Images in Focus: organizing visual content for use

ISKO, London, 28 November 2012

"Better value than some courses that cost ten times as much" said the person sitting next to me, and I think the majority of people in the room would have concurred: well worth struggling through floods and cold winter weather to get to this meeting. Condensed into half a day was a remarkable range of approaches to image metadata, ranging from a traditional commercial picture agency, to the latest developments in automatic generation of metadata at the BBC and the University of Bristol. For me, the best aspect of the whole event was that the speakers were prepared to describe in detail how they confronted the challenges before them - not always with success, but revealing what didn't work as well as what did.

There were stories of major projects, success and failure stories, and leading-edge research. It's not possible to give a full appraisal of the six sessions in this brief review, but a few points stand out:

Tom Gilmour of Mary Evans Picture Library described how keywords are chosen and input when a picture is made available on their online system. The process of adding tags looked admirably clear, and the need to be aware of what users were searching for was taken into account, but I couldn't help feeling that the metadat to be added will be influenced by the capabilities of modern search tools available from current retrieval systems. For example, many (although admittedly not all) retrieval systems will carry out automatic stemming of search terms, and in some cases will search for equivalent terms, so a search for "African" will often be converted to a search for "Africa" - there is no need to add both terms to the object metadata.

Sarah Saunders gave an enthusiastic talk about the development of image metadata, from the point of view of picture agencies, and an enthusiastic plea for the increasing adoption of embedded metadata (the kind of information that holds details of your camera for every picture you take). Her analogy was that of the sweet shop: the Web presents images as if you are in a sweet shop, and the immediate temptation is to grab the ones you want and run off and reuse them. Her definition of an orphan work as "an image that cannot legally be used at the moment" was a powerful one, but one that would be difficult for many publishers to adhere to in practice. It would have been good to take into account some of the many challenges being mounted to copyright legislation that are currently underway, and it was a little unfortunate that none of the images she used in her presentation had any attribution.

Gabriele Popp & Stephen McConnachie described the British Film Institute's large-scale project to integrate searching across all the different media types they catalogue, including not only film and TV but scripts, posters, interviews, and many other media types. The result is a "Collections Information Database" that will soon (2013) be available for public searching on the Web. This was a major project, lasting over three years, integrating searching across 35 separate datasets with no common structure, and the whole project would form a fascinating large-scale case study.

Paul Davies described an attempt at Birmingham library to use automatic image recognition for photographs in the archive - not always with very successful results. It seems clear from his presentation that software that detects shapes such as "person", "car", or "train" needs a lot of help before it can be relied on for automatic tagging.

Another presentation on automatic image recognition was by Tilo Burghardt of the University of Bristol. His team has created tools for distinguishing not just one species, but individuals within a species, for example penguins at Bristol Zoo. This exercise seemed complex enough until he pointed out that the penguin in question may be facing in any direction, necessitating the use of 3D software to identify the individual. The presentation was astonishing; I would love to try the software out in practice to see how it works.

Finally, Sam Davies of the BBC R&D metadata classification team described a way in which automatic tags have been applied to TV programmes on the BBC iPlayer. We were given a URL where the results could be seen, but I couldn't find it at the BBC R&D site. I did see a reference to the project at http://www.bbc.co.uk/blogs/researchanddevelopment/2012/01/an-interface-for-mood-based-navigation.shtml Using this tool, programs have been automatically ranked for seriousness v humour, and for fast-moving v slow-moving. It would be fascinating to see how useful this tool could be.

Michael Upshall