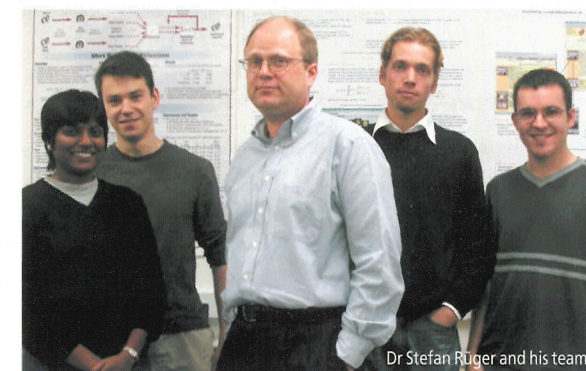


Allegretto

By Michael Boyd

Piano



Dr Stefan Rieger and his team.

Data's Rising Tide

At work, in our homes – even while we sleep – we are responsible for creating oceans of data. Yet navigating our way around this virtual sea, and especially finding what we want, is becoming ever more difficult. Could automated systems take the strain?

In many ways simply keeping abreast of the contemporary explosion of data can be compared with the unenviable task of keeping up with the proverbial Joneses. If, as often with such families, they appear to have access to unlimited funds it can be a difficult and sometimes soul-destroying task as many have found to their cost.

However, despite the truth of this adage there is at least one person who not only remains stoically optimistic but is also very determined to find ways of giving aspiring data managers a helping hand. He is Dr Stefan Rieger at Imperial College's Department of Computing. Publicly signalling that it confidently

shares his hopes is EPSRC that has given Dr Rieger an Advanced Research Fellowship to develop more efficient procedures for retrieving information, images and audio that will be tested using realistic examples. "Managing multimedia content is still an unsolved problem considering the ever increasing amount of multimedia data," he observes. Examples range from medical images for diagnosis, objects in art galleries and museums, satellite images for weather forecast and surveillance, CCTV, TV, radio, music and movie productions, journalistic archives, technical construction drawings, and digitised cultural-heritage documents. "And these are just examples of multimedia data which are used professionally," he added, "at home we have now an ever-increasing amount of personal videos and images."

The multimedia maze

As one might expect managing this data is beset with numerous problems, it is almost a question of knowing where to begin: "Current best practice for managing multimedia data is still manual annotation, such as typically the creation of a small 'library card' with a text description of the image, video or music piece," explains Dr Rieger. To try and cope commercial picture libraries often have a dedicated department using a controlled vocabulary of some 30,000 specific terms to create and deliver the annotations. Text search engines or standard database queries can then be used to find matching library cards and, hence, corresponding multimedia objects. Although this approach may be commercially viable, and even desirable, often this cannot be done. This is simply because of the prohibitive cost of funding something like the manual annotation of CCTV footage with a view to be able to

pose database queries like "find me footage of a person with black clothes jumping into a blue van and driving off at high speed". This has to be in the domain of automated, content-based analysis, says Dr Rieger.

There exist limited automated, but non-content-based methods such as Internet image search engines, like Google, that use file names and text snippets of image captions and anchor texts for image searches. Similarly, the subtitles of films or the text gained by speech recognition from radio broadcasts or audio tracks of movies can be automatically turned into library cards. However, finding video shots showing an explosion may not have the word 'explosion' spoken nor have this particular description in the subtitles.

One of the most powerful paradigms of content-based retrieval is 'searching by example' be it by relevant example images, by sketch input or by humming a tune. Content-based retrieval has thus become an active research field in recent years to overcome the limitations of traditional methods. With this aim Dr Rieger eventually hopes to help physicians

"Current best practice for managing multimedia data is still manual annotation, such as typically the creation of a small 'library card'" Dr Stefan Rieger

in their diagnosis through similar cases from a large knowledge base, support prosecution services with fingerprint matching, and identifying copyright violation or illicit

pornography on the web, increase public security by identifying suspicious movements, say, in car parks, access recorded TV programmes by content as opposed to time and channel, find the use of similar trademarks and content-based image access that is truly language independent without deploying cross-lingual retrieval for text in captions.

Four years into his five-year Fellowship Dr Rieger and his team have contributed to this research field with the development, analysis and evaluation of multimedia features for retrieval and similarity computation. These include automatically extracted, low-level properties of the multimedia objects such as pitch and rhythm from music fragments, shapes in sketches, movement elements in videos or colour distributions, structural and texture characteristics of still images. Other research questions tackled are how best to combine evidence for similarity, given partial

matches of different features. There appear to be three inherent difficulties involved; firstly, modelling human similarity perception, then disambiguation of the different possible reasons why a searcher might be interested in a particular example and finally bridging the semantic gap between low-level features (colour, shape, texture) and more abstract human concepts such as 'barbecues' or an 'uplifting mood'.

Dr Rieger has addressed these difficulties by incorporating, and making his system learn from, user feedback and responses as well as modelling context information from observed data. As an example, the system would learn barbecues appear in the context of people, bottles, grass and sky, which in turn can be found by colour, shape and texture distributions. As a consequence the deployed methods involve machine learning, artificial intelligence and statistics.

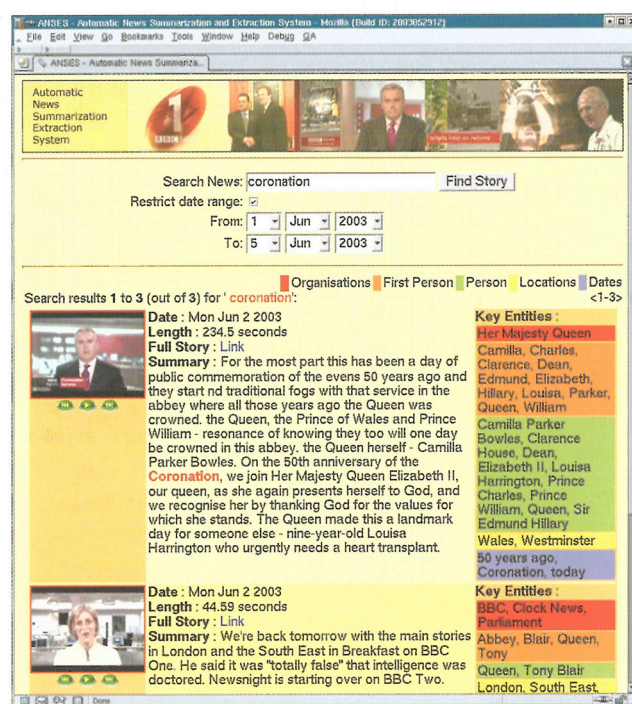
Search and evaluate

The main and overriding criterion of success, however, is that the developed prototype systems are not only user-friendly but also prove effective and efficient in thorough metrics-based evaluation. To this end the group have so far developed a prototype TV news search and summarising service, music, image and video search engines covering a huge amount of data and undergoing a constant cycle of external, international and competitive evaluation.

At the end of the Fellowship there will be a general multimedia management demonstrator incorporating novel browsing, searching, summarising and visualisation paradigms that have been tested and evaluated using large corpora. This will be the starting point for concrete research and development addressing specific application areas. One of the most promising areas appears to be the fusion of multimedia management techniques with traditional digital libraries. Dr Rieger comments: "Getting this Fellowship meant to me the freedom and flexibility to think outside specific short-term projects which is instrumental to both carrying out blue-sky research but also turning the results into useful products." ■

Contacts

For more on this project e-mail Dr Stefan Rieger at s.rieger@ic.ac.uk



Researchers at Imperial have developed a prototype TV news search and summarising service incorporating music, image and video search engines.