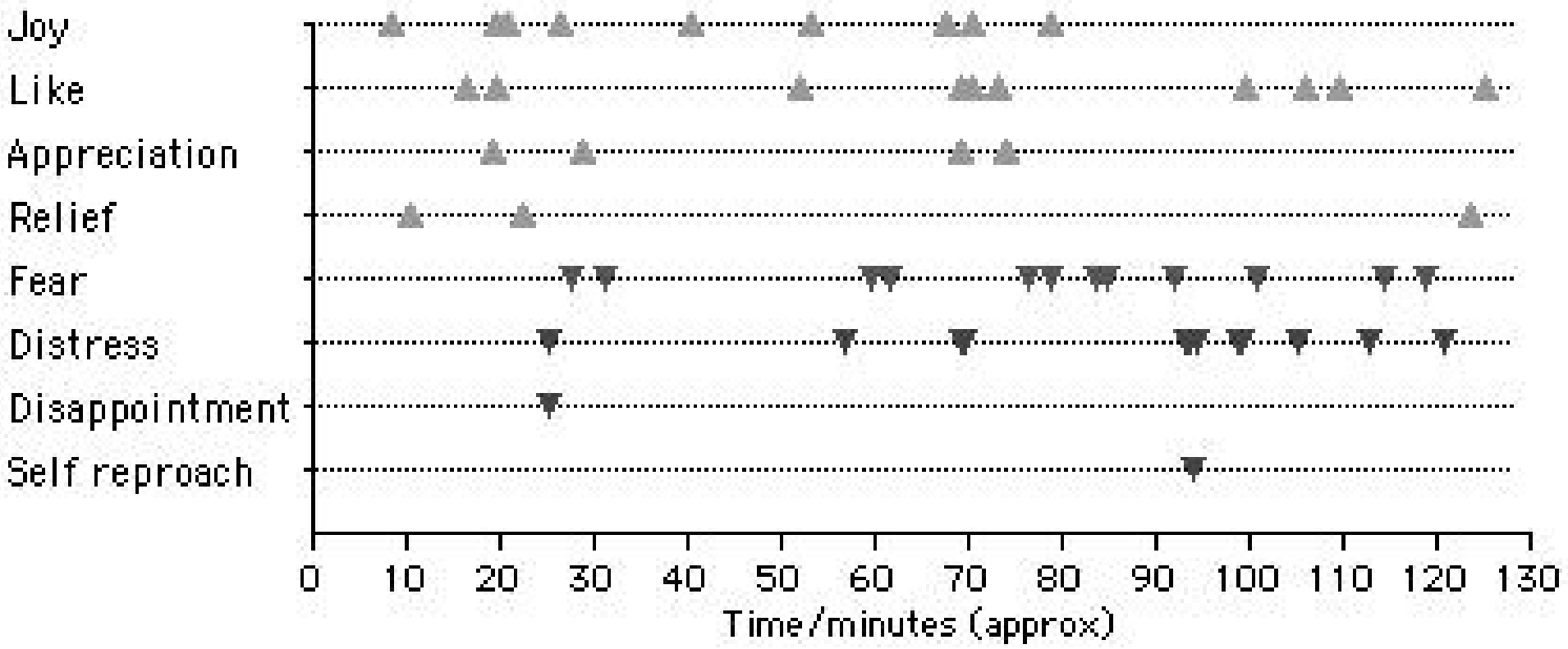
[12.35] He lifts her bodily from the ground, holding her tightly in his arms.

‘Computing Narrative’: analysis of emotion tokens

- Created lists of 600+ *emotion tokens* for 22 **EMOTION TYPES** proposed by Ortony et al. (1988):
  - **FEAR –** *afraid, fearful, petrified…*
  - **HOPE –** *anticipation, excited, expectant, optimistic…*
  - **JOY –** *euphoria, elation, happy, jolly, pleased…*
  - **DISTRESS –** *distraught, anguished, miserable, depressed…*

➔ **Take distribution of emotion tokens over time as a machine-executable representation of film content**

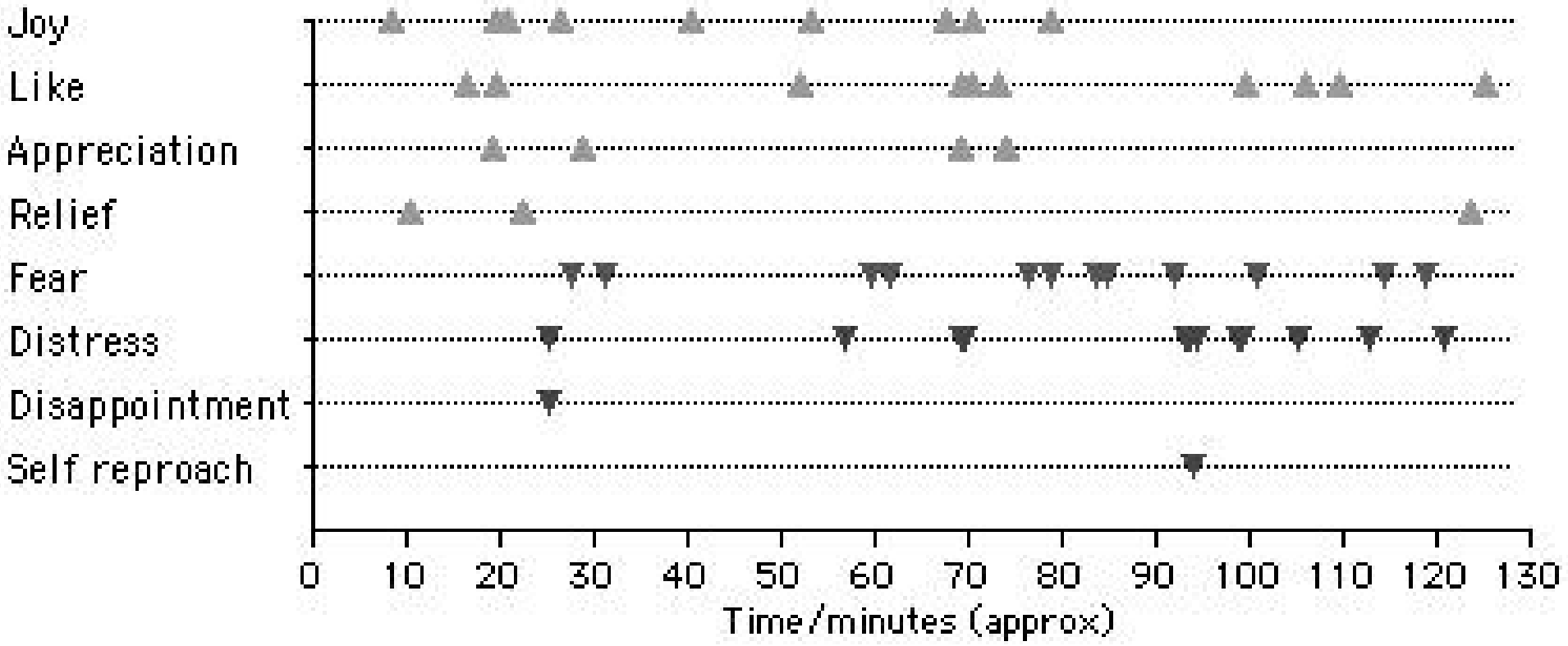
Plot of ‘Emotion Tokens’ in Audio Description for *Captain Correlli’s Mandolin*



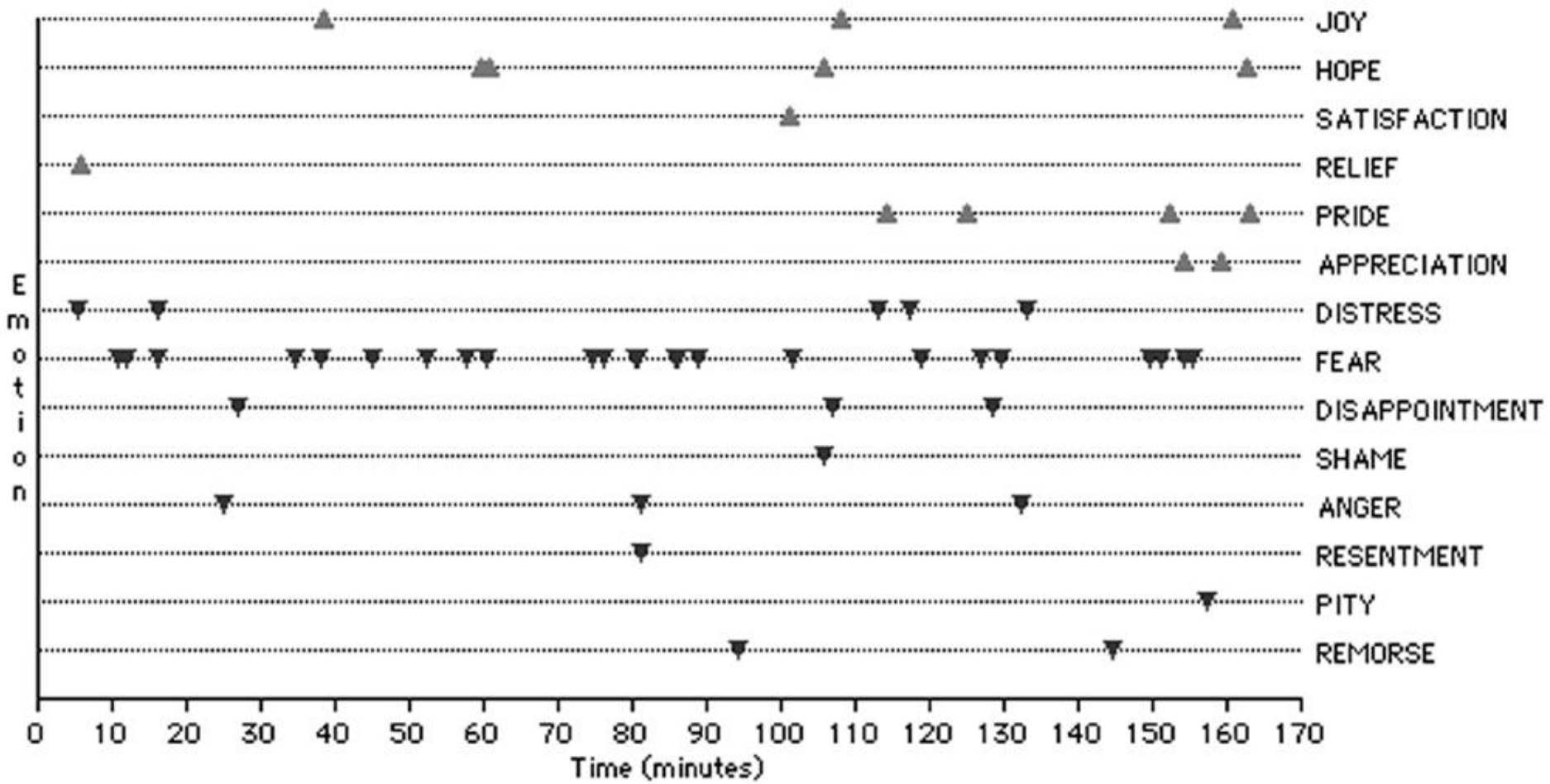
Plot of ‘Emotion Tokens’ in Audio Description for *Captain Correlli’s Mandolin*

52 tokens of 8 emotion types

15-20 minutes: Pelagria’s betrothal to Madras
20-30 minutes: invasion of the island
68-74 minutes: Pelagria and Correlli’s growing relationship
92-95 minutes: German soldiers disarm Italians



Plot of ‘Emotion Tokens’ in Audio Description for *The Postman*



‘Computing Narrative’: analysis of emotion tokens

➔Potential applications…

- Video Retrieval by Story Similarity
- Video Summarisation by Dramatic Sequences
- Video Browsing by Cause-Effect Relationships
- Browsing information space around the film, via audio description (‘DVD++’)

The Visual and The Verbal

- As it becomes increasingly easy to create, edit and distribute still and moving images, as well as text, so debates arise again about the visual and the verbal, image and text, etc.

- Perhaps it is more interesting to note that it is also becoming increasingly easy to integrate still and moving images with other kinds of media, including text, so theories about image-texts become important

How do you "read" an image-text combination? ➔

Machine-executable representations of image-text combinations

News Front Page
World
UK
England
Northern Ireland
Scotland
Wales
Business
Politics
Health
Education
Science/Nature
Technology
Entertainment

Have Your Say
Magazine
In Pictures
Week at a Glance
Country Profiles
In Depth
Programmes

BBC SPORT
BBC WEATHER
CBBC news
BBC ON THIS DAY

Last Updated: Thursday, 8 April, 2004, 07:19 GMT 08:19 UK

E-mail this to a friend
Printable version

Queen in Liverpool for ceremony

The Queen is visiting Liverpool for the ancient Maundy Thursday ceremony.

The Queen has only missed four Maundy Thursday ceremonies

During the service at the city's Anglican cathedral the Queen will hand out money to 78 men and 78 women.

Maundy Thursday marks the washing of the disciple's feet by Jesus. British monarchs used to wash subjects' feet, but James II ended the practice.

It is the first time the ceremony has been performed at the cathedral. The locals were selected for their contributions to their communities.

The ceremony is due to take place at 1100 BST.

Organisers have selected 78 men and women as a way of marking the Queen's 78th birthday celebrations which take place on 21 April.

The Queen and the Duke of Edinburgh, will arrive at Liverpool's Lime Street station before being taken the short distance to the cathedral where they will receive the traditional nosegays - posies of strong-smelling flowers.

After the service they will attend a civic reception at Liverpool Hope University.

Lord Mayor of Liverpool, Councillor Ron Gould, said: "It is a wonderful honour that the Queen has chosen to visit Liverpool this year for the Maundy Thursday celebrations.

"I am delighted that some of the people who make such a great contribution to our city are receiving this prestigious award from Her Majesty."

The Queen has only missed four Maundy Thursday ceremonies since coming to the throne in 1952.

BBC i

BBC Liverpool
Information and features on the BBC's Liverpool website

SEE ALSO:
- Bishop refuses to throw in towels
  07 Apr 04 | South Yorkshire
- Archbishops wash church-goers' feet
  17 Apr 03 | UK
- Royal visitor's silver gifts
  17 Apr 03 | Gloucestershire

RELATED BBCi LINKS:
- Maundy Thursday explained

RELATED INTERNET LINKS:
- The Queen's homepage
- Diocese of Liverpool

The BBC is not responsible for the content of external internet sites

TOP MERSEYSIDE STORIES NOW
- Funeral for murdered Shafilea
- Southport tops clean beach poll
- Queen in Liverpool for ceremony
- Drunk driver dies with friends

TOP UK STORIES NOW
- New rules for tracing relatives
- NHS 'facing new staffing crisis'
- Inquiry into WPC's killing begins
- Rate decision 'too close to call'

Tate Collections | Index - Tiscali 10.0

File Edit View Favorites Tools Help

Back Search Favorites Media

Address http://www.tate.org.uk/servlet/ViewWork?cgroupid=999999996&workid=14738 Go

Google discord chooses the apple Search Web PageRank 311 blocked AutoFill Options disco

TATE

Technology from BT Search Site: Go Tate Online together with BT

home | supporters | feedback | tickets | shop online

Collection ▸ Turner Collection ▸ Oil paintings ▸ Work

Work | On Display | Images | Subjects | Related

Joseph Mallord William Turner 1775-1851

The Goddess of Discord Choosing the Apple of Contention in the Garden of the Hesperides exhibited 1806

Oil on canvas
support: 1553 x 2184 mm
painting

Bequeathed by the artist 1856

N00477

The Hesperides were the three daughters of Hesperus, the evening star. They kept watch over a tree of golden apples in a garden on the slopes of Mount Atlas. Here, the goddess Discord chooses the apple that will eventually lead to the Trojan war. Paris offered the apple to the goddess Aphrodite who, in return, offered him the most beautiful woman in the world: the Greek Queen, Helen.

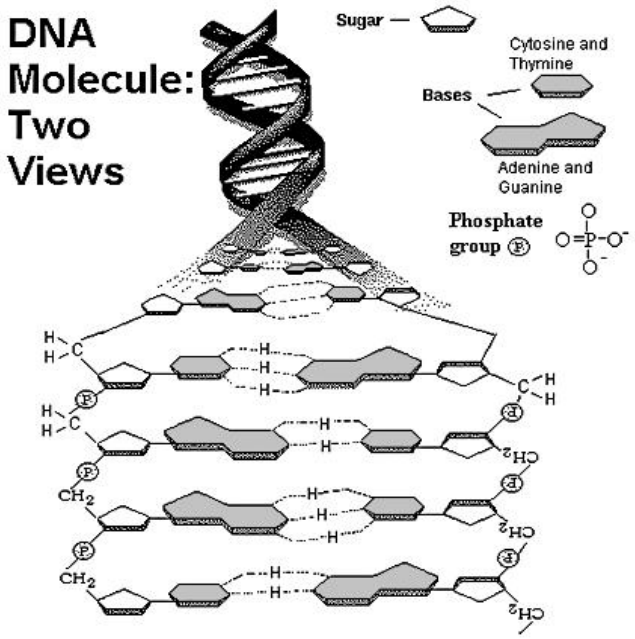
The classical grandeur of this landscape is based on the work of the seventeenth-century French artist, Nicolas Poussin. Its background reflects Turner's experience of the Alps in 1802
(From the display caption February 2004)

Tate Collections
- Search
- General collection
- Turner
- Special collections
- Works on display
- Collection highlights
- My selection
- Archive
- Library
- Print Rooms
- Insight
- Feedback

Research Services
Tate Britain
Tate Modern
Tate Liverpool
Tate St Ives
Tate Connections

Internet

DNA Molecule - Two Views



DNA Molecule - Two Views

Legend:

The double helix of the DNA is shown along with details of how the bases, sugars and phosphates connect to form the structure of the molecule.

DNA is a double-stranded molecule twisted into a helix (think of a spiral staircase). Each spiraling strand, comprised of a sugar-phosphate backbone and attached bases, is connected to a complementary strand by non-covalent hydrogen bonding between paired bases. The bases are adenine (A), thymine (T), cytosine (C) and guanine (G).

A and T are connected by two hydrogen bonds. G and C are connected by three hydrogen bonds.

News Front Page
World
UK
England
Northern Ireland
Scotland
Wales
Business
Politics
Health
Education
Science/Nature
Technology
Entertainment

Have Your Say
Magazine
In Pictures
Week at a Glance
Country Profiles
In Depth
Programmes

BBC SPORT
BBC WEATHER
CBBC news
BBC ON THIS DAY

Last Updated: Thursday, 8 April, 2004, 07:19 GMT 08:19 UK

E-mail this to a friend
Printable version

Queen in Liverpool for ceremony

The Queen is visiting Liverpool for the ancient Maundy Thursday ceremony.

The Queen has only missed four Maundy Thursday ceremonies

During the service at the city's Anglican cathedral the Queen will hand out money to 78 men and 78 women.

Maundy Thursday marks the washing of the disciple's feet by Jesus. British monarchs used to wash subjects' feet, but James II ended the practice.

It is the first time the ceremony has been performed at the cathedral. The locals were selected for their contributions to their communities.

The ceremony is due to take place at 1100 BST.

Organisers have selected 78 men and women as a way of marking the Queen's 78th birthday celebrations which take place on 21 April.

The Queen and the Duke of Edinburgh, will arrive at Liverpool's Lime Street station before being taken the short distance to the cathedral where they will receive the traditional nosegays - posies of strong-smelling flowers.

After the service they will attend a civic reception at Liverpool Hope University.

Lord Mayor of Liverpool, Councillor Ron Gould, said: "It is a wonderful honour that the Queen has chosen to visit Liverpool this year for the Maundy Thursday celebrations.

"I am delighted that some of the people who make such a great contribution to our city are receiving this prestigious award from Her Majesty."

The Queen has only missed four Maundy Thursday ceremonies since coming to the throne in 1952.

BBC i

BBC Liverpool
Information and features on the BBC's Liverpool website

SEE ALSO:
- Bishop refuses to throw in towels
  07 Apr 04 | South Yorkshire
- Archbishops wash church-goers' feet
  17 Apr 03 | UK
- Royal visitor's silver gifts
  17 Apr 03 | Gloucestershire

RELATED BBCi LINKS:
- Maundy Thursday explained

RELATED INTERNET LINKS:
- The Queen's homepage
- Diocese of Liverpool

The BBC is not responsible for the content of external internet sites

TOP MERSEYSIDE STORIES NOW
- Funeral for murdered Shafilea
- Southport tops clean beach poll
- Queen in Liverpool for ceremony
- Drunk driver dies with friends

TOP UK STORIES NOW
- New rules for tracing relatives
- NHS 'facing new staffing crisis'
- Inquiry into WPC's killing begins
- Rate decision 'too close to call'

BBC NEWS | World | Middle East | Baghdad stunned by hotel attack - Tiscali 10.0

File Edit View Favorites Tools Help

Back Search Favorites Media

Address http://news.bbc.co.uk/1/hi/world/middle_east/3522424.stm Go

Google Search Web PageRank 268 blocked AutoFill Options

BBC i | CATEGORIES | TV | RADIO | COMMUNICATE | WHERE I LIVE | INDEX | SEARCH Go

Low Graphics version | Change edition | About BBC News | Feedback | Help

BBC NEWS UK EDITION | WATCH/LISTEN TO BBC NEWS

Last Updated: Thursday, 18 March, 2004, 10:45 GMT

E-mail this to a friend Printable version

Baghdad stunned by hotel attack

Rescuers in Baghdad have given up the search for survivors in the rubble left by a massive blast which killed at least 17 people, most of them Iraqis.

The explosion on Wednesday night wrecked the small Mount Lebanon hotel and nearby houses in the city centre.

The US military says a 450kg car bomb packed with artillery ordnance killed at least 17 people and injured 45.

On Thursday, three people also died in an attack on staff at a US-funded TV station in Baquba, north of Baghdad.

The hotel would have been a soft target for bombers

US officials say the hotel attack, in the busy Karrada district, was intended to cause maximum casualties and bore the hallmarks of an al-Qaeda-related group.

One Briton was killed, and another injured, the UK government said.

US officials initially put the death toll as high as 27, but later revised it downwards. Iraqi officials said the toll was as low as six.

Hours after the blast, a mortar attack on a US logistics base north of Baghdad killed two soldiers and injured six.

'Friendly' hotel

Most of the victims of the hotel bombing were Iraqis living

WATCH AND LISTEN
The BBC's Paul Wood
...ple were killed as they were sitting down to dinner"
VIDEO

AFTER SADDAM
KEY STORIES
- US bombs mosque complex
- In pictures: Iraq turmoil
- US resolve in Iraq 'unshakable'
- New 'al-Qaeda' warning on Iraq

ANALYSIS

Anniversary blues
One year on, can the US clear up the mess in Iraq, asks the BBC's Paul Wood
- Mehdi Army
- Shia split over 'uprising'
- Allies plan crisis talks
- Shia discontent
- Handover questioned

OPEN Iraqi opinion poll results

BACKGROUND
- Who's behind the attacks?
- Who's who in Iraq
- Timeline: Saddam's Iraq

WATCH/LISTEN
VIDEO Interactive video

HAVE YOUR SAY
- Can the violence be stopped?
- Full in-depth report

RELATED INTERNET LINKS:
- Coalition Provisional Authority

Done Internet

BBC NEWS | World | Middle East | US troops face longer Iraq stay - Tiscali 10.0

File Edit View Favorites Tools Help

Back Search Favorites Media

Address http://news.bbc.co.uk/1/hi/world/middle_east/3609801.stm Go

Google Search Web PageRank 268 blocked AutoFill Options

BBC NEWS UK EDITION

WATCH/LISTEN TO BBC NEWS

Last Updated: Thursday, 8 April, 2004, 07:34 GMT 08:34 UK

E-mail this to a friend Printable version

US troops face longer Iraq stay

US Defence Secretary Donald Rumsfeld says some "seasoned" US troops in Iraq might be kept there longer than planned to deal with the growing violence.

Accounts vary about casualties in Falluja

The troops have not lost control of the security situation, despite a recent upsurge in fighting, Mr Rumsfeld said.

The clashes were the work of a few "thugs, gangs and terrorists", he said.

In the latest reported violence, US troops were involved in clashes with both Sunni and Shia insurgents in the north and west of the capital Baghdad.

Operations were also continuing in the Sunni city of Falluja west of the capital - a day after the US military bombed a compound housing a mosque.

Map locating recent clashes

Iraqi witnesses say about 40 Iraqis were killed in the strike, but US Central Command said only one "anti-coalition force member" had died and there were no civilian casualties.

It is not known if there were any deaths among the five casualties reported by the US military.

Several days of clashes have claimed the lives of well over 100 Iraqis and at least 30 coalition soldiers in what is being seen as the worst escalation of fighting since Baghdad fell to US-led forces a year ago.

> **"We will likely be managing the pace of the redeployments to allow those seasoned troops... to see the current situation through"**
>
> Donald Rumsfeld

In pictures: Iraq in turmoil

WATCH AND LISTEN
The BBC's Rebecca Pearce
"Coalition forces are fighting ...ns on two fronts"
VIDEO

AFTER SADDAM

KEY STORIES
- US bombs mosque complex
- In pictures: Iraq turmoil
- US resolve in Iraq 'unshakable'
- New 'al-Qaeda' warning on Iraq

ANALYSIS

Anniversary blues
One year on, can the US clear up the mess in Iraq, asks the BBC's Paul Wood

- Mehdi Army
- Shia split over 'uprising'
- Allies plan crisis talks
- Shia discontent
- Handover questioned

OPEN Iraqi opinion poll results

BACKGROUND
- Who's behind the attacks?
- Who's who in Iraq
- Timeline: Saddam's Iraq

WATCH/LISTEN
VIDEO Interactive video

HAVE YOUR SAY
- Can the violence be stopped?
- Full in-depth report

TOP MIDDLE EAST STORIES NOW
- US troops face longer Iraq stay
- Iran 'to press on with reactor'
- Israel breaks up barrier protest
- Arafat dismisses Sharon threat

Internet

BBC NEWS | World | Asia-Pacific | China sounds alarm on Aids - Tiscali 10.0

File Edit View Favorites Tools Help

Back Search Favorites Media

Address http://news.bbc.co.uk/1/hi/world/asia-pacific/3610281.stm Go

Google Search Web PageRank 268 blocked AutoFill Options

BBC NEWS UK EDITION

WATCH/LISTEN TO BBC NEWS

Last Updated: Thursday, 8 April, 2004, 08:56 GMT 09:56 UK

E-mail this to a friend Printable version

China sounds alarm on Aids

China's government has called for an increased effort to stop the spread of HIV/Aids, warning of severe punishment for any attempt at a cover-up.

Yu Daguan, an Aids sufferer in Dongguan village, in China's central Henan province

China hid a blood scandal in the 90s which infected entire villages

Health Minister Wu Yi told officials the epidemic was at a critical point where it could spread from high-risk groups to the wider public.

China admits to having more than 800,000 cases of HIV/Aids, but experts say the real figure could be higher.

The UN believes at least 10 million Chinese could be HIV positive by 2010.

"We can completely contain the momentum if we take it seriously. Otherwise, we will lose this best, fleeting opportunity," Ms Wu told regional health officials at a conference on HIV/Aids in Beijing.

"It must be reported timely and faithfully. And any people who intend to hide the epidemic should take responsibility and will be severely punished," the China Daily quoted the minister as saying.

The majority of Chinese still understand little about Aids

This emphasis on openness suggests China has learned lessons from the Sars (Severe Acute Respiratory Syndrome) crisis it suffered last year, which was worsened by an initial attempt by the authorities to suppress news of the disease, preventing adequate prevention measures.

China also announced on Thursday the establishment of its first formal research centre on Aids, in Shanghai.

'The genocide of a generation'

ASK THE EXPERTS
- Botswana's President Mogae
- Head of UNAids Peter Piot

FEATURES
- Gay Cubans fight own Aids battle
- Iran's drug users face Aids risk

LIVING WITH HIV
- Huang, China: Trying not to worry
- Natasha, Russia: New aim in life
- In pictures: Ray's last weeks

IN DEPTH
- The Biology of Aids
- Map: Global spread of HIV

OPEN Impact of Aids

- **Aids debate: In-depth site**

RELATED BBCi LINKS:
- BBC Chinese.com

RELATED INTERNET LINKS:
- Xinhua
- China AIDS survey

The BBC is not responsible for the content of external internet sites

TOP ASIA-PACIFIC STORIES NOW
- China sounds alarm on Aids
- Japan PM hints at retirement
- Japan probes Iraq blasts
- Indonesia vote too close to call

Internet

Classifying Image-Text Relations

- **How to 'read' an image-text combination?**
  - not simply a question of adding the result of text content analysis to the result of image content analysis!

- **The way in which the image and the text** relate **to one another seems to be crucial…**
  - in terms of relative importance
  - and, in terms of how they function to convey meaning

➔ A computational framework to ***classify image-text relations…***

Classifying Image-Text Relations

With regards to an image and a text in combination:

- How can we tell which is more important for successful communication?
- What correspondence is there between the information conveyed by one and by the other?
- What information, or other value, does one add to the other?
- If we understand the content of one, then what can we infer about the content of the other?
- What conventions are there for combining images and texts in particular genres of communication?

Proposed Classification Scheme

- In our classification of image-text relations we distinguish two kinds of relations that we take to be mutually independent.
  - **Status relations** are to do with the relative importance of the text and the image, or the dependence of one on the other.
  - **Logico-semantic** relations are to do with the functions that images and texts serve for one another.
  - NB. Different relations may hold between different parts of images and texts
- Based on Barthes (1977) and Halliday (1994)

Status Relations

- The relation between an image and a text is *equal* when:
  - both the image and the text are required for successful communication, in which case they are *equal-complementary;* ***OR***
  - both the image and the text can be understood individually, in which case they are *equal-independent*.
- The relation between an image and a text is *unequal* when either the image or the text can be understood individually - that which cannot be understood individually is *subordinate* to the other.

Logico-Semantic Relations

- A text ***elaborates*** the meaning of an image, and vice versa, by further specifying or describing it
- A text ***extends*** the meaning of an image, and vice versa, by adding new information
- A text ***enhances*** the meaning of an image, and vice versa, by qualifying it with reference to time, place and/or cause-effect

Automatic Classification?

Features of interest to us include:

- **Page layout and formatting**: the relative size and position of the image and the text; font type and size; image border
- **Lexical references in text**: for example, ‘This picture shows…’; ‘See Figure 1’; ‘on the left’; ‘is shown by’
- **Grammatical characteristics of the text**: tense – past / present; quantification – single / many; full sentences or short phrases
- **Modality of images**: a scale from realistic to abstract, or from photographic to graphic – a function of depth, colour saturation, colour differentiation, colour modulation, contextualisation, pictorial detail, illumination and degree of brightness – may correlate with use of GIF / JPEG
- **Framing of images**: for example, one centred subject, or no particular subject

Applying Image-Text Relations?

- Cross-modal Information Retrieval…
- Hypermedia Systems…
- Multimedia Generation…

NEXT STEPS

- Refine classification scheme, especially for image-text combinations with diagrams
- Build image-text corpora to train a classification system
- Evaluate image-text relations in multimedia applications
- Moving images?

Closing Remarks

Interactivity in new media is (and will be) **facilitated** (and constrained) **by machine-executable representations of media items**

Acknowledgements

- Parts of this work have been carried out with:
  - Chris Frehen – Tate analysis
  - Mike Graham, Yan Xu, Elia Tomadaki, Andrew Vassiliou, Elizabeth Jones - TIWO
  - Radan Martinec – Image-Text Relations
- Television in Words (TIWO) – Engineering and Physical Sciences Research Council, GR/R67194/01

Publications

- Salway and Graham (2003), 'Extracting Information about Emotions in Films'. *Procs. 11th ACM Conference on Multimedia 2003*, 4th-6th Nov. 2003, pp. 299-302. ISBN 1-58113-722-2.
- Salway, Graham, Tomadaki and Xu (2003), 'Linking Video and Text via Representations of Narrative', *AAAI Spring Symposium on Intelligent Multimedia Knowledge Management*, Palo Alto, 24-26 March 2003.
- Salway and Tomadaki (2002), 'Temporal Information in Collateral Texts for Indexing Moving Images.' *Proceedings of LREC 2002 Workshop on Annotation Standards for Temporal Information in Natural Language*, eds. A. Setzer and R. Gaizauskas, pp. 36-43.
- Salway and Frehen (2002), 'Words for Pictures: analysing a corpus of art texts.' *Procs. TKE 2002 – Terminology and Knowledge Engineering*.
- Salway and Ahmad (1999), Multimedia Systems and Semiotics: Collateral Texts for Video Annotation. Andrew Salway and Khurshid Ahmad. In: IEE Colloquium Digest, *Multimedia Databases and MPEG-7*, 29 Jan. 1999, London:IEE.
- Salway and Ahmad (1998), Talking Pictures: Indexing and Representing Video with Collateral Texts . Andrew Salway and Khurshid Ahmad. *14th Twente Workshop on Language Technology - Language Technology for Multimedia Information Retrieval*.

Further Information

Publications and other details available from:

www.computing.surrey.ac.uk/personal/pg/A.Salway

Journals and Conferences

For further reading on content technologies:

IEEE Multimedia – a bimonthly journal

ACM Multimedia – annual conference

Websites

- **Information Retrieval:** www.google.com
- **Information Extraction:** www.gate.ac.uk/annie/
- **Image Retrieval - visual similarity**: www.hermitagemuseum.org
- **Information Conversion – image to text:** Dr Lilian Tang, University of Surrey
- **Information Converstion – text to image:** www.semanticlight.com
- **Video Segmentation:**
  www.merl.com/images/proj_images/video-segmentation.gif
  www.cdvp.dcu.ie/manual-news-stories.html
- **Video Retrieval by Visual Similarity:** www.ctr.columbia.edu/videoq/.index.html
- **Video Retrieval with Collateral Text:** Informedia Project, CMU
- **Automatic Video Editing:** Yip, Liu and Howe – ACM Multimedia 2003 Video Program

- MPEG Standards: http://www.chiariglione.org/mpeg/

References

Barthes, R. *Image-Music-Text*. Fontana, London, 1977.

Halliday, M. A. K. *An Introduction to Functional Grammar*. Edward Arnold 2nd edition, London, 1994.

Herman, D. Story Logic: Problems and Possibilities of Narrative. Lincoln: University of Nebraska Press, 2002.

Ortony A., Clore G. L. and Collins A. The Cognitive Structure of Emotions. Cambridge University Press, 1988.

Smoliar, S.W. and L.D. Wilcox. Indexing the Content of Multimedia Documents. In Proceedings: VISual'97; Second International Conference on Visual Information Systems (San Diego, CA), 1997, pp. 53-60.

Computing Moving Images: beyond the pixel

Dr Andrew Salway (a.salway@surrey.ac.uk)
Dept. Computing, University of Surrey

Cross-overs in Audiovisual Arts and Interactive Media

Elomedia, 8 June 2004

UNIVERSITY OF ART AND DESIGN HELSINKI