Hochschule Karlsruhe Technik und Wirtschaft
UNIVERSITY OF APPLIED SCIENCES

Karlsruhe Institute of Technology

Panel

Progress in Description and Retrieval of Complex Multimedia Information

The Fifth International Conferences on Advances in Multimedia

The Fourth International Conference on Models and Ontology-based Design of Protocols, Architectures and Services

April 21 - 26, 2013 - Venice, Italy

Moderator:	**Andreas Schmidt**	**IWI - Karlsruhe University of Applied Sciences,** **IAI - Karlsruhe Institute of Technology, Germany**
Panellists:	**Fabrizio Falchi**	**ISTI - National Research Council (CNR) Pisa, Italy**
	David Newell	**SSRG - Bournemouth and Poole College, UK**
	Andreas Schmidt	**(see above)**
	Mak Sharma	**TEE - Birmingham City University**

Hochschule Karlsruhe Technik und Wirtschaft
UNIVERSITY OF APPLIED SCIENCES

Karlsruhe Institute of Technology

A very short introduction for „non multimedians“ ...

Hochschule Karlsruhe Technik und Wirtschaft
UNIVERSITY OF APPLIED SCIENCES

Karlsruhe Institute of Technology

Overview of a MIR System - Data

Progress in Description and Retrieval of Complex Multimedia Information

- **text**
- **sound**
- **video**
- **images**
- **3D-models**
- **biological data**

combination of above

example: film

- **video information**
- **audio information**
- **subtitles**
- **others**

Hochschule Karlsruhe Technik und Wirtschaft
UNIVERSITY OF APPLIED SCIENCES

Karlsruhe Institute of Technology

Overview of a MIR System

- How can this data be queried ?
- „Typical“ database query:

  Give my all autofocus cameras with at least 5 Megapixels, that cost not more than 200 €.

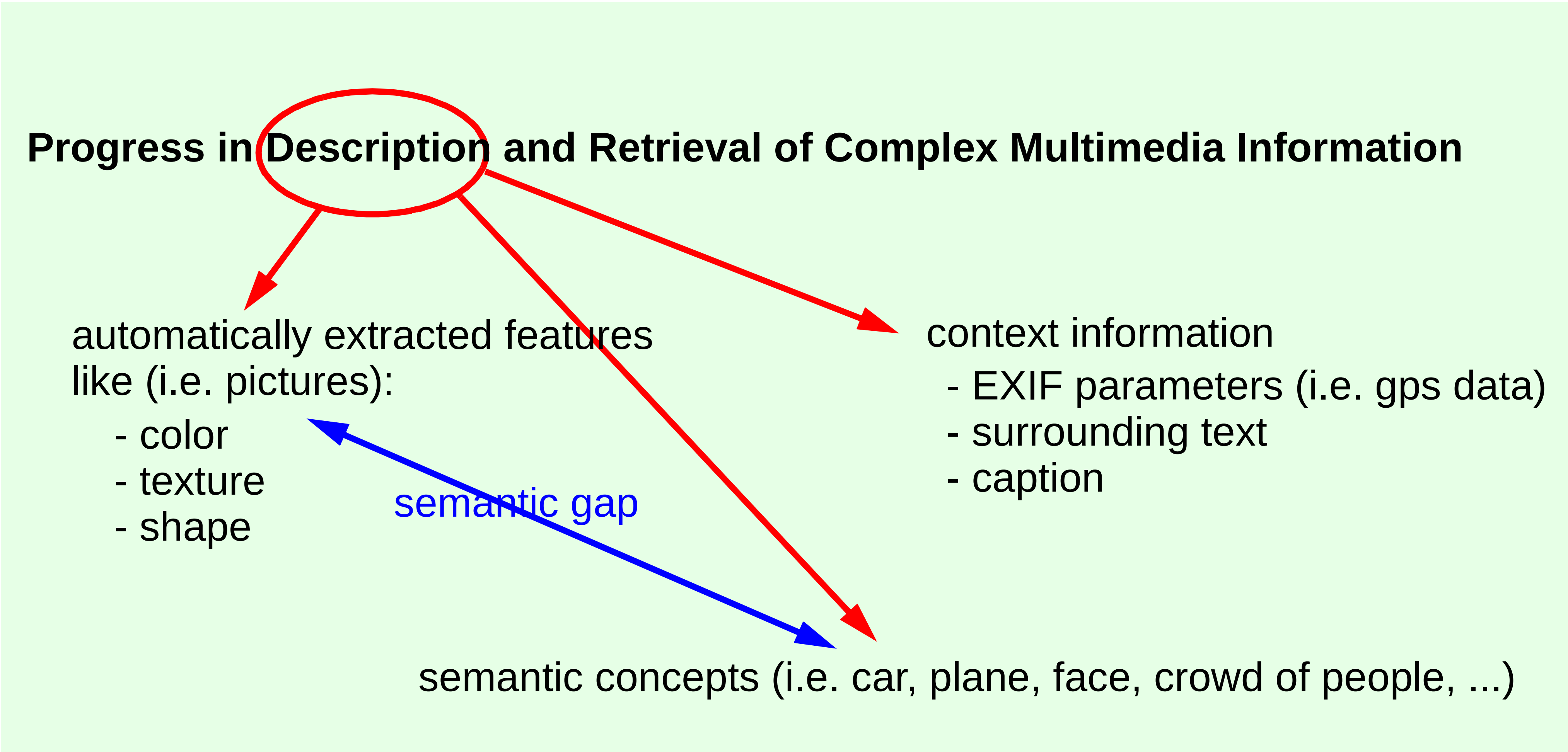
  Easy ! why ?: data is stored at a semantic level, equal to a typical query

- Problem: Multimedia raw data is stored at a very low semantical level (typically compressed)
- To query multimedia data we need an appropriate description of the data
- Content description for a multimedia object can be extracted or manually added

Hochschule Karlsruhe Technik und Wirtschaft
UNIVERSITY OF APPLIED SCIENCES

Karlsruhe Institute of Technology

Overview of a MIR System - Description



Overview of a MIR System - Retrieval

Progress in Description and Retrieval of Complex Multimedia Information

Conventional queries:

- Point query
- Analytical query
- range query (without ranking)

Multimedia query:

- often (k-) nearest neighbor
- query is interpreted as (virtual) multimedia object
- ranking of results
- definition of similarity measure

Hochschule Karlsruhe Technik und Wirtschaft
UNIVERSITY OF APPLIED SCIENCES

Karlsruhe Institute of Technology

Overview of a MIR System - Retrieval

Progress in Description and Retrieval of Complex Multimedia Information

Query By Example (QBE)
Query adaptive search strategies
Interactive retrieval
Personalized search
Browsing

Overview of a MIR System - Summary

- Multimedia objects cannot be queried directly
- Appropriate descriptors have to be found
  - Low level
  - contextual
  - semantic level

Semantic gap

- Challenges [1]:
  - data representation
  - similarity measure
  - scalability
- Very interdisciplinary work:
  - signal processing, pattern recognition,machine learning, information retrieval, NLP, HCI, psychology, computional vision, optimization theory, ...

Sources

[1] Eduard Y. Chang: Foundations of Large-Scale Multimedia Information Management and Retrieval: Mathematics of Perception. Springer, 2011

[2] A. Hanjalic, R. Lienhart, W. Y. Ma, J. R. Smith: The Holy Grail of Multimedia Information Retrieval: So Close or Yet So Far Away? Proceedings of the IEEE, Vol. 96, No. 4., 2008

Aspect Retrieval Performance ...

Aspect: Retrieval Performance

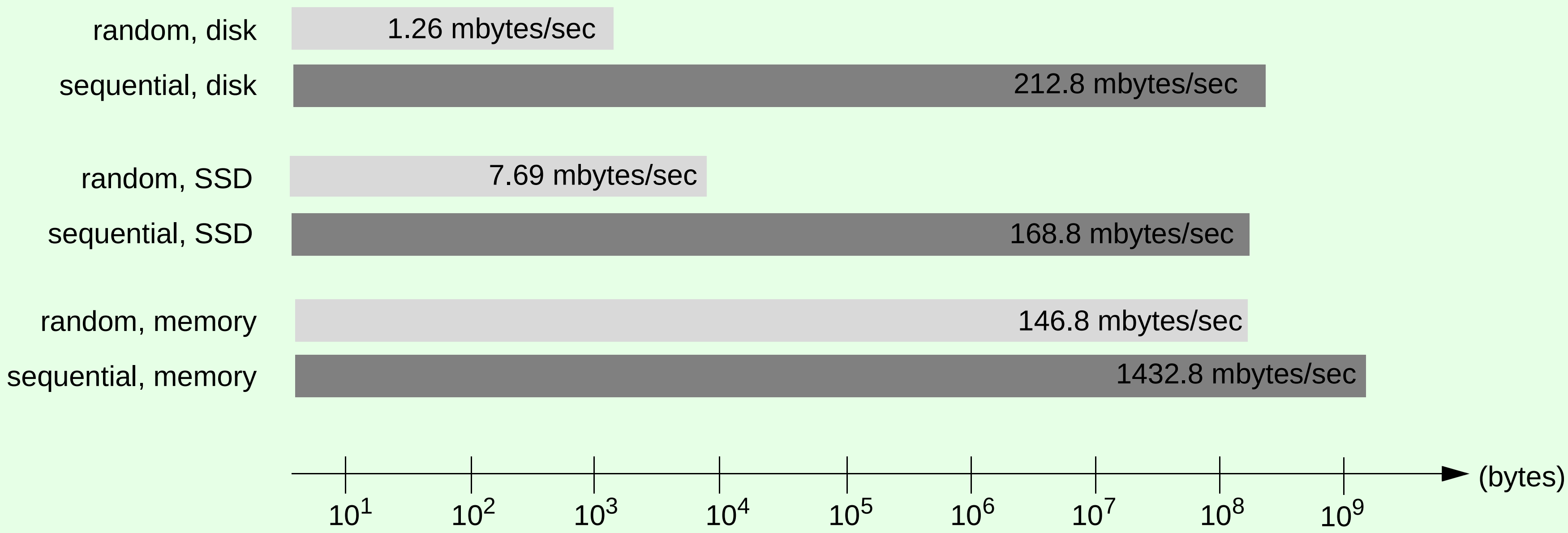
- Curse of Dimensionality
  - Dimension > 20: no k-nearest neighbor (k-nn) algorithm is significantly faster then a linear scan [1]
  - For high dimensional range queries, traditional multidemensional indexes are slower than a linear scan [2]
  - Why ? - Exponential grow of minimal bounding regions (MBR)
- Solutions:
  - Reduction of dimensionality (i.e. PCA[1], ICA[2])
  - Appropriate memory layout:
    - Approximate nearest neighbor search algorithms
    - Alternate index structure: Binned, compressed bitmap vector

---

1. Principal component analysis
2. Independant component analysis

Aspect: Retrieval Performance

Comparision of random and sequential memory access



Source: Adam Jacobs: The Pathologies of Big Data, acmqueue, July 2009

Aspect: Retrieval Performance

Approximate k-nn Algorithm *SphereDex*:

- Points are partitioned based on distance to a central point
- Make IO access sequential
- Intra partition index for effecive pruning
- Partition fits in L2 cache
- Inter partition index for finding good starting point in neighbor partition
- Neighboring partitions are sequentially stored on disk
- About one order of magnitude faster than M-Tree, LSH

Source:
Edward Y. Chang, Approximate High-Dimensional Indexing with Kernel. In Foundations of Large-Scale Multimedia Information Management and Retrieval - Mathematics of Perception, Edward Y. Chang (editor), Springer, 2011.

Aspect: Retrieval Performance

Fastbit Bitmap Index [5, 6]:

- based on WAH algorithm
- Multidimensional queries are mapped on AND/OR operations of (compressed) bitmaps
- significantly reduction of I/O
- pure sequential IO
- Binning to reduce number of bitmaps
- Further improvement with Order-preserving Bin-based Clustering (OrBiC) structure [4] (3-25 times faster than with best known multidimensional data structures)

Source:
Kesheng Wu, Kurt Stockinger, Arie Shoshani: Breaking the Curse of Cardinality on Bitmap Indexes. Lecture Notes in Computer Science 5069 Springer 2008, pages 348-365

Aspect: Retrieval Performance

Quintessence

- Sequential disk access is magnitudes faster than random access
- Appropriate memory layout can speed up the application significantly
- Approximate k-nearest neighbor is sufficiant in most cases
  - feature vector is also an approximation
  - results are inspected by humans

- Bitmaps: Binnig and run length encoding is an appropriate combination to resolve the *curse of dimensionality*

Disadvantage of the proposed solutions: For read mostly applications only (append only)

Sources

[1] P. Indyk, R. Motwani, Approximate nearest neighbors: towards removing the curse of dimensionality, in Proceedings of VLDB, 1998, pp. 604–613

[2] K. Beyer, J. Goldstein, R. Ramakrishnan, U. Shaft, When is x “nearest neighbor” meaningful? in Proceedings of ICDT, 1999, pp. 217–235

[3] Adam Jacobs: The Pathologies of Big Data, acmqueue, July 2009

[4] Kesheng Wu, Kurt Stockinger, Arie Shoshani: Breaking the Curse of Cardinality on Bitmap Indexes. Lecture Notes in Computer Science 5069 Springer 2008, pages 348-365

[5] Kurt Stockinger, Kesheng Wu, and Arie Shoshani: Evaluation Strategies for Bitmap Indices with Binning, (DEXA 2004), Lecture Notes in Computer Science 3180 Springer 2004.

[6] D Rotem, K Stockinger, Wu: Efficient binning for bitmap indices on high cardinality attributes, 2004

[7] Edward Y. Chang, Approximate High-Dimensional Indexing with Kernel. In Foundations of Large-Scale Multimedia Information Management and Retrieval - Mathematics of Perception, Edward Y. Chang (editor), Springer, 2011.

Digital Audiovisual Media Preservation

Fabrizio Falchi
fabrizio.falchi@cnr.it

Audiovisuals Digital Preservation

- **changes** in the hardware, software, and physical media used to process and store digital data will continue into the indefinite future

- Most AV files are **compressed**. Whatever 'original quality' was lost in compression will remain lost. Preservation should maximise retention of quality, a capability that needs to be defined and added to current technology.

Richard Wright, BBC

- we need **preservation metadata** in any processes that impact the communication of the digital information over time and ensure the generation of additional preservation metadata that describe the processing and its results

MP-AF Draft

Presto4U

FP7 Coordination Action

The aim of the project is to:

identify useful results of research into digital audiovisual preservation and to raise awareness and improve the adoption of these both by technology and service providers as well as media owners.

Presto4U is building **Communities of Practices** of audiovisual preservation related organizations.

We are also searching for **audiovisual for science** producers and archives to understand their preservation needs.

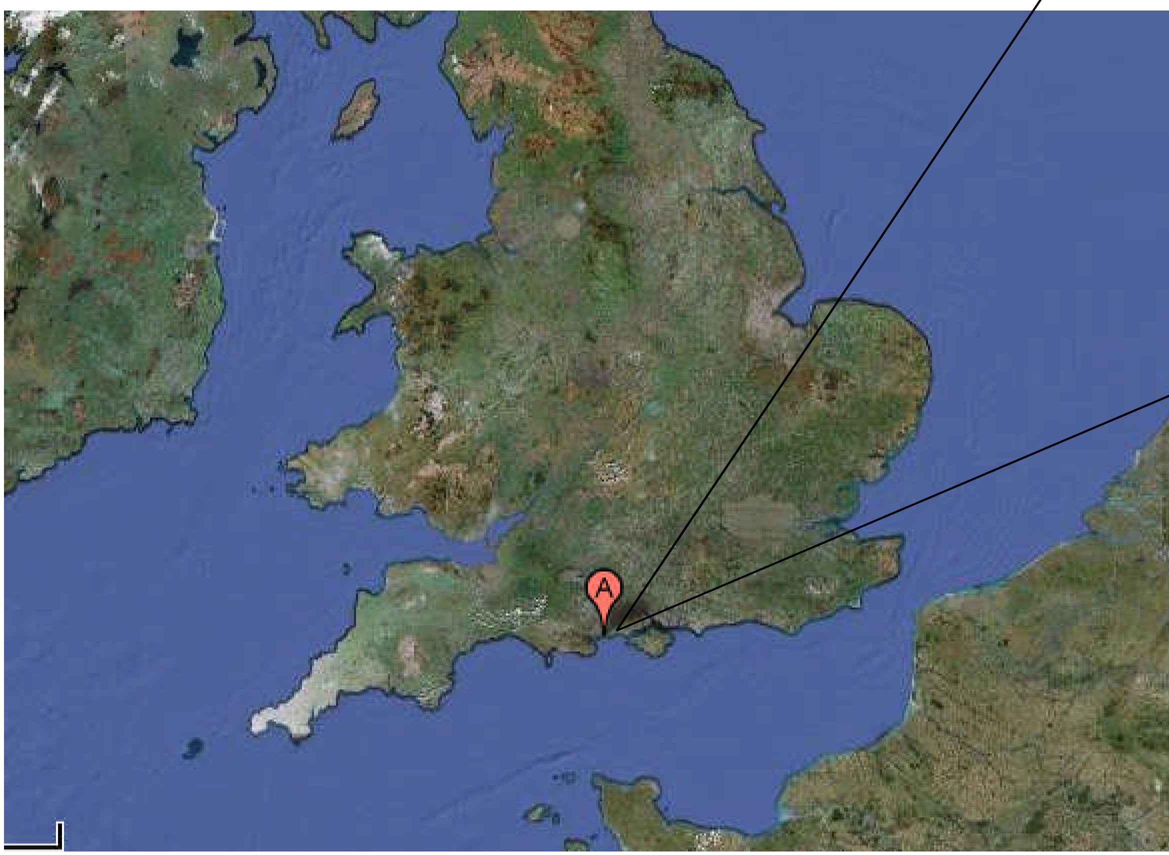
MPEG

- **Proposed Data Model and Representation for Multimedia Preservation Application Format (MP-AF)**
  - specifies the lossless representation of metadata relevant to preservation of multimedia information for exchange of multimedia content between multimedia archives
  - is specified for exchange multimedia contents among multimedia archives, avoiding the loss of metadata for preservation

Fabrizio Falchi
ISTI-CNR, Pisa
fabrizio.falchi@cnr.it

Bournemouth University

Dr Philip Davies
David Newell



MMEDIA 2013
Venice, Italy

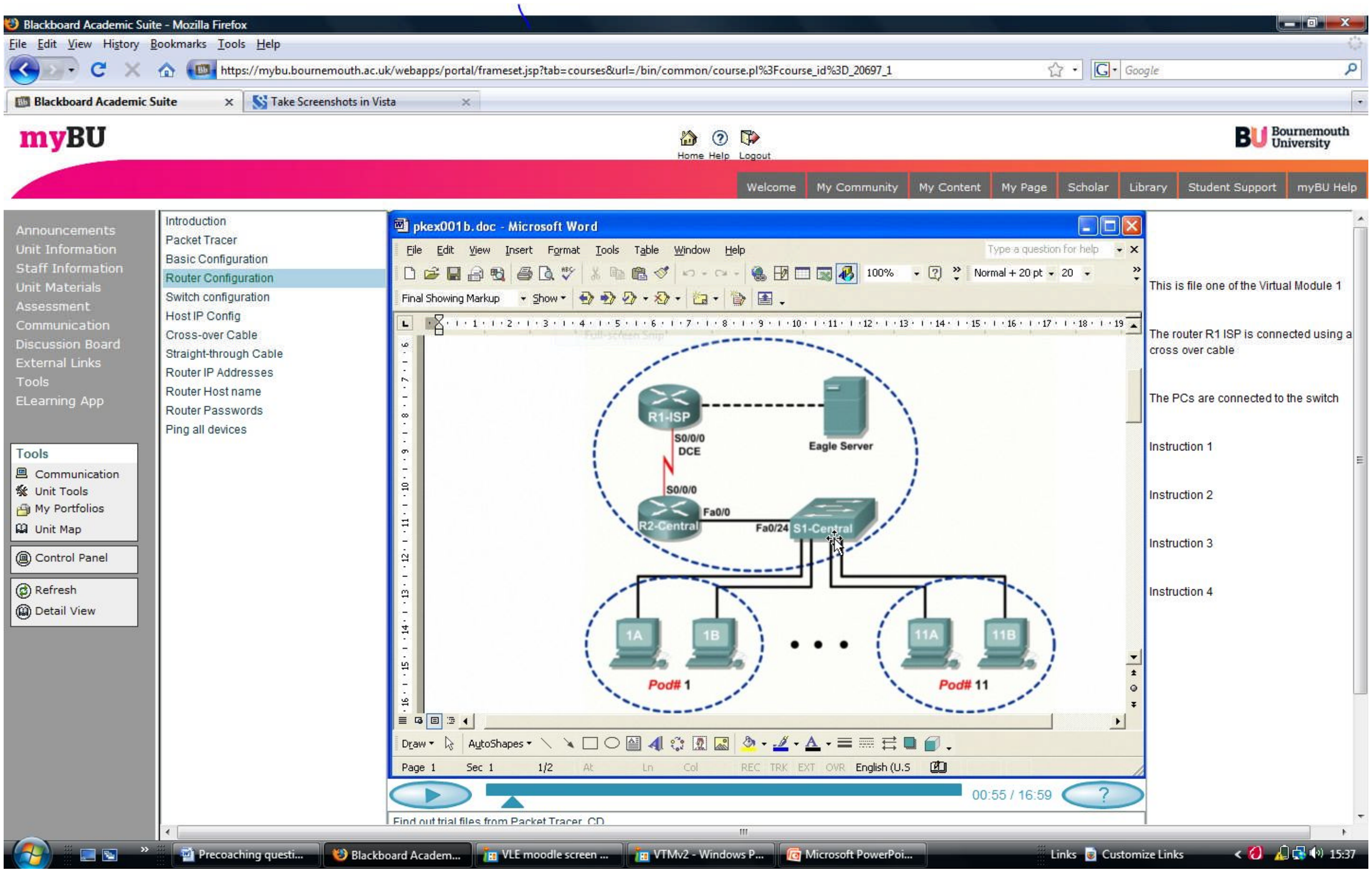
Query Methods Research for Image Databases is Widespread

QBIC (Query by Image Content) content-based retrieval method

- QBIC queries large image and video databases using:
  - example images, sketches and drawings,
  - selected colour and texture patterns, camera and object motion and other graphical information
- (1) use of image and video content-computable properties of colour, texture, shape and motion of images, videos and their objects in the queries
- (2) Use of graphical query language, in which queries are posed by drawing and selecting user-constructed graphics

Source: QBIC technology is part of several IBM products

Adaptive MM Interface (AMPS)



Towards Ontology Calculus

Query method based on 'Semantic Descriptor'

QBSD (Query by Semantic Descriptor) learning object content-based retrieval method

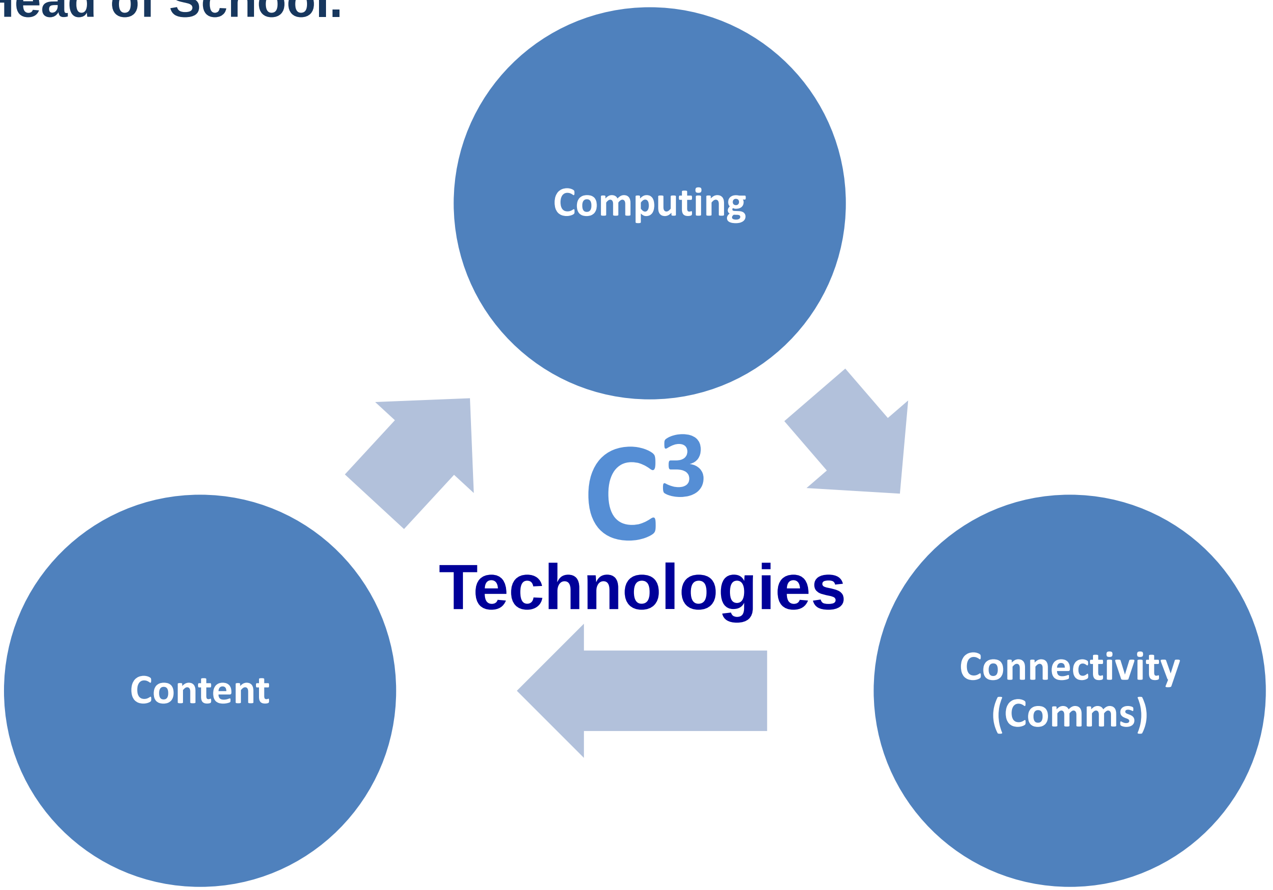
- QBSD queries large learning object video databases using:
  - Ontology meta-level description of curriculum, learning objects, fragments, user-constructed questions, and FAQs
  - selected lessons, results of MPQ tests, and other learning or evaluation meta-data used to adaptively retrieve learning objects
- (1) use of text, image and video content meta properties - difficulty, level, content type, combine responses to questions, produce supplementary videos and learning objects from queries
- (2) Use of an adaptive semantic description is a form of query language, from which queries are processed, and used to display fragments of animated graphics, images and text

Source: AMPS technology is part of several BU prototypes disseminated at MMEDIA2009-2013.

CTN

…together we can succeed and make a difference …

Mak Sharma, BEng, PgC, PgCE, FHEA, MIET, MBCS.
Head of School.



Progress in Description and Retrieval of Complex Multimedia Information

What are we Retrieving ?

How do we learn KAV ?

How do we remember ?

Throw technology at it

How can we do ?

What will it cost ?

What are the casualties ?

YouTube

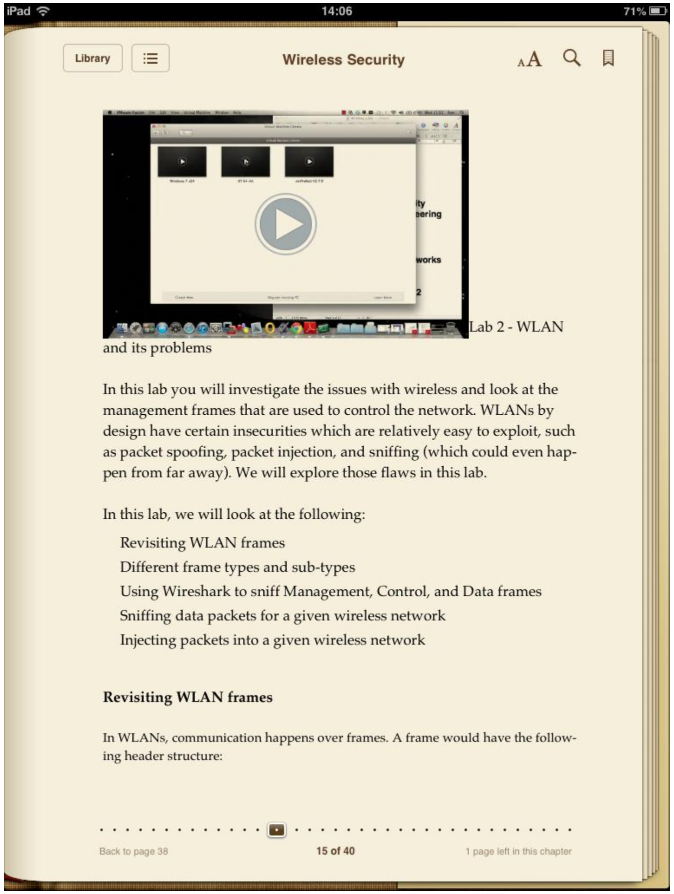
- Over 60 hours of video are uploaded every minute, or one hour of video is uploaded to YouTube every second.
- Over 4 billion videos are viewed a day
- Over 800 million unique users visit YouTube each month
- Over 3 billion hours of video are watched each month on YouTube
- More video is uploaded to YouTube in one month than the 3 major US networks created in 60+ years
- YouTube is localized in 39 countries and across 54 languages
- 1+ trillion views
- 140+ views for every person on Earth

- 500 years of YouTube video are watched every day on **Facebook**
- Etc Etc ….

Read more at http://www.jeffbullas.com/2012/05/23/35-mind-numbing-youtube-facts-figures-and-statistics-infographic/#ESOLPFtylPS5KidR.99

A better way of learning ?

- Developed from adding video clips to moodle
- More defined aims and objective within each module via BCU+ gives focused delivery of content
- Focuses students to the task



Video everything !!!!

What are considering ?

HP Cloud

Amazon Cloud

Microsoft Cloud

Cisco Servers - Cloud solutions
& define your own optimised network SDN

What is reality or are we in the Cloud ? £

Any Questions ?

Mak Sharma, BEng, PgC, PgCE, FHEA, MIET, MBCS.
Head of School.