International Symposium on Visual Analytics Science and Technology (2010)
J. Kohlhammer and D. Keim (Editors)

Visual Analytics to Check Marine Containers in the Eritr@c Project

M. Aupetit[1], L. Allano[1], I. Espagnon[1], G. Sannie[1]

[1]CEA, LIST, F-91191 Gif-sur-Yvette, France

michael.aupetit@cea.fr, lorene.allano@cea.fr, isabelle.espagnon@cea.fr, guillaume.sannie@cea.fr

Abstract

The Eritr@c project is a European project aiming at developing an inspection system to control marine containers for illicit or dangerous materials. The system is made of a neutron generator, a set of sensors, and an information system to process the data. One of the main parts of the system is an interactive visualization interface whose goal is to help custom officers to decide if a container must be opened for deeper check of its content. In this paper, we present the components of this visual interface and their use for analytic reasoning.

Categories and subject descriptors: multidimensional data; nonlinear mapping; multidimensional scaling; visual mining; visual classification; visual analytics

1. Introduction

Checking marine containers is an essential process in the international security of goods and people. About 200 millions containers pass through principal seaport each year. However, checking in depth each container by opening it is far too expensive a process.

The Eritr@c (European Ripost against Illicit TR@ffiCking) project started in 2008 and will end in June 2010. It follows the Euritrack project, and gathers seven European partners (CEA, UNIBS, UNIPD, IRB, ACT, RC and JRC[1]) and is founded by the European Union [Eur06]. The aim of these projects is to design and build a system to check the content of marine container without opening them. Eritr@ck especially improved over Euritrack, in terms of the precision of the measurements, the quantity of data and the information processing and visualization system.

2. From measures to decision support

For each container, the custom officer analyzes the accompanying manifest, which is a paper document specifying its source and destination, the owner of the goods and some crude description of its global content (“TV screen”, “plastic shoes”, “tables”, “toys”…). Based on this document, he decides to release the container or to inspect it further with an X-ray scan. In the latter case, a truck brings the container in the scan area. This scan reveals partly the geometry and the density of the materials inside the whole container. Based on this additional information, the custom officer decides to release the container or to perform a deeper scan on a small area using the neutron generator developed during the Euritrack project [CPB*08] (Figure 1). The neutron beam is a conic section about 50 cm by 50 cm tall and 2 m deep. The neutron generator throws neutrons during 10 minutes into the neutron beam in order to get sufficient statistics. Sensors measure the energy spectrum radiated by the materials at (x,y,z) position in the neutron beam with a few cm precision. Spectrum unfolding techniques are used to estimate the proportion of 14 chemi-

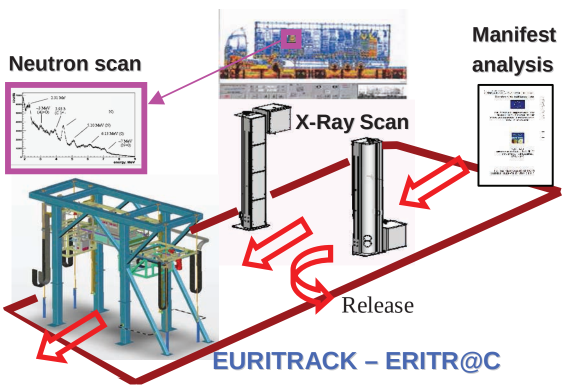

Figure 1 : *Checking process overview*

[1] CEA : French Atomic Energy Commission; UNIBS: University of Brescia; UNIPD: University of Padova; IRB: Ruder Boskovic Institute; ACT: Analysis and Control Technologies; RC: Revenue Commissioners; JRC: Joint Research Center

DOI: 10.2312/PE/EuroVAST/EuroVAST10/057-060

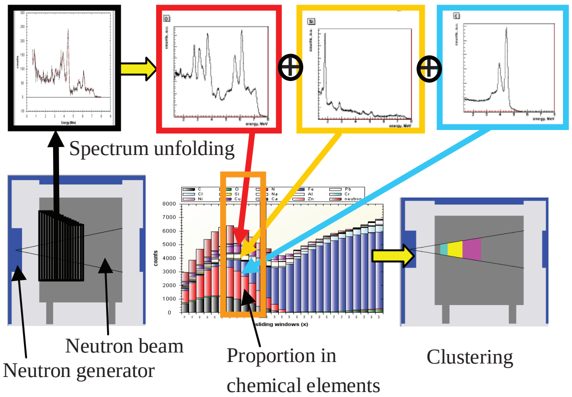

Figure 2 : *From signal to chemical proportions: energy spectrum is unfolded to get proportions of chemical elements, and clusters of homogeneous materials*

cal elements[2] in each voxel of the beam. Then a clustering technique is applied to partition the voxels in up to three spatial zones of homogeneous chemical content (Figure 2). This content is then a vector of the 14 chemical proportions averaged over the zone.

At that point, the only external information the custom officer has is the X-ray picture, and the container's manifest. We do not consider the manifest content as a possible class label for the vector of chemical proportions because of its heterogeneity, non standard format, and high level of uncertainty depending on the neutron beam location. Based on the data we have, it is not currently possible to provide automatic exploration guiding of the custom officer because there is no clear criterion about which area to check based on the X-ray picture, nor it is possible to check many areas to cover the whole container because of the 10-minute time a measure spends to get reliable chemical proportions and the high cost per time unit of holding a container at the scan area. So we propose to provide visual supports to help the custom officer to design his own exploration pathway and to decide whether to release the container, or to open it for up to several hours of manual inspection if illicit or dangerous materials are very likely to be discovered based on the analytics.

In order to support the decision of the custom officer, we develop three ways of visualizing the content of a zone: the Alert Triangle; the Material Triangle and the Content Map.

We scanned 74 containers whose content was known approximately based on their manifest, and collected the estimated chemical proportions in a database to provide reference data for the Content Map. We also use the theoretical proportions of chemical elements in 7 drugs, 8 explosives and 23 neutral materials, to provide reference data for the Alert Triangle and its automatic traffic lights.

[2] C, O, N, Fe, Pb, Al, Cu, Si, Cl, Na, Cr, Ni, Ca, Zn

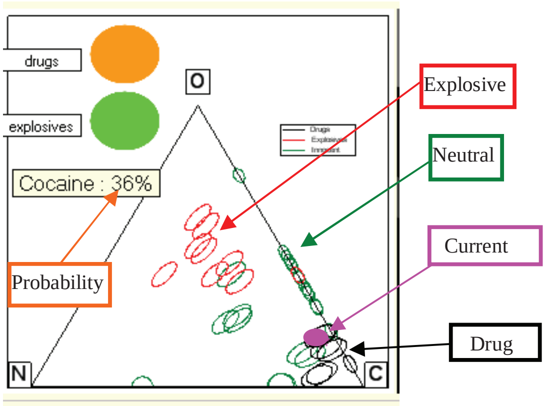

Figure 3 : *The Alert Triangle : reference materials (neutral, explosive or drugs) represented with their uncertainty as ellipse in barycentric coordinates. Traffic lights indicate the risk about drugs and explosive in the current zone based on a kernel density model.*

3. The Alert Triangle

We display some reference materials like typical explosives, typical drugs and other neutral materials (wood, iron…) in a barycentric coordinate system based on their respective proportions of carbon, nitrogen and oxygen, because drugs and explosives are mainly made up of these chemical elements (Figure 3). We model the uncertainty due to the spectrum unfolding process with Gaussian kernels estimated with a Monte-Carlo method, and show the covariance ellipse of each material to the user. The material measured for the zone under inspection is projected into this triangle with its own uncertainty area. Traffic lights display an automatic decision as "no", "possible" or "full" risk of a drug or explosive content using Bayes rule based on the Gaussian probability densities of the current and the reference materials. This Alert Triangle supports the automatic traffic light decision, and helps the user to understand why this decision is taken, and what to decide about it.

4. The Material Triangle

Chemical elements (except Cl, Na and Ca) are grouped into three classes based on the materials mainly detected in containers at the Rijeka, Croatia, seaport where the Eritr@c system is installed : Organic (C, N, O); Metallic (Fe, Cu, Pl, Al, Cr, Ni, Zn); Ceramic (Si). These classes are more meaningful to the custom officer. The current material is projected into the barycentric coordinate system based on its added proportions of elements in the above three classes. It appears as an ellipse (95% confidence) in this Material Triangle which helps the custom officer to figure out what kind of material is inside the beam (Figure 4). It is a way for instance, to detect illicit trafficking like copper rolls in a container whose manifest indicates wooden-made or plastic-made items which clearly not belong to the metallic class. A message “Non-Organic”, “Non-Metallic” or “Non-Ceramic” is automatically displayed if the proportions of Organic, Metallic or Ceramic classes respectively are lower than 5%.

5. The Content Map

All the reference data each coded as a 14-dimensional vector of proportions in chemical elements measured in a zone of a reference container, are projected into a plane using a nonlinear MDS mapping method called DD-HDS [LVG*07] which is known to well preserve the similarities between the original data (Figure 4 and 5).

Based on the work of Aupetit [Aup07], we build the Voronoï cells of the projected reference data in the plane, and color these cells to visualize the correlation measured in the original 14-dimension space between each reference data and the current query data. The main difference with our previous work [Aup07] which makes original this approach, is that the query data is not mapped with the references as a point into the plane, instead, only its original similarity to them measured as a correlation in the multidimensional space, is visualized through the coloring.

It has two benefits. First, locating a new point in a given map is not straightforward because its neighbors in the multidimensional spaces may be displayed in very different areas of the map due to manifold tearing, here we avoid this because we do not require to locate the current data as a point into the map. Second, the mapping quality is in general prone to many distortions [Aup07], but showing the original similarity as a color provides the “truth” about the similarity between the reference data and the current one as it is measured in the original space.

Voronoï cells are used to make visible the color which codes for each reference data its original similarity measure to the query, and to make the overall picture look like a mosaic work easier to remember than a set of colored points tacked on a blackboard. The size of the Voronoï cells has no meaning, it is a side effect of the scattering of the projected reference points (not displayed) whose density somehow estimates the one of the data in the original space.

Moreover, using an MDS map of the reference data has two benefits compared to random maps (Figure 6). First, it

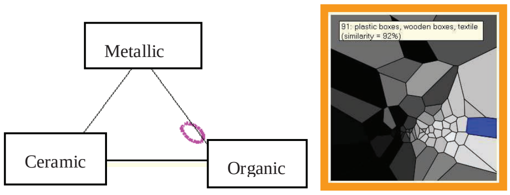

Figure 4: Left the *Material Triangle: the current container is displayed with a 95% confidence ellipse in a barycentric coordinate system. Right the Content Map: an MDS map of the reference data (formerly inspected containers). The original “true” similarity of the query data (currently inspected container) to the reference containers in the original 14-dimensional space, is shown as the brightness of their Voronoï cells (The closer, the brighter). The user can look at the content of a specific reference container (blue cells) and access to its degree of similarity to the query.*

is easier for the user to memorize patterns because similar reference data are more likely to be mapped close by and to light up together for a given query data, building up a pattern with a lower topological complexity. Second, it is safer to distinguish between different patterns in MDS because two adjacent cells in random maps are likely to show uncorrelated lighting, which if permuted for two different query data is likely to remain unnoticed, making both query data looking more similar than they are. For this same reason, it is also an easier way for the user to remember the map picture from one analysis to another, and use it as a reference and support for his memory in his recurrent analytical work.

Finally, the custom officer can fly over the map to show off the content of specific zone in reference containers as told by their manifest, and make his own idea about the content of the current zone (query data) based on its similarity to the reference data (Figure 4).

6. A visual analytics view

While the Alert Triangle is mainly focused on supporting the automatic discovery of dangerous or illicit materials, the Material Triangle and the Content Map provide decision support for consistency checking of the manifest *versus* the container content, an analytical process otherwise very difficult to automate. Thanks to these visual data representations, the custom officer can compare several zones from the same container or from different containers to each other.

For instance, as shown in the figure 5, a container which contains roofing felt as told by its manifest, shows non homogeneous densities of materials on areas of the X-ray scan. The custom officer decided to inspect two of these different areas with the neutron scan and to compare the results. The measure in the first area provided one zone of homogeneous material, while the measure in the second one provided two zones with clearly different contents. The content of one of these zones is similar to the one measured in the first area and compatible with the manifest, while the remaining zone contains non organic material immediately distinguishable from the other two based on the comparison of the three visualization paradigms. The custom officer eventually decided to open the truck.

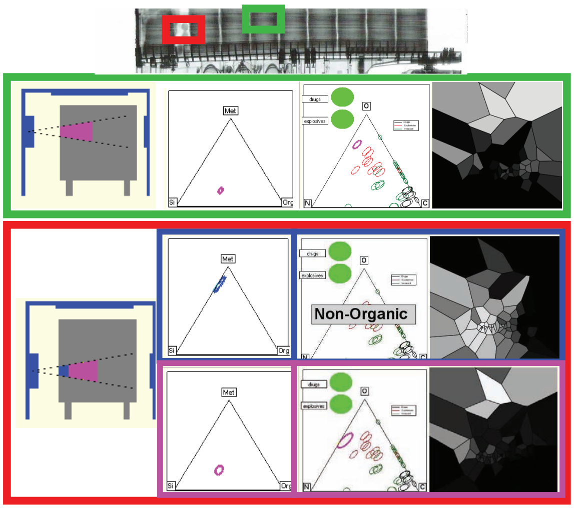

Figure 5: *Visual analytics: a first measure provides the views on the green row (Green square in the X-ray scan). A second measure in a different area of the X-ray scan (Red square) provides two clusters of homogeneous material (blue and magenta rows). While the magenta and green zones show similar content in all the views, the blue zone shows a distinct content, which appears to be not consistent with the manifest.*

7. Conclusion

We presented three visualization paradigms in an interactive framework to help the custom officer to decide whether or not to open a container for manual inspection. The Alert Triangle is used to make the user more confident about the automatic detection system of drugs and explosives. The visualization of reference data in the Alert Triangle and Content Map provides contextual information, and a recurrent mapping which could help the user to draw the landscape of his own decision space. At last the Material Triangle and the Content Map are used to complement the manifest by translating the chemical elements in terms of materials familiar to the user in order to help him to detect illicit trafficking.

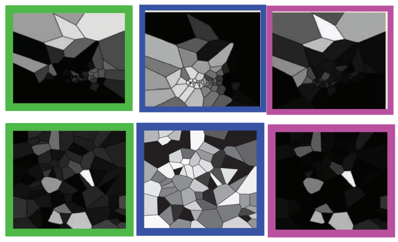

Figure 6: *Content Maps with MDS (top) and random (bottom) mappings of the same reference data. Graphics with the same color frame as in Figure 5 display the same query data.*

By the way, the Content Map shows always the same reference data, a future work would be to update this map with upcoming data while maintaining its overall visual aspect to prevent the user from the need to learn again from scratch its mental representation. Another task which follows is to gather comments from custom officers about this visualization system during a test period in real working conditions.

Acknowledgments

We are grateful to Bertrand Pérot, Wassila El Kanawati and Cédric Carasco (CEA DEN) who set up the neutron generator and sensor system during both the Euritrack and Eritr@c project, and who provide us with the reference chemical materials used in the Eritr@c information system. We also thank all the partners of the Euritrack and Eritr@c projects [Eur06].

References

[Aup07] Aupetit M., Visualizing distortions and recovering topology in continuous projection techniques" Neurocomputing, vol. 10, no. 7-9 pp. 1304–1330, 2007.

[CPB*08] Carasco C, Perot B, Bernard S, Mariani A, Szabo J, Sannie G, Roll T, Valkovic V, Sudac D, Viesti G, Lunardon M, Bottosso C, Fabris D, Nebbia G, Pesente S, Moretto S, Zenoni A, Donzella A, Moszynski M, Gierlik M, Batsch T, Wolski D, Klamra W, Le Tourneur P, Lhuissier M, Colonna A, Tintori C, Peerani P, Sequeira V, Salvato M. *In Field Tests of the EURITRACK Tagged Neutron Inspection System* . NUCLEAR INSTRUMENTS and METHODS IN PHYSICS RESEARCH SECTION A-ACCELERATORS SP 588 (3); 2008.p. 397-405.

[Eur06] Euritrack and Eritr@c projects web site : http://www.euritrack.org

[LVG*07] Lespinats S., Verleysen M., Giron A. and Fertil B., DD-HDS: a tool for visualization and exploration of highdimensional data, IEEE Trans. Neural Netw., vol. 18, no. 5, pp. 1265-1279, 2007.