A NEW ANGLE-BASED SPATIAL MODELING FOR QUERY BY VISUAL THESAURUS COMPOSITION*

Hichem Houissa, Nozha Boujemaa

INRIA-Rocquencourt, Domaine de Voluceau, BP 105 - 78153 Le Chesnay
{Hichem.Houissa, Nozha.Boujemaa}@inria.fr

ABSTRACT

Querying by Visual Thesaurus (VT) is a novel paradigm for content-based image retrieval approaches for it gives the user the possibility, in case of inappropriate starting example, to compose his query by arranging the visual patches of the starting "page zero" according to his mental image. A refinement of the willed results can be achieved by inducing a spatial description within the retrieval procedure. This paper presents a novel approach to model the spatial relations between the visual patches.We define the Weighted Angle Spatial Histogram (WASH) that combines the angular computation between pairs of regions of interest and their respective topological regularity/irregularity. WASH has shown great robustness to region shape and scale in the image because segmented regions are considered as a composition of elementary relevant and minor subregions. We tested our approach on generic database, and we compared it with other state-of-the-art techniques.

Index Terms— Image region analysis, Spatial reasoning

1. INTRODUCTION

The Query By Visual Thesaurus (QBVT) paradigm was introduced in [1] to compensate for the eventual lack of starting visual example. In fact, QBVT proposes to build a "page zero" composed of visual patches that summarize all the regions in the database. The user is then free to arrange these patches in the way that fits his mental representation of a given scene. In a previous work [2], we figured out the importance of using non traditional metrics -Earth Mover's Distance and Hausdorff distance- for comparing regions described by variable-size signatures. In fact, the regions we are dealing with are obtained by a coarse segmentation associated to fine description by means of Harris color points of interest. The regions are then clustered into visual categories using relational clustering in order to take benefit of the multi-dimensional description richness. Each category is represented by a visual patch standing for its best representative element. The set of patches composes the "page zero" that is provided to the user to compose his mental query. A binary composition of a query, for instance, query image I must **contain** regions R_i and R_j and **not** R_k is a first filtering process but is still inaccurate for it does not consider any spatial configuration of regions. Accordingly, we introduce the spatial relations through a novel spatial descriptor based on weighted angular computation between pairs of subparts of both regions to emphasize the most salient parts that are visually relevant for the user when retrieving a mental image. The spatial relations have been investigated in the literature to determine the relative position of 2 objects (especially for Geographic Information Systems) without any consideration of the context of the image nor the shape problem resulting from image segmentation. Beside 2D strings (Bounding boxes projections over X and Y axis), several Histograms were proposed: the histogram of angles was introduced by Majiyama [4] to determine the overall orientation between a query region A and a target region R. Later, Matsakis [3] defined the Histogram of forces (F-Histogram) that gives the orientation between parallel sections of A and R. Both histograms consider regions as independent entities while in the context of image retrieval, regions must necessarily be seen as a sets of subregions. This fundamental statement of fact implies the following requirements:

- Only a relevant subset of points contributes to determine the orientation between regions A and R,
- The shape of regions could bias the computation of the angle between A and R,

In order to meet these requirements, we propose a new weighted angular Histogram (called WASH) that handles the mean angle between 2 regions -provided only some parts of these regions- to determine the relative orientation between A and R. The complete description of the WASH construction is given is Section 3.

The rest of the paper is organized as follows: In Section 2, we review several Histogram-based techniques to determine the orientation between pairs of regions. Section 3 presents the details of construction of our novel Weighted Angle Spatial Histogram. In Section 4, we give some preliminary query results using spatial relations, comparisons with Angle and

*This research was supported by the European commission under Contract FP6-001765 aceMedia, http://www.acemedia.org/aceMedia.

Force Histograms are also discussed in this section. Concluding remarks and some pending issues are given in Section 5.

2. PREVIOUS WORK ON SPATIAL HISTOGRAMS

Spatial relationships between regions have been investigated first to address the problem of localization in Geographic Information Systems (GIS). The pairwise relationships between regions are assessed through the qualitative representation of spatial layout within a query image. First approaches were based on Allen's temporal interval relations and 2D-strings iconic indexing. These methods assimilate regions to elementary entities depicted either by their centroids or their minimum bounding rectangles. In addition, the ambiguity of spatial relations arises from the geometry of regions, thus requiring the use of fuzzy logic approach. The description by means of histograms was introduced by Miyajima and Ralescu [4] to cope with directional relations of an argument object A with respect to a reference object R using pairwise angle computation. In [3], Matsakis et. al proposed the Force Histogram that considers pairs of longitudinal sections instead of pairs of points. More variants were used in the literature for robot navigation. Also, in [5], Wang et al. introduced the R-Histogram in order to cope with labeled distances when computing relative orientation. In this paper, we propose a novel Histogram based on the orientation of one region regarding another, but considering that **(1)** only relevant subareas of each region are involved in the angle computation and **(2)** the (ir)regularity of the region is a hint measuring to which extent the computed orientation can be trusted.

3. WEIGHTED ANGLE SPATIAL HISTOGRAM

Although the angle values computed with the previous approaches are accurate in case of synthetic objects (geometric shapes), their computational complexity is prohibitive. In addition, these histograms show a weak robustness toward contours inaccuracy for coarsely segmented regions. Notice that the *regularity* of a region is an abstract concept depicting whether a given region is more likely to have a convex-like shape or completely irregular shape. Formally, the regularity is the ratio of the region area inside a given rectangle by the area of rectangle itself. Depending on the topological aspect of the considered couple of regions, we propose a histogram that meets the following requirement: **(1)** Only visually relevant parts of each region are considered and adaptively weighted during angle computation, **(2)** coarse segmentation resulting in rough and inaccurate contours is balanced by means of appropriate weighting functions.

3.1. WASH Histogram

The Weighted Angle Spatial Histogram (WASH) considers a given region as the sum of subparts that have variable size. Each subpart is a rectangle r_i adaptively embedded in the segmented region in a way it covers the maximum area of it as show in Fig.1.

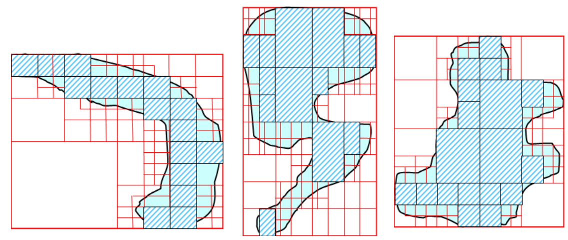


Fig. 1. Adaptive embedded rectangles construction

This adaptive grid construction has the benefit of emphasizing relevant subparts of each region of interest and getting rid of the biais induced by rough contours and coarse segmentation by assigning respective weights to each rectangle. These weights are exactly the elementary regularity of these rectangles and give more perceptual information than the global regularity of the whole region.

Let $G_A = \{(g_1^A, w_1^A), (g_2^A, w_2^A), \ldots, (g_N^A, w_N^A)\}$ and $G_R = \{(g_1^R, w_1^R), (g_2^R, w_2^R), \ldots, (g_M^R, w_M^R)\}$ be the set of centers of gravity of elementary rectangles covering regions A and R and their respective elementary weights. For each g_j^R, we define $\alpha = \{\alpha_1^j, \alpha_2^j, ..., \alpha_N^j\}$ as the set of pairwise angles between g_j^R et par g_i^A. The direction between g_j^R and region A is the weighted mean angle between g_j^R and all other g_i^A. The overall weighted angle between g_j^R and region A as follows:

$$ANG_R(j, A) = \frac{\sum_{i=1}^{N} (w_i^A * w_j^R) * \alpha_i^j}{\sum_{i=1}^{N} w_i^A * w_j^R} \tag{1}$$

$$\text{avec} \quad w_i^A = \frac{Area_A \cap MBB_i}{MBB_i} \tag{2}$$

Given that each weight w_i^A represents the topological regularity of a single cell, summing up all w_i^A, instead of considering the regularity of the whole region, is more robust to shape variations and gets rid of the effects of coarse segmentation. Typically, when segmented regions have irregular curves or unbalanced visual regions (see Fig.1), their overall regularity is not representative unless local weights are used (w_i^A, w_j^R).

Once the set of rectangles and their corresponding weights are defined, we build the WASH as a regular histogram bounded by $[-\pi, \pi]$ where each bin (corresponding to an angle obtained by quantization) is incremented by w_j^R to express the degree of local regularity of each elementary cell. As a matter of fact, if $ANG_R(j, A) = \theta$, the $WASH_\theta(A, R)$ is incremented as follows:

$$WASH_\theta(A, R) = \begin{cases} +w_j^R & if \quad ANG_R(j, A) \in [\theta - \frac{\tau}{2}, \theta + \frac{\tau}{2}] \\ 0 & else \end{cases} \tag{3}$$

Where τ is the step of quantization of the interval $[-\pi, \pi]$. We studied the importance of τ over the accuracy of WASH, we noticed that as far as the regions are convex-like (high regularity factor), the maximum still global, but once the angular computation involves complex and distorted shapes (low regularity factor), several local maxima appear denoting of partial contribution of subparts of the regions. We divide the segmented regions into 3 categories: pseudo-elliptical, pseudo-polygonal and irregular to figure out the error induced by the shape of regions. In Fig.2, we illustrate the influence of the quantization step over the accuracy of WASH. Each curve is associated to a pair of regions: 2 pseudo-elliptical (red curve), pseudo-polygonal Vs pseudo-elliptical (blue curves), irregular Vs. pseudo-elliptical (green curve) and irregular Vs. pseudo-polygonal (black curve).

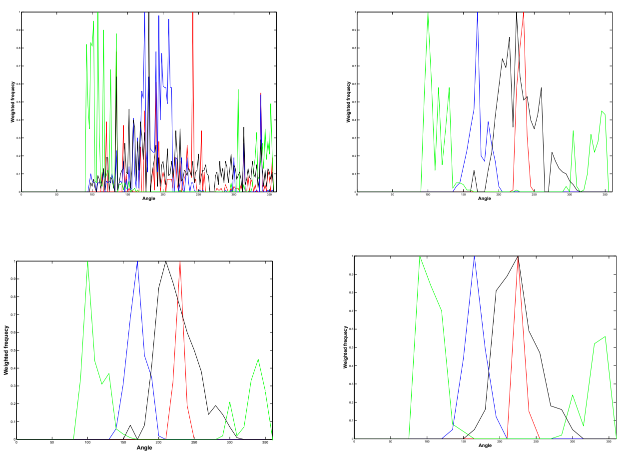


Fig. 2. The pickness of WASH varies depending on the quantization step (respectively 2, 5, 10 and 15 degrees)

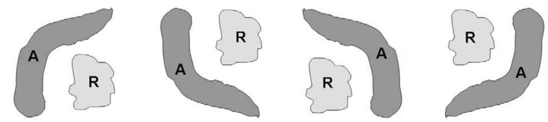
Notice that the Fig.2 points out the problem of confidence in the measures of WASH. We argue that the coarse segmentation of regions combined to the great variability in topological regularity make it difficult to rigorously assert about the angular value returned by WASH. Hence, we introduce several trigonometric functions referred as Degrees of Confidence in order to "fuzzify" our spatial histogram.

3.2. Degree of Confidence (DoC) of WASH

Like most other histograms, WASH provides a wide description of the relative orientation between 2 regions i.e. the direction of an argument region A toward a referent region R is given by a set of angular values over a quantized interval (typically $[-\pi, \pi]$). As far as the spatial localization problem is addressed, we can consider that the orientation between 2 regions is given by

$$angle(A, R) = arg \max_{\theta \in [-\pi, \pi]} WASH_\theta(A, R) \tag{4}$$

According to Fig.3, we notice that each pair of regions (A, R) is better described by, at least, a pair of relations among: ***A*** $\{$*is right of, is left of, is under, is above*$\}$ ***R***. For instance, an argument region A can be on the one hand on the right of a referent region R but also above it (Fig.3(c)).



(a) A is left-above R (b) A is left-under R (c) A is right-above R (d) A is right-under R

Fig. 3. Regions are often described by a pair of directions depending on their spatial configuration

Accordingly, every angle given by WASH can be interpreted as the combination of two directions. Each direction is trusted using a fuzzy weighting function that derives from usual trigonometric *(cosine and sinus)* functions modulated by the overall regularity of the considered regions. These directions (RIGHT, LEFT, UNDER, ABOVE) are bounded by $[-\pi, \pi]$. In other terms, if $WASH_\theta(A, R)$ returns θ_{max} as the major orientation between A and R, we assign a degree of confidence (DoC) value to θ_{max} revealing its relative relevance in favor of the directional predicates. As both regions are involved in the confidence measure through their respective regularity $DoC(A, R) = W_R * W_A$, the more regular they are, the higher (up to 1) DoC is and vice versa. This yields either to sharpen or flatten the trigonometric original functions as shown in Fig.4. As a matter of fact, if both regions are regular, DoC is close to 1 (directional functions tend to be window functions) denoting of a great confidence in the computed measure. Inversely, if the regions are quite irregular, DoC drops near to 0 (directional functions are likely to be Dirac functions) inducing a weak confidence in pairwise angle. (Fig.4).

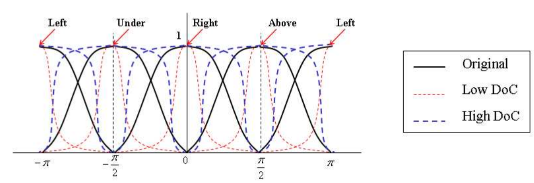


Fig. 4. DoC affects the shape of the trigonometric functions according to the regularity of the compared regions

4. EXPERIMENTAL RESULTS AND EVALUATION

We compared WASH to Force Histogram and Angle Histogram in terms of the computational time. The time issue is very important for we have to cope with hundreds of pairs of regions. Moreover, the relevance of computed angles depicted by their confidence measures determines the orientation between an argument region and a referent one. The different methods were implemented in C++ programming language on a 1.8GHz. We perform the computation for a subset of

700 images taken from a generalist database **(IDS)**[1] containing landscapes, seascapes, outdoor and natural scenes.

	reg. per img.	Angle Histo	Force Histo	WASH
Agriculture	6	3.27	2.72	1.29
Mountain	5	3.55	2.17	1.79
Architecture	7	2.68	2.25	1.92
Seascape	5	2.88	2.46	2.02

Table 1. Mean CPU time computation for Angle Histogram, Force Histogram and WASH

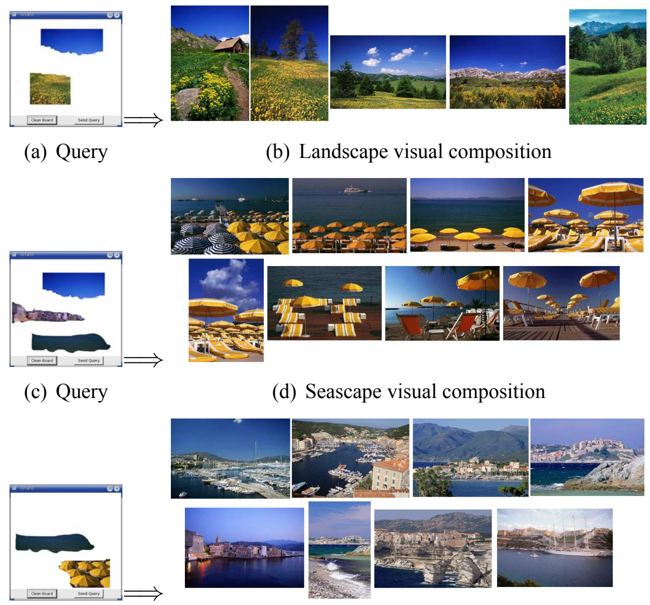
Table 1 summarizes the computational CPU time necessary to compute Force Histogram, Angle Histogram and WASH. Indeed, it is a mean CPU time over all images categorized into 4 classes: agriculture, mountain, architecture and seascape. We note that the more regular regions are, the faster computation of WASH is. Moreover, all these methods perform well for simple geometrical-like regions. It is important to note that the advantage of WASH over Force and Angle Histograms is that we compute only a small set of pairwise orientations. In fact, instead of computing the angular directions between all N and M pixels of regions A and R respectively (angle histogram) or between all sections for variable values of θ (force histogram), we confine the angular information between 2 regions to the weighted orientation between pairs of centers of gravity of elementary rectangles (respectively $n \ll N$ and $m \ll M$). These value are also modulated in order to alleviate the drawbacks of coarse segmentation. Furthermore, for force histogram, we still require to define longitudinal sections as many times as the quantization of $[-\pi, \pi]$ which is prohibitively expensive and time consuming.

A typical case of joint use of WASH and the Visual Thesaurus could be: the user is proposed a Visual Thesaurus summarizing all regions of the database in categories represented by their prototype. He selects the visual patches he is interested in and arranges them on a white board in a way that corresponds to his mental image. We illustrate some scenarios in Fig.5 where images are ranked first after a category filtering using sets intersection theory and then with respect to increasing histogram intersection distance.

5. CONCLUSION

We have proposed a novel Weighted Angle Spatial Histogram (WASH) that overcomes the heavy computational cost of the angle histogram and the integral complexity of the force histogram. Moreover, WASH introduces the adaptive elementary rectangles and their respective weights (regularity) in order to compensate for the drawbacks of rough segmentation and also to capture only visually salient parts of the regions. The fuzzy Degrees of Confidence enhance more robustness to WASH construction. The user arranges the visual patches (selected from the Visual Thesaurus) in a spatial configuration that fits to his mental image; the retrieved results are visually and spatially similar to the composed query image. We could further refine this configuration descriptor and handle distance between regions. In fact, we would like to address the metric information and especially focus on the fact that the farther regions are, the more reliable and confident WASH computation is.

[1] http://www.imagedusud.com



(a) Query (b) Landscape visual composition

(c) Query (d) Seascape visual composition

(e) Query (f) Harbor and mountain visual composition

Fig. 5. Retrieval results using both visual thesaurus and spatial relations

6. REFERENCES

[1] J. Fauqueur and N. Boujemaa. New image retrieval paradigm: logical composition of region categories. In *IEEE International Conference on Image Processing (ICIP)*, pages 601–604, 2003.

[2] H. Houissa and N. Boujemaa. On the use of metrics for multi-dimensional descriptors clustering. In *IEEE International Conference on Image Processing 2006*, 2006.

[3] P. Matsakis and L. Wendling. A new way to represent the relative position of areal objects. *IEEE Trasactions on Pattern Analylsis ad Machine Intelligence*, 7(21):634–643, 1999.

[4] K. Miyajima and A. Ralescu. Spatial organization in 2d segmented images: representation and recognition of primitive spatial relations. *Fuzzy Sets and Systems*, 2-3(65):225–236, 1994.

[5] Y.Wang, F. Makedon, J. Ford, L. Shen, and D. Goldin. Generating fuzzy semantic metadata describing spatial relations from images using the r-histogram. In *ACM/IEEE-CS joint conference on Digital Libraries*, pages 202–211, 2004.