

# défense nationale

études politiques - stratégiques - militaires  
économiques - scientifiques

**EUROSATORY**

## Europe et armement terrestre

Serge VINÇON, François LUREAU, Luc VIGNERON,  
Christian SCHMIDT, Gérard BEZACIER,  
Jean-Paul PERRUICHE, Pierre-Jean LASSALLE,  
Antonin TISSERON, Darko RIBNIKAR, François VELTZ

## 60<sup>e</sup> ANNIVERSAIRE DU DÉBARQUEMENT

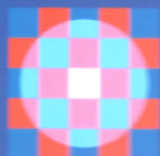
René de NAUROIS, Alexandre ORLOV,  
Louis DURTESTE, Alain DUMONTET

## L'ONU à la veille de nouveaux défis

Pascal CHAIGNEAU



revue mensuelle  
juin 2004  
10 €



# ***Image mining* :** **accelerated visual analysis**

Search, filter and categorize images

François VELTZ

As potential investigative tools, law enforcement and intelligence agencies have long been interested in using image processing and computer vision techniques (e.g. remote sensing, biometrics). For over a decade, the exponential growth of digital imaging and the *Internet* have led to a new series of phenomena: image-based investigations, image-guided computer forensics, and the monitoring of sensitive images.

LTU Technologies is a company that seems to have tackled this problem successfully. Several law enforcement and intelligence agencies in the US and worldwide are using its *image mining* technology in applications ranging from child exploitation investigations to combating counterfeiting and tracking stolen art.

This technology is not aiming at replacing the human being. To the contrary, it is a powerful tool at the service of investigators, who can let the software find image correlations and dedicate themselves to higher-level decisions and actions.

Given the importance of images and videos in today's insecure world, the perspectives of such a technology for homeland security are impressive.

In just a few years a French technology has gained momentum - because of its outstanding results - at the service of great American investigation agencies. Interoperability required for the co-operation of international institutions has led to the selection of this technology as a standard for image based investigations.

## **EXPONENTIAL GROWTH IN IMAGE PRODUCTION**

The quantity of images produced in the world has grown much faster than have methods for managing and using these images. And the trend continues to accelerate.

All sectors of human activity are concerned. Worldwide more than 100 billion photographs are taken annually, and these are increasingly in digital format. Mobile photo-phones surpassed sales projections and are becoming more and more miniaturized. Today, more images are generated by these hybrid telephone/camera devices than by any other means of digital or film photography. Each of the ten billion pages the Web contains an average of four images and Peer-to-Peer networks makes the exchanges explode. Industrial and scientific sectors also produce great images quantities. For example, pharmaceutical research generates huge quantities of photographs during combinative and robotized tests on cell cultures. Relying only on the human eye, it is impossible to examine and classify all the results. The results of simulation, technical diagrams and patent databases similarly product massive quantities of images.

Thus, a since a great deal of human knowledge is now encoded graphically, it becomes increasingly important to have advanced tools for reading and accessing this knowledge. Recourse to the techniques of computer vision, also known as artificial vision, is mandatory if one wishes to control, filter, or exploit the information contained in these images.

## **FROM IMAGE DNA TO SEMANTICS**

Giving to the machine the ability to interpret images - to find correlations between significant images or to locate subtle graphic elements - – this has been a long term challenge for scientists. Starting with fundamental research during the 1970's, the following decade focused on industrial applications of 3D perception, such as for robotics applications in manufacturing. The 1990s marked the return of interest in two-dimensional analysis, with ultimate goal of automatic content recognition of images. Today, years of high-level academic research has lead to the development and marketing of mature technologies in this field.

\* Peer-to-peer exchanges, without servers. See Napster, Kazaa, eMule and BitTorrent.

## « Image DNA »



Images are complex to analyze, since the visual elements of an image are not arranged in a conventional way, such as letters or words are arranged in the conventions of a language. The constitution of a dictionary or a thesaurus of visual data is thus quite impossible. Current technologies of "computer vision" are often inspired by human vision. The two thousand or so channels of the optic nerve do indeed send raw information to the brain, which integrates it and thus "sees" it, i.e. interprets signals into concepts.

Borrowing elegantly from biological terminology, LTU Technologies defines the concept of an "Image DNA" which defines the photometric properties (color, brightness) and geometric properties (the shapes of the objects).

For each image, the system thus extracts a single identifier, also called a digital signature. This DNA represents the image as a compact vector of numerical values, which effectively codes all information relating to the visual content. More than two thousand factors are taken into account. The Image DNA is a very light file (approximately 1Kb), robust with the traditional variations of images (resolution, contrast, rotation, distortion, cropping, change of format) and optimized for retrieval from very large databases.

Once the Image DNA is calculated, relationships with other images may be established through DNA comparison. The images may be retrieved, filtered and categorized by relative comparison of their DNA. The computer can thus truly interpret one image by its visual contents and no longer exclusively by the digital building blocks of an image (pixels). This Image DNA calculation makes it possible to find cloned images or even similar images, the latter a task being impossible with standard MD-5 or SHA-1 algorithms (any change to an original file would result in a change to the file's digital fingerprint, as calculated by these standard algorithms.)

Duplicate, cloned and similar images

	<b>Duplicate</b>	<b>binary copies (strictly identical images)</b>
	<b>Cloned</b>	<b>very close images (simple transformations like enlargement, cropping, recompression)</b>
	<b>Similar</b>	<b>similar images (semantic level)</b>

## Semantics and Recognition Training

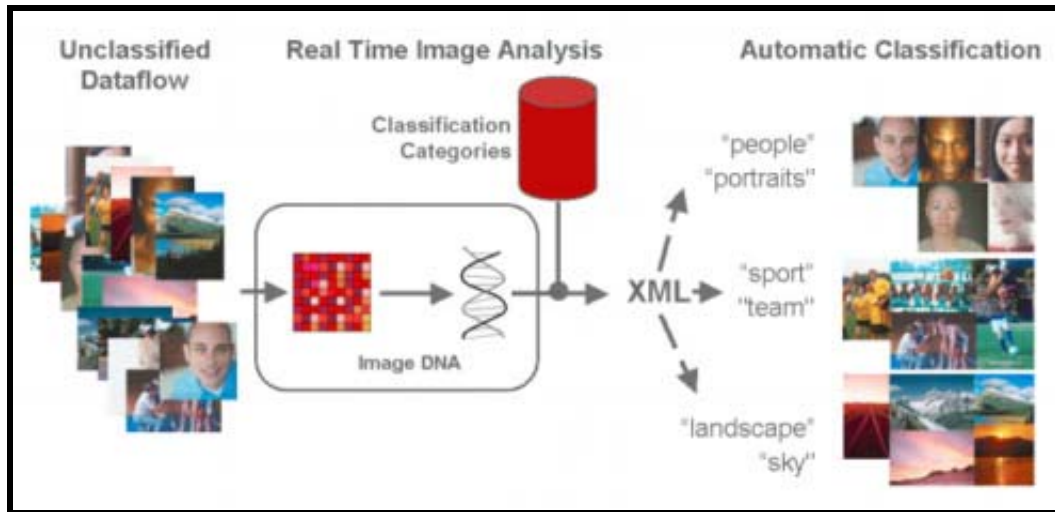
The technology also allows, within a well defined framework, to pass from the machine-calculated image DNA to a human-language or semantic description of image content. In other words, it becomes possible automatically to associate descriptive keywords to images.

The DNA is analyzed during the indexation process by various expert modules in order recognize content in relation to existing knowledge bases. The description can be absolute ("what is it ? ", "It is...") or relative ("What is similar to this image ? ", "It looks like...")

These knowledge bases can be enriched: the analysis engine can be trained to recognize new categories of objects.



## Automatic classification



## Semi-Automatic Indexation

Indexation of images determines the subsequent retrieval of images, since indexing index comes down basically to the task of numbering images. Usually, this indexation process is done manually, by adding standardized descriptive annotations to images (metadata).

In the case of very large image databases, the challenge of indexation becomes critical. In terms of time, but more especially in terms of human resources. For this reason, it is economically more efficient "to propagate" the descriptive annotations of a well referenced image database to the set of images considered to be similar by the computer. Called "Automatic Suggestion of Keywords", the process then appears as a semi-automatic system: proposal of keywords by the machine, then confirmation - or rejection - by a human being. One can thus capitalize on the existing metadata of an organization and reduce considerably the indexation costs.

## Database Navigation

Today's huge image databases are leading to the emergence of new ways of data navigation.

"Traditional" database navigation operates on annotations associated with images. This metadata can be generated directly by camera devices (date or even location, thanks to mobile global positioning systems) or be indirectly inferred (with an analysis of web page text where the photograph is

posted, for example) This navigation does not take into account the contents of the images in themselves.

However, most of the time, there is no metadata in the appropriate fields of the image file. When missing, the navigation has to be purely visual: navigation by "visual similarity" consists in submitting to the visual engine an "image query", which responds by retrieving a set of "visually similar" images. Called bouncing, this way makes it possible to converge toward a result, in a natural and interactive way. It refines the search progressively by adjustments to the weightings of color and form. Juxtaposing images considered as similar by the computer also may reveal aspects of the image that were previously unnoticed.

It remains that text-based and image-based search are complementary: the combination being better than the best of both separately. Thus it is more powerful to couple visual and text research. This complementarity is already in use and has proven to be very efficient.

This combination speeds up the investigation process.

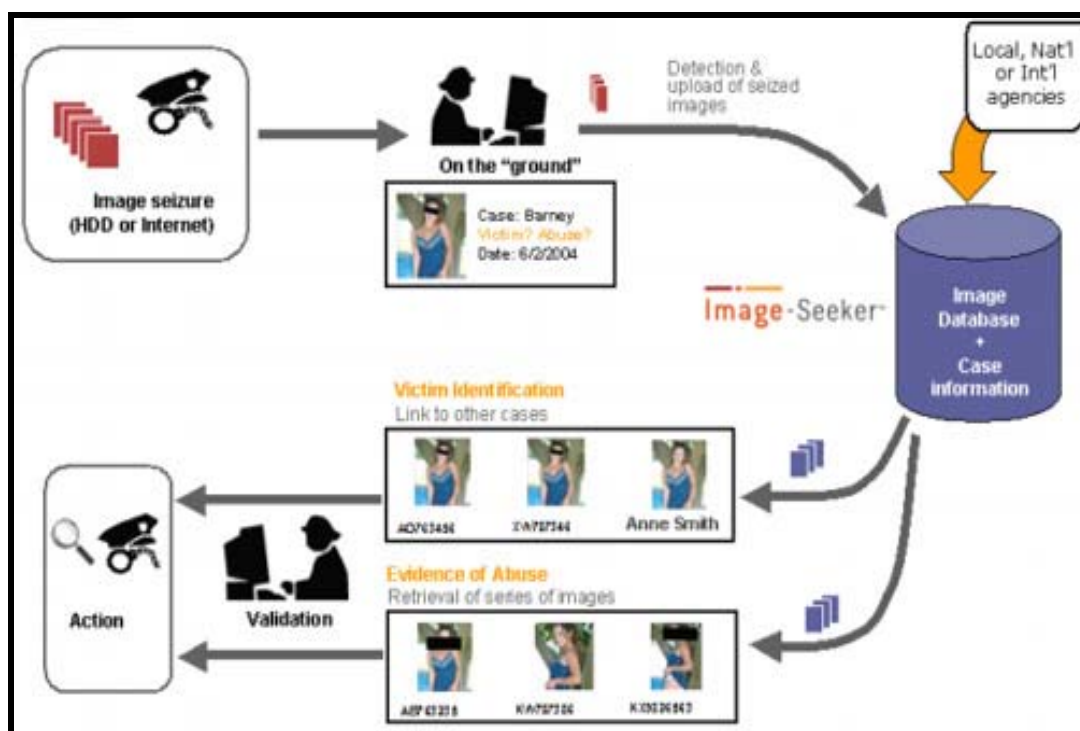
## **COMPUTER VISION APPLICATIONS**

### **Fighting Child Pornography**

Deployed by the FBI and US Customs, computer vision contributes to the dismantling of child pornography networks. The technology is also in use by several national police forces in Europe and is to be deployed in other European and international organizations targeting child abuse. Some organizations wish to remain unidentified.

The comparison of newly-seized images with images in existing law enforcement databases is largely automated, which accelerates the investigations. Detailed reports identify duplicate, cloned, and similar images as well as indicating those images which may be newly-produced (that is, completely different from any images in the existing databases) This allows a focus on tracing the distribution of the images and, most importantly, identifying the victims and the criminals.

## Image-based investigation



Pedophiles exchange both images and videos, and the majority of these are not recent and have already been identified by law enforcement organizations. Investigative agencies have identified several hundred thousand unique images ; the use of Internet multiplies the exchanges considerably.

Technology providers are in the process of adapting existing still image recognition technologies to the task of analyzing video content. There is also progress in the interception of Peer-to-Peer transfers and even the identification of the those who are distributing the files.

Assembling the Image DNA of all known child pornography images would make it possible to create the ultimate filter - impossible to circumvent - to stop the proliferation of these images on the Internet. The millions of personal pages hosted on Internet servers and the DNA of their images could easily be compared with the DNA of already identified pedophile images. This comparison would be completely automatic and not require any human intervention. The robustness of Image DNA identification would prevent the publication of a known criminal image, even if altered. This is analogous to antivirus protection that traps virus signatures.



Filtering duplicate and cloned images can thus be entirely automatic. Then the real need for human intervention and censoring would arise only when the computer flags the images as being similar to the blacklisted images, but only similar and perhaps not even criminal. In the end, only human judgment has the last word. Similarly, in the surveillance of internet chat and text message exchanges, the detection of offensive language, such as expressing racial hatred and other offensive statement, confronts the same limits: automated filters are able to block statements that clearly match prohibited statements, but the human intervention of a chat moderator is necessary in judging anything ambiguous

### **Fighting Counterfeiting and traffic in stolen art**

According to OECD sources, counterfeiting accounts for 5 to 7% of world economic trade. This is detrimental to countries, companies, and consumers. Counterfeiting has long been the domain of organized crime and even terrorist groups, and counterfeiting affects many sectors of business, particularly computer industries, toys, perfumery, pharmaceutical and automotive sectors.

LTU's technology is used in the United States in the fight against counterfeiters who falsify identity cards, passports, currencies and bank notes ,checks and credit cards (credit cards in the US and certain other countries contain no microchip; thus they are more susceptible to counterfeiting)., The counterfeiters use high resolution images to create the counterfeit items. At the time of a warranted search or an arrest, computer hard drives are seized and then examined for evidence of the high-resolution images that are used in counterfeiting . A manual review of all the files would be strictly impossible, so an automatic analysis quickly determines the presence or absence of incriminating files. Images are classified into categories, making it possible for agents to concentrate fully on their investigations.

According to Interpol, the clandestine traffic of cultural goods has become one of the dominant forms of international criminality. The two countries most particularly affected are France and Italy, but Central Europe as well as Africa are also are also concerned. An investigation system is being tested in Europe, involving industry, insurance companies, and police forces.

## **Economic intelligence and Internet surveillance**

Artificial vision technologies also find applications in the protection of corporate brands, logos, and intellectual property. Images may be gathered from all over the Web and then subjected to the same image analysis technology described above (this is done through special programs that look for particular kinds of images by “crawling” the web, hence their nicknames “spiders” or “crawlers”). The retrieved images are then analyzed, classified, and documented in automated reports. These programs can analyze even camouflaged files (such as those named with a false file extensions, for example). as well as compressed files or the enclosures of emails. The list of images to monitor on either the “visible” or “invisible” Web is long and corresponds to as many different investigative objectives: images which are illegal,, defamatory, or suspicious (amateur bomb plans ), etc.

Large companies monitor the use or misuse of their trademarks on the Internet. These companies seek mainly to detect and counter-attack the “gray” market, i.e. all unauthorized channels of distribution of brand name products, as well as the distribution of counterfeit products.

Some companies, whose trademarks represent a significant part of their intangible assets, monitor the misuse or imitation of their logos and thus anticipate denigration campaigns or attempts at destabilization. Others monitor the publication of copyrighted images without payment of rights, or even study the extension of the technology to contravene the hacking of digital video contents. This last point represents 8 billion dollars of losses for the companies of this sector.

Corporate communication policy is not the only area concerned. Certain companies working on sensitive subjects monitor and intercept accidental or malevolent document leaks (prototypes, schemas) by filtering outgoing traffic of messaging systems. While there is a plethora of solutions for monitoring incoming traffic (anti-spam devices), solutions to monitor the outgoing traffic are less developed.

## **Biometrics**

Since the attacks of September 11, 2001, and with the increase in terrorist threats, attention has focused on applying image recognition technologies to video surveillance. This relates to movement, the detection and the memorizing of people, in relation to the concept of "recognition"

In the current state of the art, the idea of trying to locate a known criminal with cameras installed in the street and coupled to a recognition software is still pure science fiction.

Out of the three components related to recognizing people (detection, verification, identification), the first two are today relatively well mastered. It is identification that really poses a problem. The precision of identification is closely related to whether the subject participates actively or passively, which determines the conditions of the acquisition of images.

Without an active participation, the performance of recognition deteriorates quickly. In order for facial identification to become reliable and achieve industry-acceptable results, the subjects photographed must play the game and be photographed according to a particularly described procedure. This calls into question the face recognition technologies designed to monitor uncooperating subjects automatically and reliably (such as in stadiums or airports).

Efficient facial recognition is thus not yet possible, and neither existing nor developing technologies offer any hope for relevant facial analysis in the near future. An entirely automated research where the system would immediately and correctly identify the suspect is implausible. On the other hand, a semi-automatic approach makes it possible to save valuable time for investigators, by combining various complementary approaches (available biometric technologies, image retrievals by similarity, and others).

Behavior recognition (as might be used in monitoring in public transport) is even less developed than facial recognition. It is more recent and, for this reason, more confidential, and very few solutions have advanced to the phase of commercialization.

## **PERSPECTIVES**

Text mining technologies have interested knowledge management industries and intelligence agencies for years. Regarding images, it is striking to notice that fewer than a dozen suppliers provide efficient and innovative solutions in the world. The most advanced image mining technology is French.

The performance of image research, filtering and categorization depends on many variables. But two variables are particularly prominent: the power of the processor and the amount of live memory. And of course since Information Technology history has shown continual increases in both processor and memory capacities, adequate computing resources are likely to be available for the increasing needs of image research technologies. This will enhance both the speed of information retrieval and the relevance and accuracy of results.

The combination of data mining technologies (text, audio, image) will outperform data flow analysis.

Lastly, progress related to man-machine interfaces (issuing commands to the computer by voice and by gesture, new techniques for the representation and visualization of data, immersive workspaces, etc) will certainly enrich the possibilities and the strategies of image search.

## **CONCLUSION**

The vast proliferation of visual content on the Internet contributed largely to increasingly important needs for content management and monitoring. In particular, the amount of visual data has experienced such significant growth that the recourse to artificial vision techniques becomes mandatory. Technology is not aiming at replacing the human being. To the contrary, it is a powerful tool at the service of investigators, who can let the software find image correlations and dedicate themselves to higher-level decisions and actions.

If massive data acquisition is not really a problem, the situation is quite different for data analysis. Articulation between Sigint (Signal Intelligence) and Humint (Human Intelligence) remains more than ever topical. In the end, only human judgment is able to decide definitively whether an image is licit or illicit, to extract "relevant" elements for a given context. Nevertheless, the computer brings a major contribution by analyzing automatically and rapidly huge amounts of data. The power of image recognition technology is a key element in the general mastery of information.

François VELTZ

François Veltz is an engineer and holds an MBA. He worked for Silicon Graphics (SGI) before joining LTU Technologies.