

Retrieving Shapes Efficiently by a Qualitative Shape Descriptor: The Scope Histogram

A. Schuldt, B. Gottfried, and O. Herzog

Centre for Computing Technologies (TZI)
University of Bremen, Am Fallturm 1, D-28359 Bremen

Abstract. Efficient image retrieval from large image databases is a challenging problem. In this paper we present a method offering constant time complexity for the comparison of two shapes. In order to achieve this, we extend the qualitative concept of positional-contrast by 86 new relations describing the position of a polygon w. r. t. its line segments. On this basis a histogram of the relations' frequencies is computed for each shape. A useful property of our approach is that, due to the underlying concept of positional-contrast, it can be intuitively decided whether its combination with other features is promising. Especially, retrieval results of about 64% are achieved in the MPEG test with constant time complexity.

1 Introduction

Recent developments show an increasing spread of digital technologies in many areas of economic, scientific, and even personal life. This process often comes along with the application of large image databases. Their management requires feasible search technologies. While a great deal of work addresses issues concerning segmentation, grouping, and feature extraction, less emphasis has been put on the representation of extracted objects so as to support efficient retrieval processes. In this paper we will concentrate on this latter topic bringing forward results in the context of qualitative representations which advance image and video retrieval systems in that they allow objects to be indexed concisely.

While colour and texture have been successfully applied to retrieval systems, shapes pose major difficulties (cf. [11]). Being able to resort to many polygonal approximation algorithms [13], and forming quite a compact description for two-dimensional outlines even with little influence on the perception of shape [2], we shall focus on the description of polygons. We use especially [12], thereby choosing a scale-invariant approximation error of one percent of a polygon's perimeter.

For the search of objects in image databases there are concurrent requirements to be satisfied: On the one hand users demand for adequate responses matching their query, on the other hand users require responses within an appropriate period of time. While the first requirement is well fulfilled for polygons [9], the second one still remains a major difficulty for large image databases due

to the time complexity needed when searching for and comparing many objects. Useful are therefore approaches that are based on simple features which only need constant time complexity when comparing objects. Although these features are rather limited for the purpose of object classification, they can be combined in order to improve retrieval results. Examples for such features include moments [8], compactness [3], radius ratio [4], and aspect ratio [3]. In this paper, we propose yet another feature which equally concisely characterises single objects, namely scope histograms, extending the approach which has been introduced in [6]. It shows, that in the MPEG test our new approach outperforms the approaches mentioned above, although still allowing to compare objects with constant time complexity. Due to its completely different representation, it is promising to combine our new approach with the other numerical features. This allows retrieval results of about 64% to be achieved in the MPEG test, staying behind [9] only twelve percentage points. However, they propose a much more complex approach in order to obtain better results, requiring $O(mn^3)$ for comparing two outlines.

In the following section we will shortly review previous work on which scope histograms are based. In Sect. 3 we introduce our new approach, which offers constant time complexity for comparing two shapes. Afterwards, in Sect. 4 the approach is compared and combined with others, and finally, we give a short summary in Sect. 5.

2 Characterising Shapes Using Positional-Contrast

The notion of positional-contrast [6] is based on the extension of the 13 qualitative relations between time intervals [1] to two dimensions. In order to distinguish two-dimensional positions between line segments qualitatively, e. g. positions between line segments of polygons, [6] uses the orientation grid which has been suggested by [15]. The orientation grid is induced by a line segment connecting two points and it consists of three lines: One line passes both points, the other two lines are oriented orthogonally w. r. t. the first one. Each of them passes one of the two reference points. This divides the plane into six sectors, allowing a third point to be located within this grid. Not only single points, but also line segments can be described by their relative position to another reference segment which induces the orientation grid. Every line segment is defined by a start and an end point. Since each of these points can be located in any of the six sectors of the orientation grid, there exist $6^2 = 36$ arrangements between two line segments that can be distinguished qualitatively. According to [5], this number can be reduced to 23, due to symmetries and the omission of intersections. The remaining 23 distinguishable bipartite arrangements, in short \mathcal{BA}_{23} , are depicted on the left hand side of Fig. 1. The relations' mnemonic labelling is shown in the centre of Fig. 1.

Applying \mathcal{BA}_{23} , it is not only possible to characterise the relative position of two single line segments. Moreover, one can describe the complete course of the polygon w. r. t. a reference segment. This is accomplished by a sequence of n \mathcal{BA} relations describing each of the n line segments, one after another [7]:

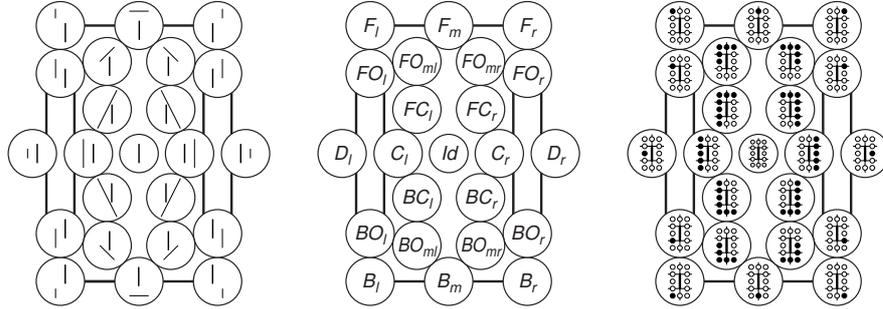


Fig. 1. Left: Example arrangements of \mathcal{BA}_{23} relations. Centre: The relations' mnemonic labels. Right: Iconic representation of the sets describing the relations' scopes

Definition 1 (Course) x is a line segment of a simple, closed polygon P . Its course $C(x)$ describes all \mathcal{BA} relations between all lines of P and x :

$$C(x) \equiv (x_{y_0}, \dots, x_{y_{n-1}}), x_{y_i} \in \mathcal{BA}, i = 0, \dots, n - 1$$

Up to now, a polygon has only been characterised w. r. t. one of its line segments. In order to gain a complete characterisation, it has to be described w. r. t. each of its segments. In doing so, we obtain a matrix containing n^2 relations. Thus, space complexity for this description is quadratic, $O(n^2)$. Time complexity for the comparison of two polygons which are described this way is even higher. Due to the fact that every matrix can be built up by starting with different reference lines, and that two matrices might be different in size (the second polygon comprises m line segments), time complexity for the comparison of two of these descriptions is $O(mn^3)$. In the following section we shall learn how space and time complexity can be reduced.

3 Reducing Time and Space Complexity

In the following, we elaborate on how time and space complexity of the approach described in the preceding Sect. 2 can be reduced. In our case, space complexity depends on the number of relations, which are needed in order to characterise a polygon. By time complexity we address the cost required for comparing two polygons. In order to reduce the complexity of our description, we start with proposing a new definition for the scope of polygons, which has been introduced in [7]. Later on, this will enable us to compute a histogram of all the scopes which are derivable from a polygon. Since two histograms can be compared with constant time complexity by computing the sum of the distances of their corresponding entries, we obtain a description offering both constant time and constant space complexity. Note that the number of entries in the histogram does not depend on the number of line segments of the polygons but on the number of scopes conceivable, which is a fixed number.

As seen before, a sequence of n relations is needed in order to describe a polygon's course, $C(x)$, by characterising the position of each of its n segments w. r. t. x . In proceeding this way we obtain a detailed description of the polygon with n^2 relations. This level of detail is not always needed. Instead, it often suffices to describe the position of a polygon as a whole w. r. t. a reference segment; Figure 2 illustrates this. While the courses of the depicted polygons are quite different, both courses have the same position w. r. t. their highlighted reference segments: They are solely located in the left half of the orientation grids induced by their respective reference segments.

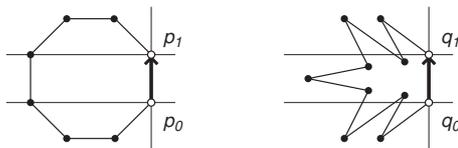


Fig. 2. Two different polygons, both located solely in the left half of the orientation grids induced by their respective reference segments

Instead of taking n relations as before, we now get along with a single relation in order to characterise the polygons. Thus, space complexity for this description is constant. When applied to both polygons depicted in Fig. 2, we obtain the same description, i. e. they cannot be distinguished. At first glance, this seems to be a major loss of information. Although the difference of both polygons is obvious for humans, they would be indistinguishable according to our new description. As we shall see below, this problem can be solved by describing polygons not only w. r. t. one of their line segments. It is rather necessary to use each line segment as a reference segment, inducing the orientation grid to each line segment, one after another. Doing this, we observe that the left polygon in Fig. 2 always lies on the left w. r. t. its reference segments. By contrast, the right example results in many different descriptions.

Hitherto, we described the position of polygons rather at a conceptual level, for instance, having denoted the relation in the preceding example “left”. In the following, it is necessary to determine a set of relations which is adequate for our purposes. The \mathcal{BA}_{23} relations introduced in Sect. 2 do not fulfill this requirement. Segments characterised by them only pass through up to four sectors of the orientation grid. Instead, a subset of \mathcal{BA}_{23} containing only atomic bipartite arrangements is worth to be further examined. We classify a \mathcal{BA} relation as being atomic if it either is located in one single sector or if it passes exactly one of the orientation grid's singularities connecting two adjacent sectors. This holds for twelve relations, namely $B_l, BO_l, D_l, FO_l, F_l, F_m, F_r, FO_r, D_r, BO_r, B_r$, as well as B_m . In the following, these relations will be referred to as $\mathcal{BA}_{12} \subset \mathcal{BA}_{23}$.

3.1 Scope of Bipartite Arrangements

Using the atomic relations identified above, the position of all \mathcal{BA}_{23} relations can be characterised w. r. t. the reference segment. For this purpose, each relation is represented by the set of \mathcal{BA}_{12} relations describing its position. In the following, this set is referred to as the relation's scope σ . The scope of each of the twelve atomic relations is rather simple. It contains only the respective \mathcal{BA}_{12} relation and no further relations. More interesting are the scopes of the remaining non-atomic relations, namely BO_{ml} , BC_l , C_l , FC_l , FO_{ml} , FO_{mr} , FC_r , C_r , BC_r , and BO_{mr} . Since we deal with sets, these scopes can simply be considered as unions of atomic scopes. For instance, the position of C_l is determined by the set $\{B_l, BO_l, D_l, FO_l, F_l\}$. Each scope can be visualised by twelve circles, each of which being located in accordance to its respective atomic relation in the orientation grid. A circle is depicted opaque if its \mathcal{BA}_{12} relation is member of the scope, transparent otherwise. The right hand side of Fig. 1 depicts this iconic visualisation for all \mathcal{BA}_{23} relations.

3.2 Scope of Courses

After having defined the scopes for \mathcal{BA}_{23} relations, we are now prepared to determine the scope of a polygon's course, i. e. for the whole polygon. As seen before, the scope of non-atomic relations can simply be determined by combining their respective atomic relations using the well-known union operation for sets. Extending this to obtain the scope of the complete polygon is straightforward. For the time being, we consider the scope σ' simply as being determined by the union of the scopes of all relations r_i contained in the course $C(x)$:

$$\sigma'(C(x)) \equiv \bigcup_{i=0}^{n-1} \sigma(r_i), r_i \in \mathcal{BA}_{23} \quad (1)$$

Every course of a simple, closed polygon describes a connected sequence of segments. However, while the scopes depicted on the right hand side of Fig. 1 form gapless chains of atomic relations, this does not hold for every scope determined using (1). This observation can be illustrated by taking a closer look at the left hand side of Fig. 1. These visualisations of the \mathcal{BA}_{23} relations show that two relations which pass singularities of the orientation grid can follow each other without being interrupted by the relation in the sector in between. For example, BO_l can be followed directly by FO_l without D_l in between. However, one could argue that D_l has nevertheless been passed through and therefore should be considered in the resulting scope. In order to achieve this, the following auxiliary function is used:

$$close(\sigma') = \{sel(\sigma', 1), sel(\sigma', 2), \dots, sel(\sigma', 12)\} \quad (2)$$

This function itself uses another auxiliary function:

$$sel(\sigma', i) = \begin{cases} \mathcal{BA}_{12}(i) & \text{if } \mathcal{BA}_{12}(i) \in \sigma' \\ & \text{or } \mathcal{BA}_{12}(i-1) \in \sigma' \wedge \mathcal{BA}_{12}(i+1) \in \sigma' \wedge odd(i) \\ \{\} & \text{else} \end{cases} \quad (3)$$

In order to arrive at a scope having all gaps closed, (2) uses (3) twelve times. The parameters of (3) are the preliminary scope σ' determined with (1) as well as a number from one to twelve. These numbers denote the atomic relations according to their position in the scope as depicted in Fig. 1. Starting with B_l , ending with B_m this allows to denote the atomic relations by $\mathcal{BA}_{12}(1)$ to $\mathcal{BA}_{12}(12)$ instead of their mnemonic labels. Equation (3) first checks whether an atomic relation is part of the preliminary scope. If so, its return value is the atomic relation itself. A second case handles the situation if, instead of the relation itself, both of its adjacent relations are contained in the preliminary scope. If the relation itself additionally has an odd number, it fulfills the condition of being located in one single sector and being left out by its neighbours. In this case, (3) also returns the respective atomic relation. Otherwise it returns an empty set. Applying (2) we get the final definition for the scope of courses:

Definition 2 (Scope) *x is a line segment of a simple, closed polygon and $C(x)$ is its course. The set of atomic relations describing the position of $C(x)$ is called the scope of the course, in short $\sigma(C(x))$. It is defined by:*

$$\sigma(C(x)) \equiv \text{close} \left(\bigcup_{i=0}^{n-1} \sigma(r_i) \right), r_i \in \mathcal{BA}_{23}$$

Since every scope is a set of up to twelve atomic relations, in theory $2^{12} = 4096$ different scopes can be distinguished. As outlined above, the scopes realised by simple, closed polygons always form a gapless chain of atomic relations. This restriction leads to an extremely confined number of scopes that actually exist. Each of these chains can contain up to twelve atomic relations. Those chains containing only up to eleven relations can start at each of the twelve atomic positions. If a scope contains all twelve relations its starting point is irrelevant. This scope is referred to as the universal scope. Altogether $11 \cdot 12 + 1 = 133$ gapless scopes can be distinguished. However, it can be observed that 47 of these 133 scopes cannot be realised by simple, closed polygons. This is due to the fact that a realisation of these scopes would require a mathematically negative order of the polygon's points. By restricting ourselves to polygons with a mathematically positive ordering, only $133 - 47 = 86$ scopes can be realised by simple, closed polygons.

To summarise, we are now able to qualitatively characterise the position of any simple, closed polygon w. r. t. a reference segment by a set of twelve atomic relations. This means, that space complexity for this description is constant instead of linear (as before), when characterising a polygon by the position of each of its segments. Thus, we are also able to characterise the polygon w. r. t. all of its segments with linear space complexity instead of quadratic space complexity, as before. In the following section we will see how to reduce this complexity even further.

3.3 Scope of Polygons

In Sect. 2 we recognised that matrices are permutable and that two matrices may not have the same size, which is the reason for the high runtime complexity when comparing two polygons on the basis of this description. In principle, the same holds for a polygon characterised by n scopes as described in Sect. 3.2: The sequence of scopes is permutable and may not have the same length for two polygons. These observations are related to the fact that we deal with an ordered sequence of scopes.

Instead of dealing with ordered sequences, we shall now discard this ordering: The qualitative partitioning into classes by the 86 scopes allows a histogram of frequencies to be computed (Fig. 3). Thus, instead of being linear, space complexity for this description is constant now, since only the frequencies of a fixed number of 86 scopes are to be described. In order to achieve a certain invariance against different approximation results, it makes sense to weight the histogram’s entries by the relative length of the respective reference segments. Time complexity also decreases by this procedure, since the distance of two histograms can easily be computed by the sum of the distances of their corresponding entries. Thus, we are now able to characterise polygons with constant space complexity. Furthermore, it is possible to compare two polygons simply by considering the histograms of their scopes. In the following Sect. 4, we shall learn which retrieval results can be achieved by this method.

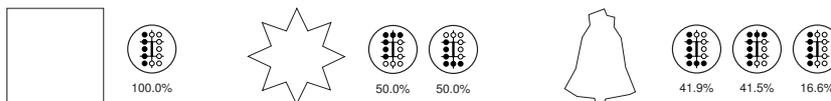


Fig. 3. Three simple scope histograms and their respective polygons. Left: The square’s histogram contains only one entry, since all of its line segments have the same scope relation. Centre: Due to the polygon’s regularity, the star’s histogram consists of two entries with equal frequency. Right: The bell’s scope histogram contains three entries

4 Comparison with Other Approaches

In order to measure the retrieval performance of our method a comparison to other approaches has been carried out. As an evaluation method, we use the well-known core experiment CE-Shape-1 [10] for the MPEG-7 standard, which allows a comparison of approaches to be accomplished by taking into account only their retrieval results. Part B tests the capability of similarity-based retrieval techniques with a database of 1400 images: These images are semantically grouped into 70 classes of various shapes, each class containing 20 objects. Each image is used as a query, all other images in the database are ordered w.r.t. their similarity by the approach under consideration. For each query the number

of images belonging to the same class are counted in the first 40 results. Since every class contains 20 instances, the maximum number of correct matches is 20 for each single query. Altogether, the total number of correct matches for all 1400 queries is 28000. As explained in [10], a retrieval rate of 100% is not possible using only shape knowledge, since some classes contain objects, which are semantically similar but differ significantly regarding their shape. Conversely, using a hypergeometric distribution, it is easy to show that a random ordering of the search results achieves about 2.86% in the MPEG test. This is a lower bound showing how much better an approach is in comparison with mere chance.

Since the scope histogram offers constant time complexity for the comparison of two objects, first of all it makes sense to compare it to other approaches also having this property. This holds for the seven invariant moments proposed by [8]. Applying [14], these moments can directly be computed for polygons like our approach. Even simpler are quantitative numeric features characterising polygons by a single numeric value. An example is the compactness as defined by [3], which corresponds to the ratio $\frac{4\pi A}{P^2}$ of area and perimeter. Further examples are the radius ratio $\frac{R_{min}}{R_{max}}$ of the minimum enclosing circle and the maximal contained circle [4] as well as the aspect ratio $\frac{H_r}{W_r}$ of the minimal enclosing rectangle [3]. Performing the MPEG test for the approaches described before leads to the results listed in Table 1: It shows that the numeric features, namely compactness, radius ratio, and aspect ratio, which characterise a shape by one single number gain results approximately between 16% and 24%. This is already significantly better than when ordering the shapes randomly, which achieves not even three percent correct matches. Even better results can be achieved using the seven Hu moments. Their results are at least ten percentage points better as they retrieve about 34% of the total number of correct matches. The scope histogram introduced in this paper outperforms all other examined approaches and retrieves about 46%. To conclude, using the scope histogram it is possible to achieve a retrieval result which is about 16 times better than a random ordering. Furthermore, our approach outperforms the other examined approaches which also offer constant time complexity.

Table 1. Classification results of compactness (CO), radius ratio (RR), aspect ratio (AR), Hu moments (HU), and scope histogram (SH) for CE-Shape-1 Part B

CO	RR	AR	HU	SH
21.86	16.82	24.12	34.13	45.52

After having examined the retrieval results for the scope histogram, the question arises, whether its results can be improved by combining it with one or more of the other approaches with constant time complexity. Table 2 lists the classification results for some of these combinations. It shows that by combining all numeric features, namely compactness, radius ratio, and aspect ratio, a result

of about 52% can be achieved; including the Hu moments, we obtain 54%; the scope histogram in combination with the Hu moments, 54%; the scope histogram in combination with the three numeric features, we achieve 64%. When taking all five features together into consideration we achieve a retrieval result of approximately 64%. When comparing the retrieval result of all features (AL) with all but the Hu moments (NS), we learn that the Hu moments do not significantly improve the results. By contrast, a comparison of the retrieval results of all features (AL) and all features excluding our scope histogram (NH) shows that the scope histogram improves the results by about ten percentage points. Eventually, it is worth mentioning that a retrieval result of about 64% is only about twelve percentage points less than the results achieved by the correspondence of visual parts of [9] (which is 76.45%), that has a significantly higher time complexity of $O(mn^3)$ for the comparison of two objects.

Table 2. Classification results for combined features: All numeric features (NF), all numeric features and Hu moments (NH), all numeric features and scope histogram (NS), Hu moments and scope histogram (HS), and all these features together (AL)

NF	NH	NS	HS	AL
51.58	53.99	63.75	53.81	64.26

As mentioned above, when performing the complete MPEG test, 1400 queries, each consisting of 1400 comparisons of two objects, have to be processed. Altogether, this results in nearly two million comparisons. In our Java implementation it takes only about 20 seconds to perform the MPEG test for the scope histogram on a computer with Windows XP and an AMD mobile Athlon processor with about 1.5 Gigahertz.

5 Discussion

Using the orientation grid we describe objects by the configuration of their parts, i. e. we obtain self-referring descriptions. In doing so, a certain robustness can be achieved regarding changes of the viewpoint. Our approach is particularly suitable for rigid objects, since they do not change w. r. t. the configuration of their parts. Nevertheless, there exist many areas in which the approach is applicable, that is, there are many domains which deal with rigid artificial objects, and it is the notion of positional-contrast which allows us to intuitively decide whether two object categories can be distinguished. Robustness against noise in the underlying image data is realised by applying approximation algorithms and by using a qualitative approach, completely abstracting from exact quantitative data. Furthermore, it is worth mentioning that the application of an intrinsic reference system brings in invariance against scale, translation, as well as rotation.

Coming to a conclusion, we introduced a new approach offering constant space complexity for the description of a polygon by its scope, which is its relative position w. r. t. one of its line segments. By computing the frequencies of all occurring scopes, it is also possible to describe the whole polygon with constant space complexity. This histogram representation also allows two objects to be compared with constant time complexity. The evaluation results show that our new approach outperforms other approaches which also offer constant time complexity. Furthermore, due to the underlying concept of positional-contrast it can be intuitively decided, whether it is worth combining it with other approaches, namely with those ones which are not based on the description of the configuration of their parts (e. g. square measures, roundness, etc. — not to mention texture and colour). This combination leads to retrieval results of about 64% in the MPEG test.

References

1. J. F. Allen. Maintaining Knowledge about Temporal Intervals. *Communications of the ACM*, 26(11):832–843, 1983.
2. F. Attneave. Some Informational Aspects of Visual Perception. *Psychological Review*, 61:183–193, 1954.
3. R. O. Duda and P. E. Hart. *Pattern Classification and Scene Analysis*. John Wiley and Sons, Inc., 1973.
4. G. D. Garson and R. S. Biggs. *Analytic Mapping and Geographic Databases*. Sage Publications, Newbury Park, CA, 1992.
5. B. Gottfried. Reasoning about Intervals in Two Dimensions. In *IEEE Int. Conference on Systems, Man and Cybernetics*, pages 5324–5332, The Hague, 2004.
6. B. Gottfried. *Shape from Positional-Contrast — Characterising Sketches with Qualitative Line Arrangements*. Doctoral dissertation, University of Bremen, 2005.
7. B. Gottfried. Querying for Silhouettes by Qualitative Feature Schemes. In *DIAL'06, 2nd IEEE Inter. Conf. on Document Image Analysis for Libraries*, 2006.
8. M.-K. Hu. Visual Pattern Recognition by Moment Invariants. *IRE Transactions on Information Theory*, 8(2):179–187, 1962.
9. L. J. Latecki and R. Lakämper. Shape Similarity Measure Based on Correspondence of Visual Parts. *IEEE PAMI*, 22(10):1185–1190, 2000.
10. L. J. Latecki, R. Lakämper, and U. Eckhardt. Shape Descriptors for Non-rigid Shapes with a Single Closed Contour. In *IEEE CVPR*, pages 424–429, 2000.
11. S. Loncaric. A Survey of Shape Analysis Techniques. *Pattern Recognition*, 31:983–1001, 1998.
12. D. A. Mitziias and B. G. Mertzios. Shape Recognition with a Neural Classifier Based on a Fast Polygon Approximation Technique. *Pattern Recognition*, 27:627–636, 1994.
13. P. L. Rosin. Assessing the Behaviour of Polygonal Approximation Algorithms. *Pattern Recognition*, 36:505–518, 2003.
14. C. Steger. On the Calculation of Arbitrary Moments of Polygons. Technical Report FGBV-96-05, Informatik IX, Technische Universität München, 1996.
15. K. Zimmermann and C. Freksa. Qualitative Spatial Reasoning Using Orientation, Distance, and Path Knowledge. *Applied Intelligence*, 6:49–58, 1996.