



HAL
open science

A histogram semantic-based distance for multiresolution image classification

Camille Kurtz, Nicolas Passat, Pierre Gañçarski, Anne Puissant

► **To cite this version:**

Camille Kurtz, Nicolas Passat, Pierre Gañçarski, Anne Puissant. A histogram semantic-based distance for multiresolution image classification. International Conference on Image Processing (ICIP), 2012, Orlando, United States. pp.1157-1160, 10.1109/ICIP.2012.6467070 . hal-01695060

HAL Id: hal-01695060

<https://hal.univ-reims.fr/hal-01695060>

Submitted on 15 Feb 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A HISTOGRAM SEMANTIC-BASED DISTANCE FOR MULTIREOLUTION IMAGE CLASSIFICATION

Camille Kurtz, Student Member, IEEE, Nicolas Passat, Pierre Gañcarski

Anne Puissant

LSIIT, UMR CNRS 7005
Université de Strasbourg, France
{ckurtz, passat, gancarski}@unistra.fr

LIVE, ERL CNRS 7230
Université de Strasbourg, France
anne.puissant@live-cnrs.unistra.fr

ABSTRACT

Image classification methods based on histogram analysis generally require to use relevant distances for histogram comparison. In this article, we propose a new distance devoted to compare histograms associated to semantic concepts linked by (dis)similarity correlations. This distance, whose computation relies on a hierarchical strategy, captures the multilevel semantic relations between these concepts. It also inherits from the low complexity properties of standard bin-to-bin distances, thus leading to fast and accurate results in the context of multiresolution image classification. Experiments performed on satellite images emphasize the relevance and usefulness of the proposed distance.

Index Terms— Histogram distance, Classification, Background knowledge, Multiresolution images, Remote sensing.

I. INTRODUCTION

THANKS to their ability to model image information in a synthetic fashion and with a high robustness to transformations (translation, rotation, scaling), histograms have been early and widely used in various image processing and analysis fields, *e.g.*, segmentation, retrieval or classification [1]–[3]. In this context, histograms have been frequently used to model low-level information/features intrinsically linked to the images, *e.g.*, intensity (grey-level or multispectral) distribution, texture parameters, etc.

However, when dealing with object-based approaches (*e.g.*, segmentation or classification), and in particular in the case of hierarchical strategies and/or the handling of multiresolution data [4], [5], it may also be relevant to use histograms in order to model information of higher level [6]–[8]. For instance, the segments extracted from an image by a segmentation process at a given scale/resolution may be classified by using their (semantic) composition in the image at another scale/resolution. Such compositions can be naturally modelled by nominal histograms (*i.e.*, non-ordered ones) where each bin is then associated to a specific semantic class.

This article deals with the issue of comparing such nominal histograms, and in particular those for which (dis)similarity information is available between the semantic concepts associated to each bin. This comparison is indeed mandatory to ease and improve the classification of image segments whose semantic composition is modelled by such histograms.

The comparison of nominal histograms is not a trivial task, and several measures of similarity have been proposed to this end in the last decades [9]. These histogram distances can be divided into two

categories, namely *bin-to-bin* and *cross-bin* distances. The bin-to-bin distances consider a histogram as a fixed-dimensional vector and only compare the content of corresponding histogram bins (by using, for instance, the Manhattan or the Euclidean distances), while the cross-bin distances consider a histogram as an estimation of a probability density function and compare corresponding bins as well as non-corresponding ones.

On one hand, bin-to-bin distances are well fitted to process large datasets, in particular thanks to their low computational cost [10]. However, by definition, such distances cannot consider the semantic proximity between the bins (in particular, they suffer from both problems of histogram translations and bin size changes). On the other hand, cross-bin distances, which compare more exhaustively both corresponding and non-corresponding bins, enable to consider the semantic proximity between them [11], [12]. Practically, this can be done by assigning, to each pair of instances, a numerical weight modelling the degree of semantic proximity between the compared instances. These weights are often stored by using a (dis)similarity matrix. The counterpart of such strategies is the polynomial cost induced by these multiple bin comparisons. It then appears that, when comparing nominal histograms, the handling of semantic proximity between their instances seems incompatible with a low computational cost.

Following the idea that organizing the information carried out by the data into hierarchies constitutes a successful strategy for knowledge extraction tasks [13], we propose to use this approach for the definition of a new distance on nominal histograms. (To the best of our knowledge, such a hierarchical approach has been considered only once [14], but for a distance devoted to ordinal histogram comparison.) The purpose of the proposed distance, called *Hierarchical Semantic-Based Distance* (HSBD, for brief) is twofold. First, it aims to capture the multilevel semantic relations intrinsically carried out by the (dis)similarity matrix associated to the instances of the histograms. Second, by relying at each stage of the hierarchy on a bin-to-bin distance, it leads to a lower computational cost than cross-bin distances. It then provides a relevant tool in the context of multiresolution image classification.

The article is organized as follows. Section II provides information on the proposed distance, including its computation scheme (Section II-A) and a computational cost analysis (Section II-B). Section III describes experimental validations of this distance in the context of multiresolution satellite image classification. These experiments aim at reclassifying urban images preprocessed by a hybrid segmentation/classification method [8], recalled in Section III-A. The considered datasets and their classification results are detailed in Section III-B. Section III-C concludes the article by providing perspective works.

The research leading to these results has received funding from the French Agence Nationale de la Recherche (Grant Agreement ANR-10-COSI-012-03).

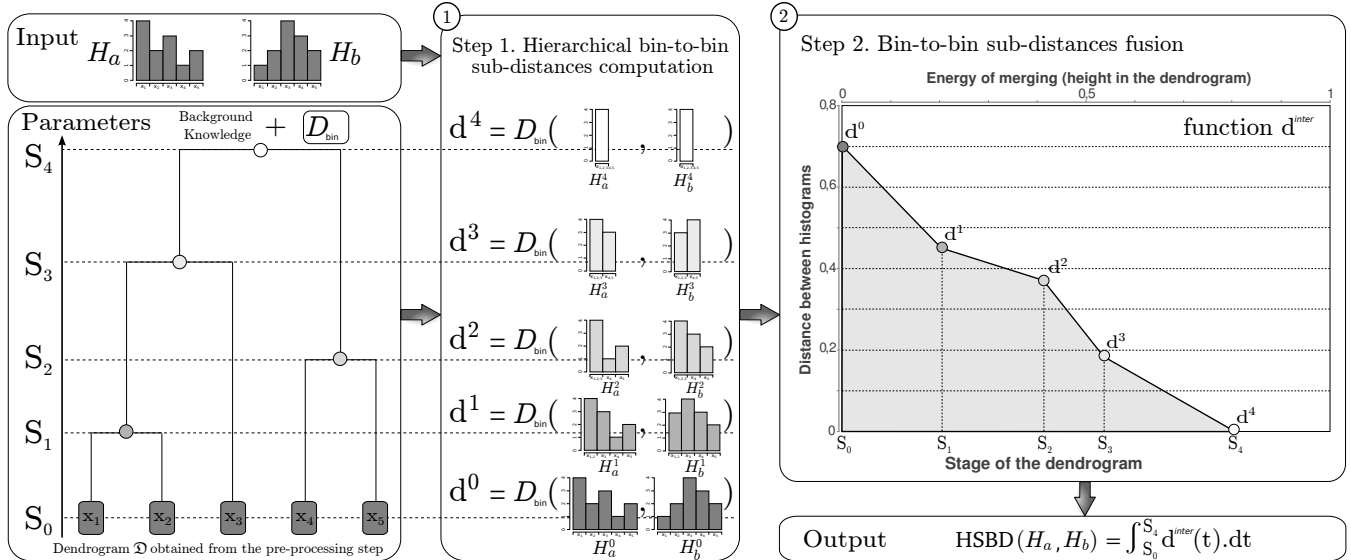


Fig. 1. Computation workflow of the HSB distance.

II. THE HSB DISTANCE

II-A. Distance computation

The computation of the HSB distance between two histograms H_a and H_b of v bins, requires two parameters:

- (1) A dissimilarity matrix \mathcal{M}^{dis} modelling the semantic proximity (from 0: equal, to 1: totally different) between the v instances of the concept represented by these histograms (Table I provides a matrix example for the concept URBAN OBJECT).
- (2) A bin-to-bin histogram distance D_{bin} .

Before actually computing the distance between H_a and H_b , it is necessary to define a way to hierarchically merge the different instances of the histograms into clusters (*i.e.*, instances of higher semantic levels). To this end, it is possible to build a dendrogram \mathcal{D} (Fig. 2) induced by \mathcal{M}^{dis} (modelling this merging hierarchy) by using the Ascendant Hierarchical Clustering (AHC) algorithm [15]. Each one of the s stages of \mathcal{D} corresponds to a particular semantic level.

Once the dendrogram \mathcal{D} has been built, the HSB distance can be computed. This computation is organized in two main steps:

– **Step 1. Hierarchical bin-to-bin sub-distances computation** During an iterative merging process (scanning each stage of the dendrogram from the leaves to the root), the histograms H_a^k and H_b^k associated to H_a and H_b , which are induced by the merging of the instances composing each cluster of the stage S_k , are built (Fig. 1-①). After each iteration, the bin-to-bin sub-distance D_{bin} is then computed between the couple of (coarser) histograms H_a^k and H_b^k created previously. The resulting series of bin-to-bin sub-distances d^0, \dots, d^{s-1} enables to assess the histogram similarities at different semantic levels.

– **Step 2. Bin-to-bin sub-distances fusion** The bin-to-bin sub-distances d^k computed for all the stages of the dendrogram, and the “energy of merging” required to go from one stage to the next, are then fused into a function d^{inter} which is finally integrated to provide the HSB distance (Fig. 1-②).

II-B. Properties and complexity analysis

Metric properties As HSB can be defined as a sum of sub-distances d^k , it is equivalent to a sum of bin-to-bin distances D_{bin} . Thus, it inherits from the metric properties of D_{bin} (non-negativity, symmetry, identity, triangle inequality) and is then a distance.

Compliance with bin-to-bin distances It is possible to use HSB with a “partial” matrix \mathcal{M}^{dis} . Indeed, in the case where a proximity cannot be established between some of the instances, the associated values of \mathcal{M}^{dis} are set to 1. For such parts of the matrix, HSB will be equivalent to the bin-to-bin distance D_{bin} . *A fortiori*, HSB and D_{bin} will be fully equivalent if all the values of \mathcal{M}^{dis} are set to 1 (*i.e.*, if no background knowledge is available).

Computational complexity To compute HSB, it is first necessary to build the dendrogram \mathcal{D} . To this end, we use the AHC algorithm which has a complexity of $\Theta(v^2 \log v)$ when using a priority-queue algorithm (v is the number of basic instances). Since this operation is only performed once, its complexity can be considered as insignificant when comparing large datasets.

It is then necessary to compute the series of sub-distances d^0, \dots, d^{s-1} where s is the number of stages in \mathcal{D} . The complexity of the computation of each d^k is linked to the one of D_{bin} which is, in general, $\Theta(v)$ where v is the number of bins in the histograms. Thus, the complexity of HSB depends on both the number of bins v in the histograms and the number of stages s in \mathcal{D} .

Depending on the value of s , two extremal cases can occur:

- if \mathcal{D} is a fully balanced dendrogram, then $s = \log_2(v)$. The complexity of HSB is $\Theta(v + \frac{v}{2} + \frac{v}{4} + \dots + 1) = \Theta(v)$.
- if \mathcal{D} is a unbalanced dendrogram, then $s = v$ and the complexity of HSB is $\Theta(v + (v-1) + \dots + 1) = \Theta(v^2)$.

The computational cost of HSB is thus bounded by $\Theta(v)$ and $\Theta(v^2)$. Then, it requires more time/operations than the computation of classical bin-to-bin distances, but much less than the cross-bin distances that can require, in the worst cases, supercubic time. HSB can then be relevantly used to compare large histogram datasets.

III. EXPERIMENTS

III-A. Experimental context

Urban planning organizations need to follow the increase of urban settlements. To this end, it is necessary to map urban areas from satellite images at different scales/levels of analysis. Some efforts have been conducted to extract features from such images, in order to involve them into learning systems. The classical methodology consists of classifying the data into land cover classes by using supervised or unsupervised object-based approaches.

Recently, a hybrid segmentation/classification approach has been proposed [8] enabling to map urban areas at different scales (*i.e.*, districts, urban blocks, simple objects). It takes as input a multiresolution set of n satellite images and provides as output a set of n classification maps computed in a top-down fashion.

This approach classifies the segments extracted from an image (at a particular resolution) in terms of their semantic composition into the next resolution image. For instance, if we focus on the scale of the urban blocks, they can be characterized by their compositions in terms of basic urban objects (*e.g.*, individual houses, gardens, roads, *etc.*). A block U_i can thus be characterized by a histogram H_i where each bin is linked to an instance of URBAN OBJECT (Table I).

The main issue is to succeed in classifying into a same cluster similar objects that are not characterized by similar histograms. For instance, let us consider a block U_i characterized by a histogram H_i composed in particular of 21 red tile roofs and 4 slate roofs, and a block U_j characterized by a histogram H_j composed of 3 red tile roofs and 22 slate roofs. These blocks have to be grouped into a same class “Urban fabric with individual houses” since they are both composed of houses (with red roofs or slate ones).

We propose to validate the usefulness of HSBD by using it to improve the results provided by the approach described in [8]. In order to do so, we integrate HSBD into a post-processing step to re-classify these images in a bottom-up fashion.

III-B. Datasets and classification results

We consider three datasets (denoted **DATASET-1**, **-2** and **-3**) composed each of (1) a SPOT-5 MSR (9.6m) multispectral image and (2) a couple of QUICKBIRD images composed by a HSR (2.4m) multispectral image and a VHSR (60cm) pan-sharpened one. The HSR image of the **DATASET-1** is presented in Fig. 3(a).

Once the urban blocks have been extracted from the HSR image by using the hierarchical top-down segmentation approach proposed in [8], it becomes possible to characterize them by a “composition” histogram H_i which models the composition of the urban block U_i in terms of the distribution of the eleven instances of the semantic concept URBAN OBJECT in the VHSR image. The eleven instances considered in this experiment are listed in Table I.

To model the semantic relations between each bin of the histograms, a 11×11 dissimilarity matrix \mathcal{M}^{dis} has been provided by the expert (Table I). From this matrix, a 7-stage dendrogram has been built to model the merging order of the bins (Fig. 2).

We have applied the K -MEANS clustering algorithm, which does not require *a priori* parameters, to re-classify the urban blocks created previously. To process, the distance HSBD (associated to the Manhattan bin-to-bin sub-distance) has been directly integrated into the K -MEANS clustering algorithm to compare the classified histograms. From these datasets, we have chosen, in agreement

Table I. Dissimilarity matrix \mathcal{M}^{dis} associated to the instances of the concept URBAN OBJECT.

x_i	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11
x1 - Shadow	–	0.10	1.00	1.00	1.00	1.00	0.80	0.80	0.80	0.80	0.80
x2 - Water	–	–	0.50	0.50	1.00	1.00	1.00	1.00	1.00	1.00	1.00
x3 - Road	–	–	–	0.10	1.00	1.00	1.00	1.00	1.00	1.00	1.00
x4 - Railway	–	–	–	–	1.00	1.00	1.00	1.00	1.00	1.00	1.00
x5 - Bare soil	–	–	–	–	–	0.10	0.80	1.00	1.00	1.00	1.00
x6 - Herbac. veget.	–	–	–	–	–	–	0.80	1.00	1.00	1.00	1.00
x7 - Forest	–	–	–	–	–	–	–	1.00	1.00	1.00	1.00
x8 - Slate roof	–	–	–	–	–	–	–	–	0.10	0.40	0.90
x9 - Red tile roof	–	–	–	–	–	–	–	–	–	0.40	0.90
x10 - Gray tile roof	–	–	–	–	–	–	–	–	–	–	0.10
x11 - Metallic roof	–	–	–	–	–	–	–	–	–	–	–

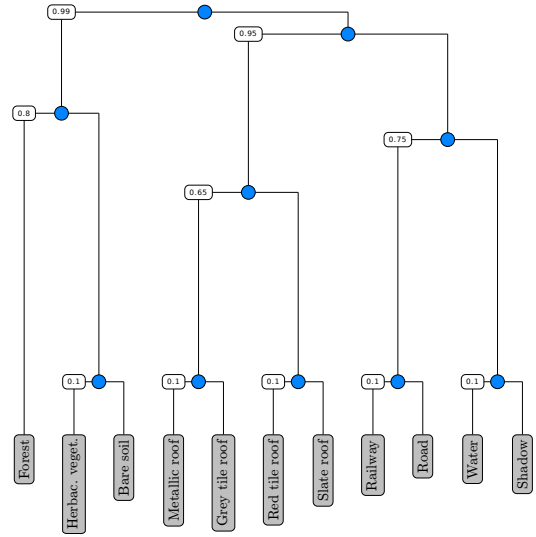


Fig. 2. Dendrogram associated to the dissimilarity matrix (Table I).

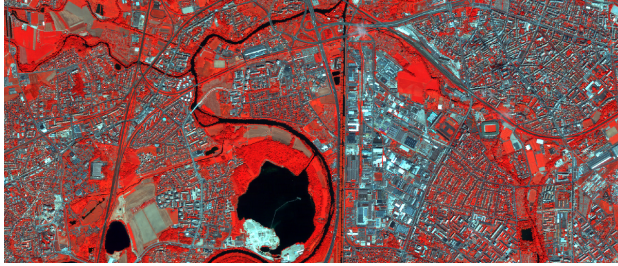
with the expert, to extract ten classes of urban blocks. Thus, the K -MEANS algorithm has been run with ten clusters.

Classification results obtained have then been compared to the ones obtained previously in [8]. To this end, all the results have been compared to certified ground-truth maps provided by the expert using the F-measure \mathcal{F} , and the Kappa \mathcal{K} indexes. Fig. 3(b) presents the ground-truth map of the **DATASET-1** at the urban blocks scale. These comparisons enable to assess the advantages of using HSBD instead of a classical bin-to-bin distance D_{bin} (used in [8]).

The results obtained are presented in Table II. From these different results, one can see that the proposed distance outperforms the classical bin-to-bin ones when comparing semantic nominal histograms. For a visual comparison purpose, the classification map obtained (initially) in [8] is presented in Fig. 3(c) while the one obtained by using the HSBD distance is proposed in Fig. 3(d).

III-C. Perspectives

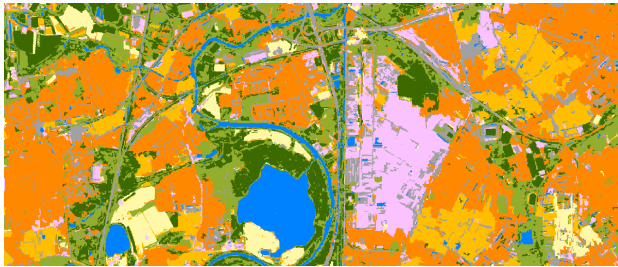
The HSBD distance enables to compare nominal histograms equipped with a dissimilarity matrix modelling the semantic proximity relations between the bins. Thanks to a hierarchical strategy, this distance can consider the multilevel semantic correlations



(a) HSR image (© DigitalGlobe Inc.).



(b) Ground-truth map (urban blocks level, 9 classes).



(c) Classification result obtained with [8].



(d) Classification result obtained with HSBD.

Fig. 3. Illustration of the **DATASET-1** and the associated results.

between the bins. Moreover, by opposition to cross-bin distances, it inherits from the low computational cost of bin-to-bin distances, while keeping the advantages of cross-bin ones, namely robustness to histogram translation and histogram bin size issues.

Experiments on multiresolution satellite images were carried out for validation purpose. The results have shown that HSBD can enhance the comparison of semantic nominal histograms and thus, the accuracy of classification results. Such validations emphasize the relevance and usefulness of HSBD for image processing tasks.

This work opens up several perspectives. In particular, we plan to study more formally the possible behaviors of the sub-distance function \mathbf{d}^k and to integrate an approach enabling to help the user for building the dissimilarity matrix.

Table II. Global evaluation results. F-measure and Kappa indexes.

Dataset	$\overline{F} \pm \sigma$		$\mathcal{K} \pm \sigma$	
	D_{bin} [8]	HSBD	D_{bin} [8]	HSBD
DATASET-1	0.65 ± 0.02	0.71 ± 0.01	0.76 ± 0.02	0.79 ± 0.01
DATASET-2	0.67 ± 0.02	0.72 ± 0.01	0.77 ± 0.02	0.86 ± 0.02
DATASET-3	0.63 ± 0.01	0.66 ± 0.01	0.73 ± 0.01	0.76 ± 0.01

IV. REFERENCES

- [1] J. N. Kapur, P. K. Sahoo, and A. K. C. Wong, "A new method for gray-level picture thresholding using the entropy of the histogram," *Comput. Vision Graph.*, vol. 29, no. 3, pp. 273–285, 1985.
- [2] R. Brunelli and O. Mich, "Histograms analysis for image retrieval," *Pattern Recogn.*, vol. 34, no. 8, pp. 1625–1637, 2001.
- [3] O. Chapelle, P. Haffner, and V. N. Vapnik, "Support vector machines for histogram-based image classification," *IEEE T. Neural Netw.*, vol. 10, no. 5, pp. 1055–1064, 1999.
- [4] M. Pietikainen and A. Rosenfeld, "Image segmentation by texture using pyramid node linking," *IEEE T. Syst. Man Cyb.*, vol. 11, no. 12, pp. 822–825, 1981.
- [5] R. Goffe, L. Brun, and G. Damiand, "Tiled top-down combinatorial pyramids for large images representation," *Int. J. Imag. Syst. Tech.*, vol. 21, no. 1, pp. 28–36, 2011.
- [6] C. Wemmert, A. Puissant, G. Forestier, and P. Gançarski, "Multiresolution remote sensing image clustering," *IEEE Geosci. Remote S.*, vol. 6, no. 3, pp. 533–537, 2009.
- [7] C. Kurtz, N. Passat, P. Gançarski, and A. Puissant, "Multiresolution region-based clustering for urban analysis," *Int. J. Remote Sens.*, vol. 31, no. 22, pp. 5941–5973, 2010.
- [8] C. Kurtz, N. Passat, P. Gançarski, and A. Puissant, "Extraction of complex patterns from multiresolution remote sensing images: A hierarchical top-down methodology," *Pattern Recogn.*, vol. 45, no. 2, pp. 685–706, 2012.
- [9] S. H. Cha and S. N. Srihari, "On measuring the distance between histograms," *Pattern Recogn.*, vol. 35, no. 6, pp. 1355–1370, 2002.
- [10] S. H. Cha, "Taxonomy of nominal type histogram distance measures," in *Proc. American Conference on Applied Mathematics*, 2008, pp. 325–330.
- [11] J. Hafner, H. S. Sawhney, W. Equitz, M. Flickner, and W. Niblack, "Efficient color histogram indexing for quadratic form distance functions," *IEEE T. Pattern Anal.*, vol. 17, 1995.
- [12] Y. Rubner, C. Tomasi, and L. J. Guibas, "The Earth Mover's Distance as a metric for image retrieval," *Int. J. Comput. Vision*, vol. 40, no. 2, pp. 99–121, 2000.
- [13] A. M. Tousch, S. Herbin, and J. Y. Audibert, "Semantic hierarchies for image annotation: A survey," *Pattern Recogn.*, vol. 45, no. 1, pp. 333–345, 2012.
- [14] Y. Ma, X. Gu, and Y. Wang, "Histogram similarity measure using variable bin size distance," *Comput. Vis. Image Und.*, vol. 114, no. 8, pp. 981–989, 2010.
- [15] J. H. Ward, "Hierarchical grouping to optimize an objective function," *J. Am. Stat. Assoc.*, vol. 58, no. 301, pp. 236–244, 1963.