# ON THE IMAGE CONTENT OF THE ESA EUSC JRC WORKSHOP ON IMAGE INFORMATION MINING

*Marco Quartulli, Mikel Zorrilla, Igor G. Olaizola*

Vicomtech-IK4
Mikeletegi Pasealekua 57, Parque Tecnológico
20009 Donostia / San Sebastián, Spain

**ABSTRACT**

Due to the specific topic the meeting is devoted to, the images appearing in the papers and presentations published through the years on the websites of the ESA EUSC JRC IIM Workshop often play a central role in the ongoing public discourse on methodology and applications.

While tools for analyzing textual content are widespread, we introduce instead a methodology for semi–automatically organizing published material based on its image content. The sources of quality degradation that are introduced by the publication process are taken into account and partially inverted. Specific methods are introduced to take into account the Earth observation–related specificities of the published image data. A semi-supervised data analysis and labeling step is used to produce semantically enriched descriptions in a set of linked RDF ontologies.

This information is not intended anymore for mere presentation on paper or on screen.

It can be aggregated and re–used, e.g. to answer complex queries mixing content– and metadata–based attributes of the items in the dataset.

***Index Terms*—** Remote sensing, Content-Based Image Retrieval

## 1. INTRODUCTION

Workshop and conference proceedings typically present a corpus of separate contributions and related proceeding paper and presentation files, organizing them by publication year and by session.

Textual analysis can be used to e.g. recognize named entities referring to specific sensors, methodologies and geospatial locations. Yet, in addition to textual content, most contributions also include valuable and potentially re–usable image elements: conceptual diagrams, input and output datasets, result and ground truth maps.

We consider the possibility to analyze and aggregate published contributions on the basis of their image content as a valuable Image Information Mining objective.

We have analyzed the presentations of the 2011 ESA EUSC JRC workshop on Image Information Mining for Geospatial Intelligence from Earth Observation by the semi–supervised methodology described in this paper. To better demonstrate the capabilities of the system, we also included in the input dataset a number of contributions from the Sixth Conference on Image Information Mining held in 2009, and notably all available presentation in OpenDocument–compatible formats from the November 4, 2009, Conference Day 2, Morning Session "Bridging the gap with operations", which focused on results obtained by different groups starting from a common dataset.

## 2. QUALITY DEGRADATION SOURCES

The published images are degraded with respect to the original acquired signals: spatial downsampling, spectral reduction to a tri–stimulus color space, dynamic range limitations, compression artifacts and perceptual optimization–induced degradations such as an increase in the saturation or the application of logarithmic curves to SAR intensity data all play a role. A symptom is visible in fig. 1: the edges of a pyramid shape are well perceivable in a feature space plot projection on the first 2 PCA components derived by color descriptors decorated by unsupervised class identifiers. Degradation due to pre-publishing perceptual enhancement/"beautification" is probably responsible for the relative importance of the edges with respect to the less saturated, less contrasted volume space. While a reconstruction of the original data would entail the elimination of compression artifacts, the estimation of likely multi–spectral values from observable colors and some sort of super–resolution, the analysis of published image content in terms of semantic content models taking into account the degradation effects (fig.2) can be achieved by simpler inversion methodologies.

## 3. INVERSION METHODOLOGY

Ground truth is defined by supervised tagging of the automatically extracted image files with one of four classes ('EO

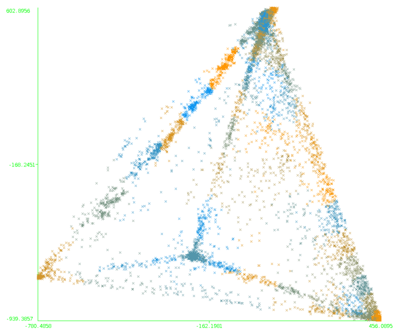


**Fig. 1**. An empty color pyramid volume as a symptom of signal degradation in published material.

**Table 1**. Confusion matrix for C5.4 tree–based classification of reference dataset.

| a | b | c | d | ← classified as |
|---|---|---|---|---|
| 3422 | 0 | 1 | 0 | a = drawing |
| 0 | 3423 | 0 | 0 | b = eo_multispectral |
| 0 | 0 | 3423 | 0 | c = eo_monochrome |
| 0 | 0 | 0 | 3423 | d = natural_photo |

monochrome', 'EO multispectral', 'Drawing' and 'Natural photo' — fig. 3a). The manual tagging is aided by a preliminary unsupervised X-means [3] classification based on the extracted primitive descriptors (see below). Multiple instances of the same image appear in the considered corpus of presentations (methodology comparison sessions on a specific given dataset). Metadata are extracted from the inputs (file size, format, color mode). EO-specific content descriptors are extracted. For example, global optical/SAR discrimination is operated by decision theory ratios of max likelihoods of Gaussian mixture, single Gaussian, and exponential distribution histogram fitting. All images are then converted to the $YC_bC_r$ color space and resampled to a standard size. Both image–level "global" and tile–level "local" primitive descriptors are extracted including Haralick [4], local binary patterns [5], threshold adjacency statistics [6] and Zernike texture estimators [7]. The sensitivity of SURF [8] and SIFT descriptors [9] to local gradients is complemented by considering the local color histogram in the $YC_bC_r$ space around the interest points identified by the feature extractor. TIFF LZW codebooks [10] and LZW-based Kolmogorov complexity estimates are generated. To evaluate the structure of the generated parameter space, a decision tree generated by the C5.4 algorithm [11] (table 1) is considered: only one instance is misclassified in the whole considered image set of 12,000 instances.

A primitive feature selection process by 1R [12] is carried out after equal size sampling based on the defined ground truth in order to avoid biasing classifiers with the largely different population sizes in the 'Drawing' and in other labeled

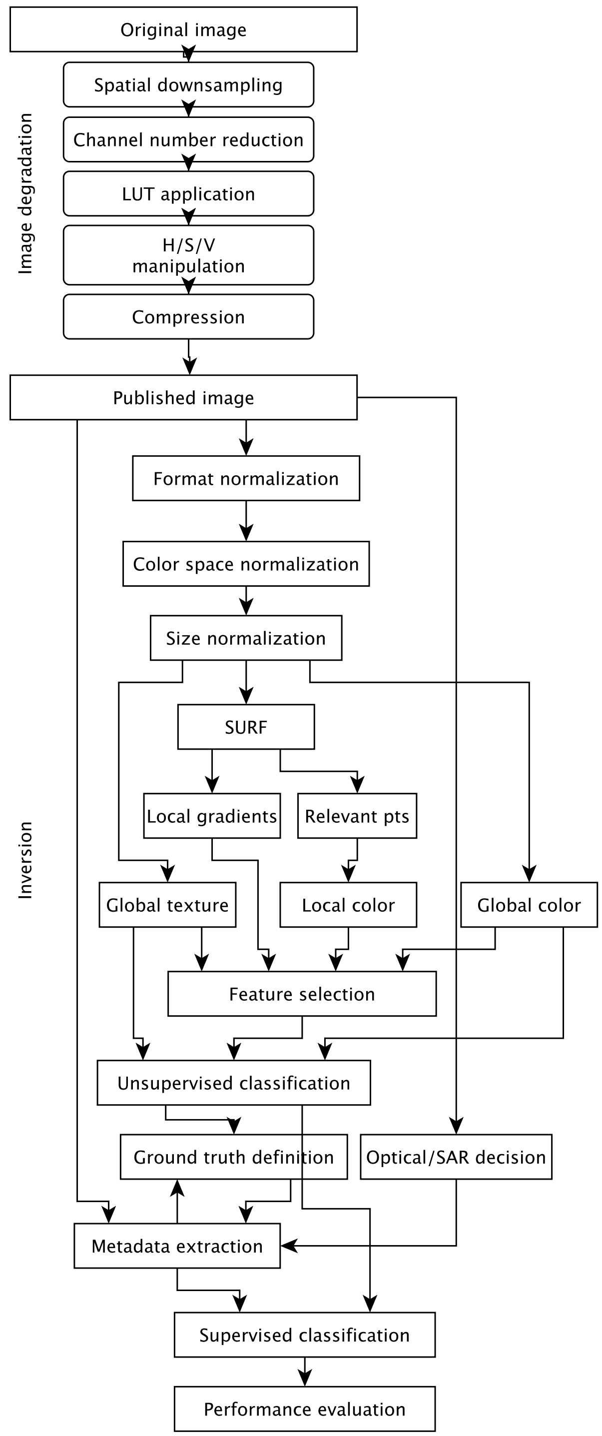
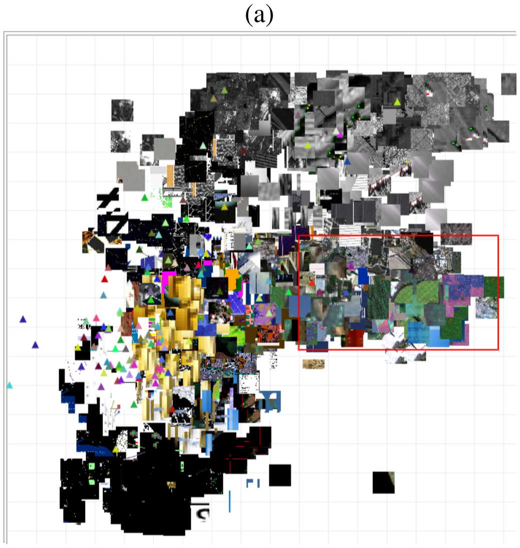



**Fig. 2**. Conceptual diagram: assumed degradation flow and proposed inversion methodology.

(a)



(b)

2009/15_30_PPT_Younan
2009/10_15_PPT_Soille
2009/10_40_PPT_Ramiro
2009/09_20_PPT_DElia
2009/09_50_PPT_Gomez
...
2011/09_40_Vaduva
2011/11_00_Bruzzone
2011/15_50_Trianni
2011/11_30_Barb
2011/10_20_Blancha
...

(c)

**Fig. 3**. (a) ESA EUSC JRC 2011 reference dataset with ground truth classes 'Drawing' (232 instances), 'EO multispectral' (68), 'EO mono/panchromatic' (33), 'Natural photo' (8 instances). (b) 2,500 2009-2011 image thumbnails in global color/texture feature projection (c) a subset of presentations intersecting the red area in (b).

sets. The resulting proportional relevance of global descriptors such as color, texture and edges corresponds to the nature of the considered ground truth classes.

Further descriptors are added to take into account the specific Earth observation origin of the considered images. For example, a decision ratio between maximum likelihoods of fit to Gaussian mixture and exponential distributions is used to differentiate between SAR and optical quick-looks.

PostgreSQL is selected for data handling due to its efficient support for multidimensional data based on GiST, a balanced tree–structured access method, that acts as a base template in which to implement arbitrary indexing schemes.

We approach the issue of massive semantic labeling by an HTML5 image tagging application based on a `CANVAS` element displaying tens of thousands of thumbnails in a user selectable 2D projection of the full feature space represented in a complete n–dimensional scatter plot right to its side. Regions defined in the feature space aggregate data with similar content. They can be saved and linked into an ontology represented as a graph in a side pane. Semantic tagging operations are implicitly performed by intersecting data points with the defined regions (fig. 3b).

We consider that, while a classification into drawings, natural photos, panchromatic and multispectral images is definitely insufficient in order to assess the relevance of the considered methods for the needs of the IIM/EO community, the fact that the system is able to accurately find identical (sub-)images used by different authors in different presentations — e.g. by SIFT descriptor nearest neighbor matching in a Euclidean space projected and reduced in dimensionality by PCA [13] — is suggestive of the performance of the system with respect to much more specific ground truth classes.

## 4. CONCLUSIONS

We introduced an image analysis tool (fig. 4) aimed at aggregated repositories of published material based on the content of their imagery. Its usability was investigated on content from the ESA EUSC JRC Image Information Mining Workshop from 2009 and 2011.

We aim at extending the system so that it can be used to obtain ground truth collections from published EO images to be used for large scale algorithm validation in open competitions.

## 5. REFERENCES


[1] M. Quartulli and I. Olaizola, "A review of EO image information mining," http://arxiv.org/abs/1203.0747 - To be published on ISPRS Journal of Photogrammetry and Remote Sensing, February 2012.

[2] L. Ratinov and D. Roth, "Design challenges and misconceptions in named entity recognition," in *Proceed-*

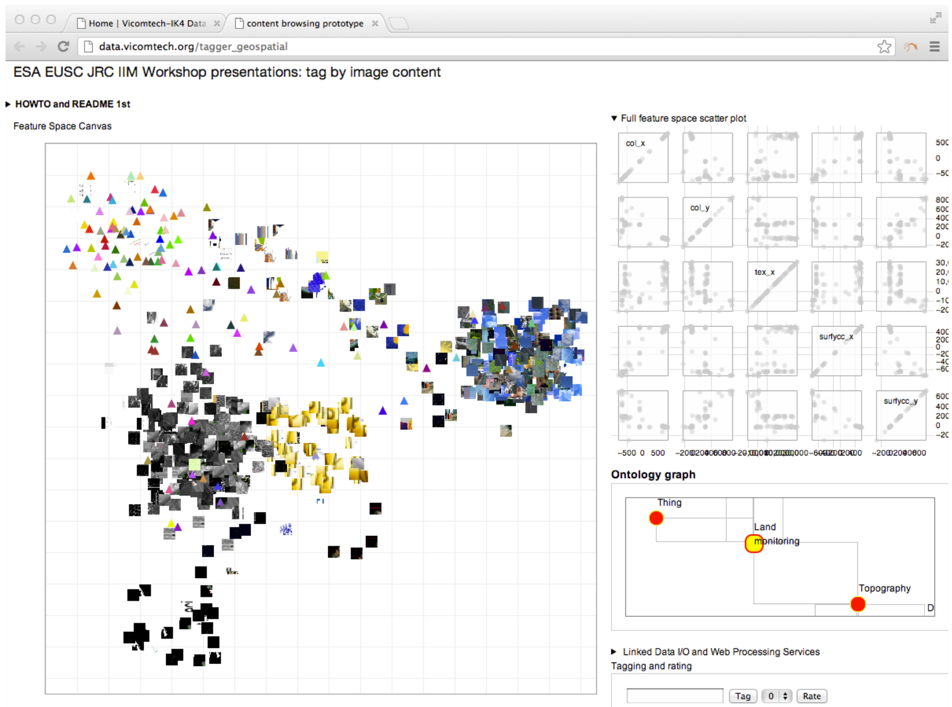



**Fig. 4**. Prototype web–based user interface. Image region thumbnails are presented in a user–selectable 2D projection (top left) of a space represented in a scatter matrix (top right). Data selection hypercubes are connected with the concepts in an application ontology (bottom right).

*ings of the Thirteenth Conference on Computational Natural Language Learning*. Association for Computational Linguistics, 2009, pp. 147–155.

[3] A.K. Jain, "Data clustering: 50 years beyond k-means," *Pattern Recognition Letters*, vol. 31, no. 8, pp. 651–666, 2010.

[4] R.M. Haralick, K. Shanmugam, and I.H. Dinstein, "Textural features for image classification," *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 3, no. 6, pp. 610–621, 1973.

[5] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Gray scale and rotation invariant texture classification with local binary patterns," *Computer Vision-ECCV 2000*, pp. 404–420, 2000.

[6] N.A. Hamilton, R.S. Pantelic, K. Hanson, and R.D. Teasdale, "Fast automated cell phenotype image classification," *BMC bioinformatics*, vol. 8, no. 1, pp. 110, 2007.

[7] M.R. Teague, "Image analysis via the general theory of moments," *JOSA*, vol. 70, no. 8, pp. 920–930, 1980.

[8] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, 2008.

[9] D.G. Lowe, "Object recognition from local scale-invariant features," in *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*. Ieee, 1999, vol. 2, pp. 1150–1157.

[10] D. Cerra and M. Datcu, "A fast compression-based similarity measure with applications to content-based image retrieval," *J. Visual Communication and Image Representation*, vol. 23, no. 2, pp. 293–302, 2012.

[11] A. Cufoglu, M. Lohi, and K. Madani, "A comparative study of selected classifiers with classification accuracy in user profiling," in *Computer Science and Information Engineering, 2009 WRI World Congress on*. IEEE, 2009, vol. 3, pp. 708–712.

[12] R.C. Holte, "Very simple classification rules perform well on most commonly used datasets," *Machine learning*, vol. 11, no. 1, pp. 63–90, 1993.

[13] Y. Ke and R. Sukthankar, "Pca-sift: A more distinctive representation for local image descriptors," in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*. IEEE, 2004, vol. 2, pp. II–506.