Overview of approaches to Multimedia Information Retrieval

Lecture 10
Example Queries:

- Find images in which Don Quixote attacks the giants
- Find examples of Don Quixote attacking the giants
- Find a video clip of a train crossing a bridge over a valley
- Find the song which goes (hum the tune here...)
- Find an image of a scene of sunset in a horseshoe bay which has two palm trees growing in an X-shape on the left-hand side.
- Find examples of Captain Kirk saying "Beam me up, Scotty!"
Problems

- Ubiquitous MIR very hard (at the moment!)
- Scalability: high volume of multimedia data compared to textual document
- Semantics: already very hard in text collections
- Versatility: usually, solutions are domain and medium specific
Approaches

• Query-by-Example
• Query-by-Content
• Description Logics
Query-by-Example

- Use low-level syntactic properties of media to locate "similar" multimedia documents

- Statistical features
  Image (colour frequency, texture)
  Video (image features, variation over time)
  Audio (loudness, pitch, brightness [higher frequency content of signal])

- Structural information
  Image (spatial colour, contour)
  Video (shot length, shot type)
  Audio (bandwidth [variation of frequencies], harmonicity [comparison of spectra to white noise])
QBE Advantages

• Pattern recognition and matching
• Supports similarity-based retrieval
• Can classify documents (or rather, document segments)
• Can train QBE systems to auto-recognise/classify new material
• Don't need to mark-up with text

QBE disadvantages

• Not expressive enough
• Exceptions hard to classify/quantify
• Requires the user to know composition of relevant documents
• Can't search using a textual description
Query-by-Content

- Based on QBE, but with extensions for shape and motion.