



Visual Search Technology for Publishers

Version: 1.0
Release Date: 29 September 2011

PUBLISHERS

Newspaper and magazine publishers have unique requirements for dynamic content and it is not unusual for a national newspaper publisher to process more than 50,000 new images each day. These are typically presented by picture agencies and commercial image libraries via feeds, with the images then managed by specialised picture desk and digital asset management tools. The workflow involves new images being delivered and indexed before being presented to the picture desk. Pictures arrive with metadata so that after indexing they are immediately searchable.

This metadata also describes usage rights and, in some cases, price. To improve usability publishers will sometimes also add additional captioning to each image. Following this, pictures are then available to be selected for use to address immediate editorial requirements. If any pictures are seen as particularly topical or useful, then they may also be selected and the editorial team notified. Due to the volumes involved, images are not archived unless they are used and so after a period of time (usually 14 days) they are simply discarded.

ISSUES WITH THE TRADITIONAL PICTURE DESK APPROACH

Publishers get the majority of their content from subscription feeds although they will also commission their own photography as well as buy images individually from stock photo providers as needed. Due to the nature of editorial content however, the various feeds will often contain many duplicate images. This is either because a photographer has an arrangement with more than one vendor, so images appear more than once; or that simply the same images from different sources are presented with slightly different (or just incorrect) tags and so are indexed under different categories. As a result of the limitations of text based search engines, sometimes a required image may not be found by traditional search, even if the image is known to be in the cache or the archive. This, in turn, can cause unnecessary spend as frustrated users go directly to commercial image or agency web sites to purchase the same image.

WHAT IS NOW POSSIBLE WITH IMAGE RECOGNITION TECHNOLOGY?

Visual search technology can be used to not only provide another level of search filtering but also to improve how new images are managed by rapidly identifying duplicates.

Image similarity search is designed to work alongside and complement existing text based search engines and provides users with the ability to search using a reference image for other visually similar images, even offering the ability to do this by refining things such as the image colour, or whether there is a face in the image or not. This means that visually similar images regardless of their metadata can always be found, thus improving the user experience and reducing the frustration of not being able to find the image they want. An extensible solution, such as Imprezzeo, also provides the ability to automate this process and so offer an assisted search experience. This means that whilst navigating content, similar images can be suggested automatically to the user, prioritizing, of course, those that are royalty free.

The technology can also be used at the front end of the image workflow. New images that are delivered via feeds can be identified as duplicates by comparing them in near real time to the existing cache or archive. From this, rules can be set up to handle the duplicates in the most appropriate way for the business. For example, where there might be three copies of the same image, the copy with the most beneficial rights can be passed through to the users; the remainder can simply be ignored. This is especially relevant where images are already owned and available for use in the archive but using the newly arrived and duplicate images has cost implications. Thresholds for close matches can also be set, so very similar but, say, royalty free images from the archive could also be passed through.

The ability to identify duplicate and close matches can also be used to address substantive (multi million image) archives and so provide publishers with the ability to improve the quality of their archives.

Finally, for those publishers concerned with compliance, image technology can also provide a way to audit whether the images published by them are being appropriately used in accordance with the usage rights and that the

Visual Search Technology for Publishers

appropriate fee has been paid to the appropriate vendor. This is done by comparing the published content with vendors' libraries using the same similarity search technology.

BUSINESS BENEFITS

Similarity Search:

- Reduce wasted time searching for images;
- Improve user experience, by making it easier to find the desired image;
- Find 'lost' images and so avoid unnecessary spend;
- Identifying Duplicates;
- Identify images in near real time that are already 'owned' thus avoiding unnecessary spend;
- Remove duplicates, therefore improving IT utilisation i.e. storage, processing;
- Usage compliance;
- Reduced risk and reduced cost of non-compliance.

ABOUT

Imprezzeo is a visual search technology company. Imprezzeo Image Suite provides an enterprise level solution that complements existing search systems, making it easier and faster for users to find the images they need. Our patented technology provides developer tools to enable rapid integration of visual similarity search and facial recognition functionality into existing search and content management systems.

www.imprezzo.com