

A Geometric Data Structure Applicable to Image Mining and Retrieval

T.Iwaszko, M.Melkemi, and L.Idoumghar

Université de Haute Alsace,
LMIA-MAGE, 4 Rue des Frères Lumière,
68093 Mulhouse, France

{thomas.iwaszko,mahmoud.melkemi,lhassane.idoumghar}@uha.fr

Abstract. Due to improvements in image acquisition and storage technology, terabyte-sized databases of images are nowadays common. This abundance of data leads us to two basic problems: how to exploit images (image mining)? Or how to make it accessible to human beings (image retrieval)? The specificity of image mining/retrieval among other similar topics (object recognition, machine vision, computer vision, etc.) is precisely that their techniques operate on the whole collection of images, not a single one. Under these circumstances, it is obvious that the time complexity of related algorithms plays an important role. In this paper, we suggest a novel general approach applicable to image mining and retrieval, using only compact geometric structures which can be pre-computed from a database.

1 Introduction

In recent years, the amount of “non-standard” or multimedia data (in contrast with standard alphanumeric data) has greatly increased. Terabyte-sized databases of images are now available for various purposes: medicine, astronomy, physics, etc. but also digital photography: monitoring, online photo albums or entertainment.

In general, the problem of extracting implicit relevant information has already been studied for decades by researchers from *data mining*. However, and as described in [1], data mining techniques are not sufficient or fully appropriate for *image* databases. Singularities of information in images make necessary the design specific techniques and tools. These are being developed in the young area of *image mining* [2, 3].

The other way to deal with such large collections of images is to make them easily accessible to human beings. To tackle this problem of *image retrieval*, one must provide a user interface to make the collection browsable, and a relevant method for specifying search queries. In addition, the image collection has to be ordered (*indexing*) in such way the system can quickly compute database matches with user queries. Finally, the user should be able to give feedback on

the relevance of the results so the searching engine can possibly improve its performance afterwards. Content-Based Image Retrieval systems [4, 5] i.e. CBIR systems are the realization of these ideas.

Image mining and image retrieval share the fact they both operate on whole collections of images, in contrast to fields of *object recognition*, *machine vision*, *computer vision*, *etc.* which analyse a single image, try to recognize a single scene. Consequently, the time complexity of image mining/retrieval-related algorithms must be taken into account, as well as the size of intermediate representation.

Until now, many indexing techniques have been reported in the literature [6, 7]. Thanks to an indexing schema, it is possible to filter the complete list of elements in the database, in order to reduce the actual number of considered images.

In this paper, we suggest a general approach for working with image collections, which can be seen as an alternative or a complement to indexing. It offers the possibility to reduce the dataset by converting images to sets of points. This compact representation along with a pre-computation step might speed-up detection of spatial patterns in image mining or retrieval techniques. We present a geometric data structure, variant of the Voronoi diagram, for recording in advance locations of empty shapes (i.e. spatial patterns) and thus saving time on later treatments.

The next section deals with the “feature extraction step” for converting raw image data to geometric data. Section 3 introduces the notations used throughout the paper. A particular shape representation is presented in section 4. Then, in section 5, we define the new geometric structure based on shapes. We outline computation and present some results in section 6. Finally, we conclude by describing future work and other possible applications in the last section.

2 Image Analysis and Computational Geometry

Recently in image analysis, some research has been done in order to detect so-called *interest points* in images. Interest points are sometimes called Spatial Interest Pixels, [8]. Intuitively, an interest point corresponds to a pixel that has stronger interest in strength than most of pixels in an image. Interest point detection is often a particular form of edge/corner detection, but it can also concern search methods in the color space [9]. For a list and evaluation of interest point detectors, see [10]. This method constitutes a fast pre-processing step and allows to work on more compact representations. A remarkable advantage is that it can be combined easily with other tools of image analysis (histograms for instance, as demonstrated in [11]) or computer science.

In connection with point sets, computational geometry (or CG) is a field devoted to the study of algorithms which can be stated in terms of geometry. The algorithms and data structures (e.g. the Voronoi decomposition) of CG are designed for efficiency and have found numerous applications in various fields of computer science, in particular: image processing [12], analysis, indexing [9], retrieval [4].

In this paper, we present a novel data structure, based on points and geometric shapes, suitable for image mining/retrieval tasks. This structure works with specified models of shape, but these models could be learned on labelled images as well.

Let us first consider the abstract problem and the structure in general before showing its application to the field of Image Analysis. Let S be a set of points in the plane. Being given a plane geometric shape, (that is, an open bounded region of \mathbb{R}^2) is it possible to translate and rescale the shape in such way it has at least one point of S on its boundary, while remaining empty? More generally what is the set of solutions to the problem? i.e. how to locate all the free spaces for fitting a particular shape into the set of points? The shape representation and geometric structure introduced in the following bring an answer to this question. The structure can be seen as a variant of the Voronoi diagram.

3 Notations and Basic Terminology

In this work, we use the following notations:

- pq : the euclidean distance between p and q
- $[pq]$: the segment of extremities p and q
- $\mathbf{b}(x, r)$: the *open disc* of radius r centered on x . Its boundary is a *circle*, we note it $\partial\mathbf{b}(x, r)$. Mathematically:

$$\begin{aligned}\mathbf{b}(x, r) &= \{y \in \mathbb{R} \mid yx < r\} \\ \partial\mathbf{b}(x, r) &= \{y \in \mathbb{R} \mid yx = r\}\end{aligned}$$

- $R(p)$: the Voronoi region of a point p of S , S being a given set of points. Mathematically, we have:

$$R(p) = \{x \in \mathbb{R}^2 \mid px \leq qx, \forall q \in S\}$$

- $L = (l_1, \dots, l_n)$ denotes a n -tuple (i.e. a sequence) of objects, where l_1 is the first one, l_2 the second one, etc.

Moreover, we introduce the following terms:

- Given a set of points S , an open bounded region A is said to be *empty* if and only if $A \cap S = \emptyset$.
- We call *weighted point* a pair constituted by a point and a real positive number, formally: $w = (p, r)$ where $p \in \mathbb{R}^2, r \in \mathbb{R}$ and $r > 0$.

4 Shape Representation

For convenience, in this section we shall abuse language slightly: Given a tuple T of weighted points, a *disc of T* refers to an open disc $b(c, r)$ where (c, r) is an object of T .

Definition 1 requires two preliminary concepts introduced below. Given $T = ((c_1, r_1), (c_2, r_2), \dots, (c_m, r_m))$ a m -tuple of weighted points, it is accepted that:

- Two discs $b(c_i, r_i)$, $b(c_j, r_j)$ of T are said to be **adjacent discs** iff: the smallest has its center on the boundary of the biggest, i.e.

$$\begin{cases} c_j \in \partial b(c_i, r_i) & \text{if } r_j \leq r_i \\ c_i \in \partial b(c_j, r_j) & \text{otherwise} \end{cases}$$

- Let $V = \{c_1, \dots, c_m\}$ be the set of all the points listed in T . Let E be the set of segments $[c_i c_j]$ such that the discs $b(c_i, r_i)$ and $b(c_j, r_j)$ of T are adjacent discs. The resulting straight-line graph (V, E) is **the adjacency graph** of T . An example of an adjacency graph is shown on Fig. 1.

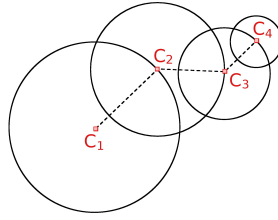


Fig. 1: Representation of a shape-parameter list and its adjacency graph (dashed line)

Definition 1. A *shape-parameter list* $\mathcal{C} = ((c_1, r_1), (c_2, r_2), \dots, (c_m, r_m))$ is an m -tuple of weighted points which satisfies the three conditions:

1. The two first discs of \mathcal{C} are adjacent discs
2. The adjacency graph of \mathcal{C} is connected
3. Let \mathcal{C}' be the list of the k first discs of \mathcal{C} , where $1 < k < m$. The adjacency graph of \mathcal{C}' is also connected.

Definition 2. Given a shape-parameter list $\mathcal{C} = ((c_1, r_1), \dots, (c_m, r_m))$, we define the **shape-model** $\mathfrak{p}_m(\mathcal{C})$ as being the open bounded region obtained by the union of all the discs of \mathcal{C} :

$$\mathfrak{p}_m(\mathcal{C}) = \bigcup_{i=1}^m b(c_i, r_i)$$

The centers of the two first discs of \mathcal{C} (i.e. the points c_1, c_2) are called **reference points** of the shape-model $\mathfrak{p}_m(\mathcal{C})$.

An example of a simple shape-model is shown on Fig. 2a (its shape-parameter list is represented on Fig. 2b).

As we will see in section 6, that representation allows us to create complex shape-models and even good approximations of real objects silhouettes. Some examples are given on Fig. 3.

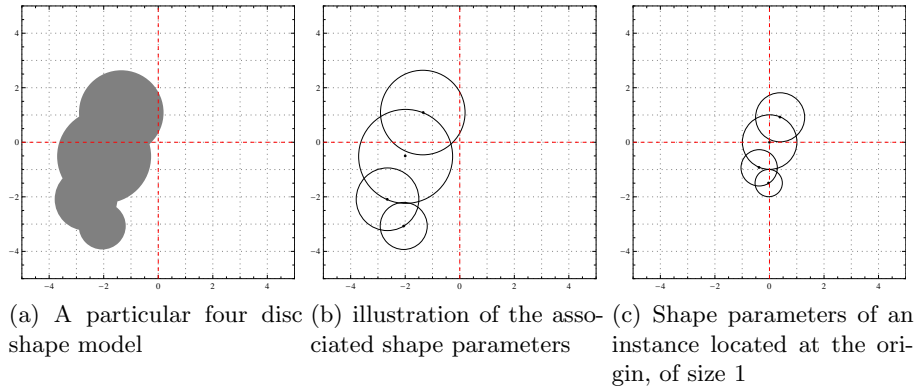


Fig. 2: Concepts for shape representation

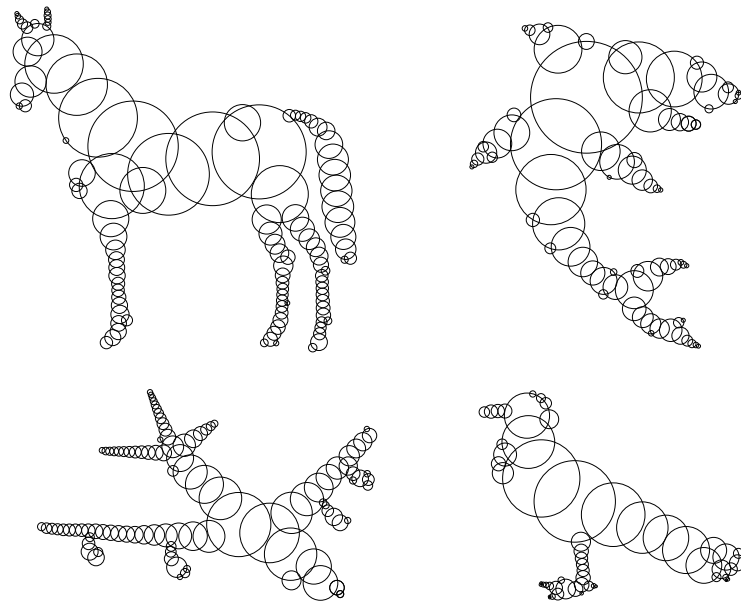


Fig. 3: Sophisticated shape-models, approximating real-world 2D pictures

Definition 3. Given a shape-model $\mathbf{p}_m(\mathcal{C})$, we call **instance of $\mathbf{p}_m(\mathcal{C})$ located at x and of size λ** the region of the plane defined by:

$$\mathbf{p}_m(\mathcal{C}, x, \lambda) = \bigcup_{i=1}^m b(c'_i, r'_i)$$

Where c'_i, r'_i are given by:

- $c'_1 = x$, and $r'_1 = \lambda$
- $c'_i = x + \alpha(c_i - c_1)$, and $r'_i = \alpha r_i$ for $2 \leq i \leq m$
- α is the rescaling factor: $\alpha = \frac{\lambda}{r_1}$, and λ the resulting size.

For short, in the following we shall use the term **instance** for referring to: instance of a shape-model located at a certain point and of a new size.

An illustration of this concept is shown on Fig. 2c.

Interpretation: Given a shape-model $M_1 = \mathbf{p}_m(\mathcal{C})$, its instance is a new shape-model $M_2 = \mathbf{p}_m(\mathcal{C}, x, \lambda)$, which has x and $x + \alpha(c_2 - c_1)$ for reference points while being similar to M_1 (similar: there exists an affine transformation that takes M_2 to M_1).

Interpretation of the Calculi: The coordinates c'_i and numbers r'_i are defined in order to perform an affine transformation which combines: translation and homothety (no rotation). This transformation is fully parameterized by x, λ .

5 Regions of Expanded Empty Shape-models

We have defined shape-models precisely. Thanks to this preliminary work, given a shape-model, new instances can be computed. An instance is parameterized by a point and a real number, and the original shape parameters are known. Thus all the geometric information (boundary, disc overlap, etc.) is computable. We can determine whether a particular instance is empty or not, has a point on its boundary or not, etc.

Definition 4. Given S a set of points of the plane, and $\mathbf{p}_m(\mathcal{C})$ a m -disc shape-model, we call **region of expanded empty shape-model** associated to $p \in S$ (REES of $p \in S$ for short) the region defined such that:

$$R_{\mathcal{C}}(p) = \{x \in \mathbb{R}^2 \mid \mathbf{p}_m(\mathcal{C}, x, px) \cap S = \emptyset\}$$

Fig. 4 illustrates this definition with few simple shape-models and associated regions.

Intuitively, $R_{\mathcal{C}}(p)$ represents the locations $x \in \mathbb{R}^2$ where the shape-model $\mathbf{p}(\mathcal{C})$ can be translated (ie. its first center becomes x) and expanded until it has p on its boundary, while remaining empty of S .

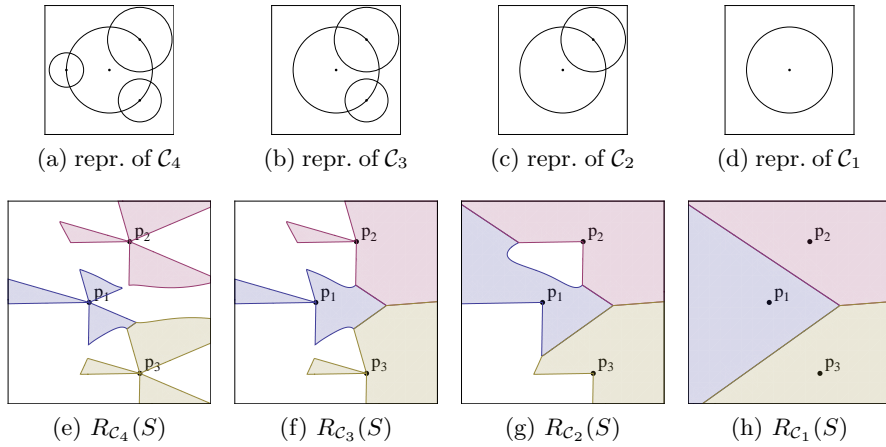


Fig. 4: Trivial shape-models and associated regions for a 3-point set. The regions are the set of points where the shape-model can be translated then rescaled until it has a site on its boundary, while remaining empty.

6 Practical Results

Shape-models as presented in this paper can be build from 2d silhouettes as shown on Fig. 5. Actually, that disc-based shape representation has already been introduced in previous work. Despite of the fact the representation has been slightly modified and reformulated since then, the construction of approximations from silhouette remains identical. For details on this process, see [13].

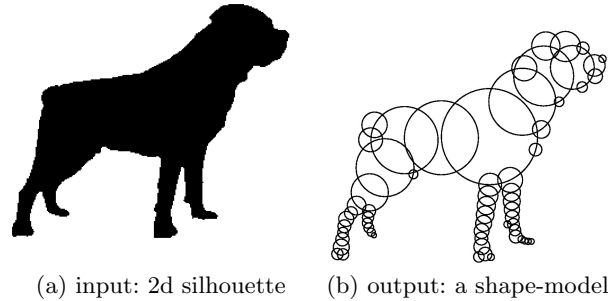


Fig. 5: Building process of a shape-model for approximating a given silhouette as well as possible (notably using the so-called medial axis or topological skeleton)

In the section 2, several methods for interest points detection, along with the possibility to used it as a pre-processing and combine it with other tools, were mentionned.

Accordingly, we can use our geometric structure with images. After choosing a suitable method for interest points detection (“suitable“ would be application-specific), a whole image collection used for data mining/retrieval can be converted in advance into a collection of point-sets.

Firstly we have implemented the well-known Harris detector [14] and ran the algorithm on grayscale images in order to find mostly spatial pixels of interest. Some results are shown on Fig. 6. Note that by pre-computing point sets for each image in a database, it is possible to save space and the structure proposed previously can still work on these points sets for detecting spatial pattern (empty of points).



Fig. 6: Computation of pixels of interest using the Harris detector on grayscale images, in order to convert images to simple point sets

Having computed both point sets and shape-models from silhouettes, the computation of REES can be made. The definition mentioned in previous sections being equivalent to a system of inequalities (to test the emptiness of the shape is to test the distances between centers of discs constituting the shape-model and points/sites), the calculation boils down to approximating each region using linear algebra and algebraic methods, like the resultants. Thus each elements of the system is simply considered as the part of a two variable polynomial.

The result of such regions computation is shown on Fig. 7.

For application to CBIR systems, geometric models could be specified in advance and the user would select one and indicate its approximate location in the picture he is searching. We have chosen this scenario but other possibilities are offered by the geometric structure. Two facts are worth of interest with this approach:

- The computation of regions is still a part of the pre-computation (before any user query)
- The classic matching step is replaced by a simple point in polygon algorithm (REES being approximated by polygons).

Therefore good performances are expected but this cannot be strictly demonstrated yet.

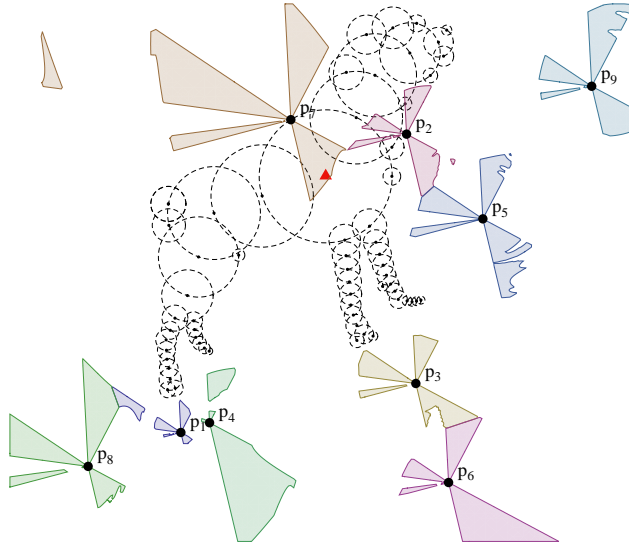


Fig. 7: REES and an empty instance. The triangle represents the location of the instance's reference point. According to the definition, it is known the whole shape-model is empty of points iff its reference point is located inside a region $Rc(p)$

Similarly, at this point of our work it was not possible to make comparisons with other CBIR methods as a whole framework integrating all steps is required (image to point sets conversion, shape-models construction and computation of regions). However in this section we have presented meaningful results already obtained for each separate step.

7 Conclusion

This paper presented a theoretical structure and explained in what way it could be used for image mining and retrieval. The actual computation of this structure relies on algebraic calculus. Indeed, the introduced REES can be decomposed into simpler regions (just like the Voronoi diagram can be decomposed in halfplanes). Each simpler subregion can then be expressed with an inequation. All in all, the computation boils down to algebraic system of inequations solving.

Currently we use the computational software *Mathematica* for this task. Our implementation produced the illustrations presented in this paper and scales up well, up to hundreds of discs and points.

If fully described and developed, this new structure might have numerous applications, like the classic Voronoi diagram, because problems involving proximity informations are general, and found in many of areas of science. New shape-based algorithms for image mining and retrieval could be designed. The presented structure could also be worth of interest for robotics path planning problems.

As future work, we aim at: achieving the REES computation and coding in a stand-alone application. It would let us studying precise time complexity for both region construction and point query. The most interesting future prospect is the setting of a whole CBIR framework (using together all the steps presented in the previous section) in order to test the structure in practice and notably compare it to existing techniques.

References

1. Wynne Hsu, Mong Li Lee, and Ji Zhang. Image mining: Trends and developments. *J. Intell. Inf. Syst.*, 19(1):7–23, 2002.
2. Ji Zhang, Wynne Hsu, and Mong-Li Lee. An information-driven framework for image mining. In *DEXA '01: Proceedings of the 12th International Conference on Database and Expert Systems Applications*, pages 232–242. Springer-Verlag, 2001.
3. Michael C. Burl, Charless Fowlkes, and Joseph Roden. Mining for image content. In *in Systemics, Cybernetics, and Informatics / Information Systems: Analysis and Synthesis*, 1999.
4. Remco C. Veltkamp, Mirela Tanase, and Danielle Sent. Features in content-based image retrieval systems: a survey. In *State-of-the-Art in CBIR*, pages 97–124. Kluwer, B.V., 2001.
5. Ritendra Datta, Jia Li, and James Z. Wang. Content-based image retrieval: approaches and trends of the new age. In *MIR '05: Proceedings of the 7th ACM SIGMM international workshop on Multimedia information retrieval*, pages 253–262. ACM, 2005.
6. Tzi-cker Chiueh. Content-based image indexing. In *VLDB '94: Proceedings of the 20th International Conference on Very Large Data Bases*, pages 582–593. Morgan Kaufmann Publishers Inc., 1994.
7. D. Doermann. The indexing and retrieval of document images: A survey. *CVIU*, 70(3):287–298, June 1998.
8. Qi Li, Jieping Ye, and Chandra Kambhampettu. Spatial interest pixels (sips): useful low-level features of visual media data. *Multimedia Tools Appl.*, 30(1):89–108, 2006.
9. Yi Tao and William I. Grosky. Spatial color indexing: A novel approach for content-based image retrieval. In *ICMCS'99*, pages 530–535. IEEE Computer Society, 1999.
10. Cordelia Schmid, Roger Mohr, and Christian Bauckhage. Evaluation of interest point detectors. *Int. J. Comput. Vision*, 37(2):151–172, 2000.
11. Fan-jie Meng, Bao-long Guo, and Lei Guo. Image retrieval based on 2d histogram of interest points. In *IAS '09: Proceedings of the 2009 Fifth International Conference on Information Assurance and Security*, pages 250–253, Washington, DC, USA, 2009. IEEE Computer Society.
12. Qiang Du, Max Gunzburger, Lili Ju, and Xiaoqiang Wang. Centroidal voronoi tessellation algorithms for image compression, segmentation, and multichannel restoration. *J. Math. Imaging Vis.*, 24(2):177–194, 2006.
13. L. Idoumghar and M. Melkemi. Pattern retrieval from a cloud of points using geometric concepts. In *Proc. of ICIAR'07*, pages 460–468. M. Kamel and A. Campilho Eds, 2007.
14. C. Harris and M. Stephens. A combined corner and edge detector. In *Proceedings of the 4th Alvey Vision Conference*, pages 147–151, 1988.