# Shape Retrieval with Qualitative Relations

## The Influence of Part-Order and Approximation Precision on Retrieval Performance and Computational Effort

Arne Schuldt

Centre for Computing and Communication Technologies (TZI)
University of Bremen, Am Fallturm 1, 28359 Bremen, Germany

**Abstract.** Manifold approaches exist in the field of similarity-based shape retrieval. Although many of them achieve good results in reference tests, there has been less focus on systematically examining the factors influencing both retrieval performance and computational effort. Such an investigation, however, is important for the structured development and improvement of shape descriptors. This paper contributes a thorough investigation of the influence of the shape part-order and approximation precision. Firstly, two shape descriptors based on qualitative spatial relations are introduced and evaluated. These descriptors are particularly suited for the intended investigation because their only distinction is that one of them preserves the part-order, the other abandons it. Secondly, the recall and precision values are related to the degree of approximation in three-dimensional recall-precision-approximation diagrams. This helps choose an appropriate approximation precision. Finally, it turns out that remarkable retrieval results can be achieved even if only qualitative position information is considered.

## 1 Introduction

Similarity-based shape retrieval poses a demanding task in computer vision. 2D object shapes originate from projections of 3D objects. Latecki et al. [10] point out that, consequently, the shape of the 2D projection may change because

- the observer of the original 3D object changes his point of view,
- the original 3D object performs non-rigid motion, or
- digitisation or segmentation cause noise.

Existing shape descriptions range from compact [7, 2, 5, 14] to complex [13, 8, 11, 15]. While complex shape descriptors focus on maximising the retrieval performance, compact descriptors aim at minimising the computational effort.

Although manifold shape descriptors exist, less effort has been spent on a systematic investigation of the factors influencing both retrieval performance and computational effort. Examining such factors, however, could lead to important insights for creating both effective and efficient shape descriptions. In this context, the main research questions addressed by this paper are:

1. What is the influence of the part-order on shape retrieval?
2. What is the influence of the approximation precision on shape retrieval?

The first question is motivated by earlier experiments [15] on the influence of preserving and abandoning the order of parts (line segments in the case of polygons) in the shape representation. The second question is due to the insight that often a comparatively coarse shape abstraction suffices for object recognition [1]. As the precision may have a significant influence on the computational effort, it is crucial to determine an appropriate degree of approximation.

Obviously, one cannot expect exact quantitative answers (such as "the part-order has an influence of 15.7 % on the retrieval performance") to the above research questions that hold for all conceivable shape descriptors. However, if a significant influence can be verified for a specific shape descriptor, this may outline a clear tendency towards a general qualitative answer (such as "the part-order has a positive influence on the retrieval performance"). In particular, this paper investigates two shape descriptions that are based on cognitively motivated qualitative spatial relations. These relations characterise the relative positions of the polygon segments of the shape's contour outline qualitatively. This leads to the following additional research question to be answered in this paper:

3. How significant is qualitative position information for shape retrieval?

The remainder of this paper is structured as follows. Section 2 introduces the two shape descriptors that serve as the subject of investigation throughout this paper. Section 3 discusses the influence of the shape precision in order to identify an appropriate degree of approximation. Section 4 presents extensive experiments that examine the influence of part-order and precision on retrieval performance and computational effort. Subsequently, Section 5 gives a discussion of the evaluation results. Finally, Section 6 summarises the findings and gives an outlook on future research.

## 2 Shape Description with Qualitative Spatial Relations

The first research question asks whether sticking to the order of the shape parts is advantageous or disadvantageous. To answer this question, it is important to compare two shape descriptions that only distinguish with respect to this property. Section 2.1 introduces the qualitative bipartite arrangement relations which are the foundation for the shape descriptors investigated in this paper. Subsequently, Section 2.2 introduces how part-order preserving matrices of bipartite arrangements can be applied to characterise polygons. Finally, Section 2.3 defines part-order abandoning histograms of bipartite arrangements which are a very compact shape description.

### 2.1 Qualitative Bipartite Arrangement Relations

Gottfried [6] proposes the positional-contrast framework of qualitative spatial relations. An important building block are the 23 bipartite arrangement relations, in short $\mathcal{BA}_{23}$ (Figure 1). These relations describe the relative position of two line segments with respect to each other. To this end, a double cross [17] is induced on one of the line segments, $\overrightarrow{bc}$ in the example (Figure 1 left). It consists of three auxiliary lines that tessellate the two-dimensional plane into six sectors that can be perceived easily by humans. One auxiliary line characterises the left/right dichotomy, the other two distinguish front, during, and back. The start and end points of the other line segment $\overrightarrow{ad}$ can then be located in any of the six sectors. If symmetries and intersections are left out, the 23 $\mathcal{BA}_{23}$ relations remain (Figure 1 right).

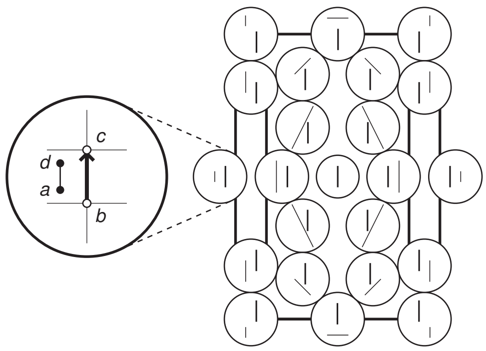


**Fig. 1.** The double cross induced on $\overrightarrow{bc}$ divides the two-dimensional plane into six sectors (left). Any other line segment can start and end in any of these sectors. This leads to the 23 $\mathcal{BA}_{23}$ relations (right).

This representation exhibits important properties with respect to the three challenges for similarity-based shape retrieval (Section 1). On the one hand, the representation is invariant against the affine transforms scale, translation, and rotation. This is accomplished by the double cross as in intrinsic reference system which is induced on one of the characterised line segments and thus exposed to the same transforms. Consequently, the representation is to a certain extent invariant against changes in the point of view and non-rigid object motion. On the other hand, compared to exact quantitative representations, the $\mathcal{BA}_{23}$ relations have a rather coarse resolution. Even if qualitatively characterised points move (due to any of the three challenges mentioned), they will frequently not leave their double cross sector and thus not lead to another relation.

### 2.2 Part-Order Preserving Bipartite Arrangement Matrix

The $\mathcal{BA}_{23}$ relations characterise the relative position of two line segments. Generally, real-world shapes are more complex than just two line segments. To this end, Gottfried's positional-contrast [6] introduces a matrix in which all line segments of a polygon are related to all others. Consequently, such matrices have a quadratic space complexity, $O(n^2)$. A row in the matrix is referred to as the *course* of the polygon. Each row characterises the qualitative positions of all other line segments with respect to a reference segment. Each column characterises the position of the reference segment with respect to all other segments.

The off-line complexity for computing a $\mathcal{BA}_{23}$ matrix is $O(n^2)$ as each of the $n$ polygon segments has to be related to all $n$ others. The distance of two matrices is the percentage of non-matching entries whereby only identical entries match. Each pair of matrices has to be compared $n$ times due to cyclic permutation (the permutation with the minimum difference determines the distance). Hence, the complexity for on-line comparison is $O(n^3)$. Therewith, this shape description can be characterised as complex.

Note that only matrices of the same size can be compared to each other because the matching process becomes even more expensive otherwise. However, there is nothing to be said against approximating triangles with more than three points if this helps reduce the overall computational complexity.

### 2.3 Part-order Abandoning Bipartite Arrangement Histogram

While the $\mathcal{BA}_{23}$ is a complex shape descriptor that aims at maximising the retrieval performance, other applications require minimal computational effort. Think, for instance, of the real-time constraints in the RoboCup domain [3]. The complexity arises from the fact that the order with which segments occur in the polygon is preserved in the matrix. The complexity can be significantly reduced if this part-order is abandoned. This can be achieved by computing a histogram of the frequencies with which the $\mathcal{BA}_{23}$ relations occur in the matrix. While this representation considers which relations occur, their particular order is neglected. Such a histogram has 23 entries, independently from the number of polygon segments characterised. The space complexity it is thus $O(1)$.

The off-line computational complexity for computing a histogram is still $O(n^2)$ as it requires computing the matrix first. However, the on-line complexity is significantly reduced because two $\mathcal{BA}_{23}$ histograms can be compared with constant complexity, $O(1)$. Therewith, abandoning the part-order has actually a significant influence on the asymptotic computational complexity.

## 3 Appropriate Precision of Approximation

So far, this paper has implicitly assumed that shapes are characterised by polygons of their contour points. This is motivated by Attneave's finding that the contour points are of particular importance for recognition while the remaining pixels can be neglected [1]. However, even the contour points are not of equal importance. With a cat approximated by only 38 points (Figure 2), Attneave demonstrates that even comparatively few points suffice for object recognition [1]. In particular, points laying on straight lines are redundant for object perception [12], while points of high curvature are important [1].

This raises the question how important points can be identified. The discrete curve evolution approach [9] iteratively deletes the contour points with the least

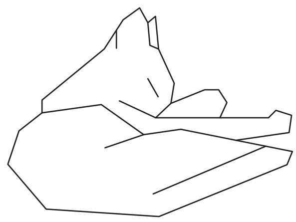

**Fig. 2.** Although this object has been approximated by Attneave [1] with only 38 connected points, it can still be recognised as a sleeping cat.

influence on the appearance of the shape (based on lengths and angles). However, ordering the contour points based on their importance does not answer the question how many of them are actually required. Instead, McNeill and Vijayakumar [11] propose to simply choose a fixed number of arbitrary equally-spaced contour points. Proceeding this way has the following advantages:

1. The approximation can be computed quickly because the original polygon is simply split into equally-spaced points.
2. There is no need to decide which and how many points are relevant for the appearance of an individual shape.
3. Choosing a fixed number of contour points for all shapes simplifies the on-line matching problem (as assumed for the $\mathcal{BA}_{23}$ matrix).

Although this approximation strategy relieves one from finding out how many points are necessary for an individual shape, the question remains how many points are required overall. Figure 3 depicts the contour outline of a beetle with different precision of approximation starting with a triangle. In the first row of shapes, it is virtually impossible to recognise the beetle. In the second row, the beetle becomes gradually recognisable. In the third row, the contour is sufficiently precise to identify that it is a beetle outline. The increased precision in the fourth row does not provide much more information. As a hypothesis, it can therefore be assumed that approximating this shape with 30 equally-spaced points should suffice for shape retrieval. A more general answer will be derived based on extensive retrieval experiments.

## 4 Evaluation

As a foundation for answering the research questions posed in Section 1, extensive experiments have been conducted. The foundation for this evaluation is the shape test data set of the MPEG-7 CE-Shape-1 reference test [10]. This well-known data set comprises $1,400$ shapes that are organised in 70 categories with 20 instances each. Example instances for each category are depicted in Figure 4. Section 4.1 evaluates the retrieval performance of the $\mathcal{BA}_{23}$ histogram and the $\mathcal{BA}_{23}$ matrix in an established reference test. Section 4.2 relates these results to the computational effort to be spent for shape retrieval with both approaches.

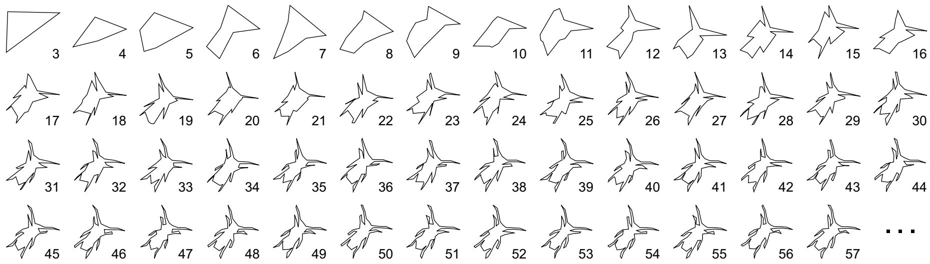


**Fig. 3.** The contour outline of a beetle from the MPEG-7 data set [10] approximated with increasing precision. The index denotes the number of polygon points.

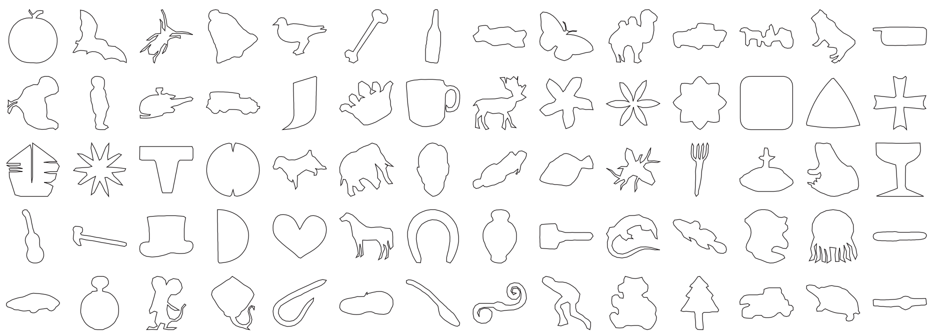

**Fig. 4.** Example instances for all 70 categories of the test data set for the MPEG-7 CE-Shape-1 reference test [10], each example represents 20 instances.

### 4.1 Shape Retrieval Performance

The so-called bullseye test [10] is a well-established method for measuring the retrieval performance of shape descriptors. The approach enables the comparison of shape descriptors with completely different mathematical foundations simply based on their retrieval performance. It is usually carried out by the developers of the respective shape description themselves in order to ensure the best parameter choice. The bullseye test is defined as follows [10]. Each of the $1,400$ shapes (Figure 4) is used as a query, one after another. All other shapes are then ordered based on their similarity with the query. The correct matches within the first 40 results are summed up for all queries. This number is related to $1,400 \cdot 20 = 28,000$ which is the total number of possible correct matches (each of the 70 classes has 20 instances). It is worth mentioning that the classes are grouped semantically. Therefore, retrieval results of $100.0\,\%$ are unlikely solely based on shape knowledge.

The bullseye values for the $\mathcal{BA}_{23}$ histogram in relation to the degree of approximation are depicted on the left hand side of Figure 5. In addition to the aggregated bullseye values, the three-dimensional recall-precision-approximation diagram in Figure 6 relates the recall and precision values to the degree of approximation. Both diagrams show that, initially, the retrieval performance increases with the number of polygon points. Having reached a precision of about 30 points, however, the retrieval performance fluctuates only slightly. The best bullseye value of 41.8 % is achieved with 37 polygon points.

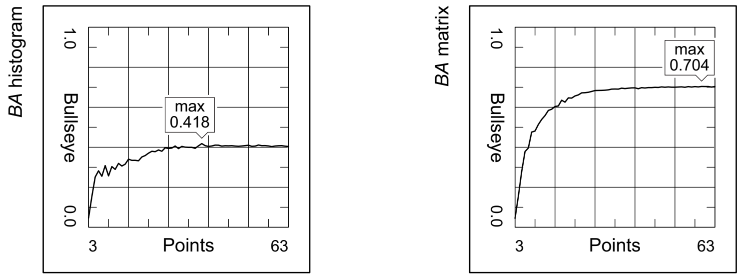


**Fig. 5.** Retrieval performance in the bullseye test in relation to the degree of approximation. The $\mathcal{BA}_{23}$ histogram (without part-order) achieves a result of up to 41.8 % (left) while the $\mathcal{BA}_{23}$ matrix (with part-order) achieves even up to 70.4 % (right).

Compared to the $\mathcal{BA}_{23}$ histogram, the $\mathcal{BA}_{23}$ matrix is a more complex shape representation. Hence, it is not surprising that it achieves better retrieval results. As depicted on the right hand side of Figure 5, the best bullseye value achieved is 70.4 %. The better retrieval performance also becomes clear when comparing the more detailed recall-precision-approximation diagrams for $\mathcal{BA}_{23}$ histogram (Figure 6) and matrix (Figure 7). To recapitulate, the effort for comparing two $\mathcal{BA}_{23}$ matrices depends on the number of polygon points. Even more important than the pure retrieval performance is therefore the degree of approximation required. Like for the $\mathcal{BA}_{23}$ histogram, the results improve significantly until about 30 polygon points. The best result is achieved with 59 points. However, a bullseye value that is only about one percentage point below can be achieved already with 35 polygon points. For a result of only about two percentage points below the best result, 27 polygon points suffice.

### 4.2 Computational Effort for Shape Retrieval

To recapitulate, the foundation for the evaluation is the test data set of $1,400$ shapes. This means that, for one query, the query shape has to be compared to $1,400$ other shapes. As every instance serves as a query, one after another, this makes $1,400^2 = 1,960,000$ for the whole experiment. Each $\mathcal{BA}_{23}$ histogram consists of 23 real numbers, independently from the precision of approximation. Consequently, $23 \cdot 1,400 = 32,200$ real numbers have to be compared for one query and $23 \cdot 1,960,000 = 45,080,000$ for the whole experiment (Figure 8 left). Each qualitative relation in a $\mathcal{BA}_{23}$ matrix is represented by an integer number. This means, that for a whole experiment, $1,960,000 \cdot n^3$ integers have to be compared with $n$ being the number of polygon points. The computational effort

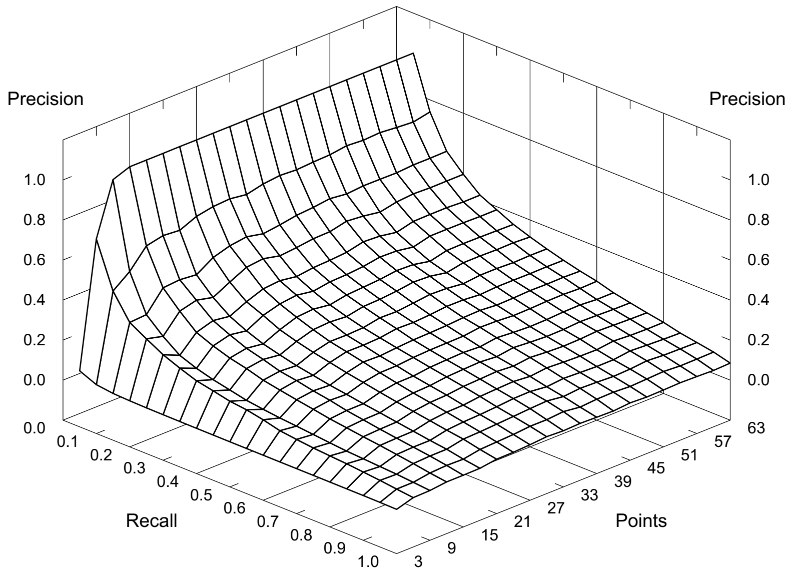


**Fig. 6.** Recall and precision values for the $\mathcal{BA}_{23}$ histogram (without part-order) in relation to the approximation from 3 to 63 polygon points.

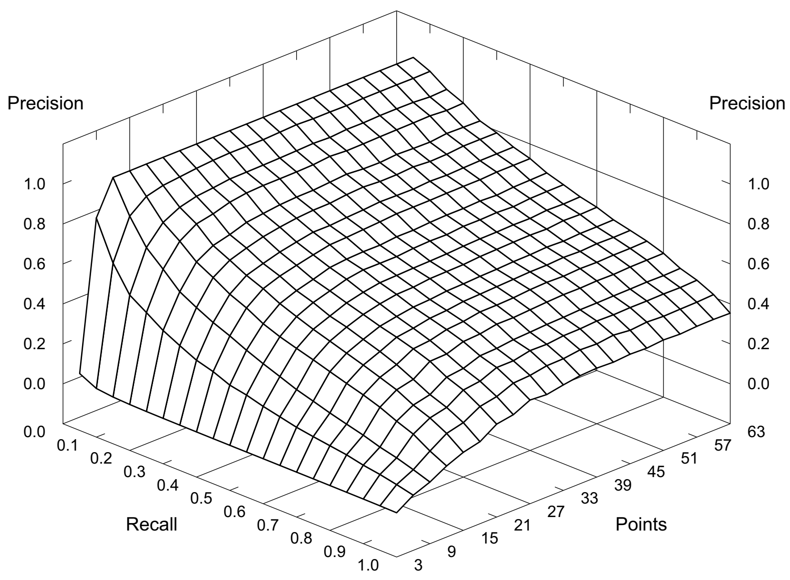


**Fig. 7.** Recall and precision values for the $\mathcal{BA}_{23}$ matrix (with part-order) in relation to the approximation from 3 to 63 polygon points.

ranges from 52,920,000 to 490,092,120,000 integer comparisons for polygons with 3 and 63 points, respectively (Figure 8 right). In contrast to the $\mathcal{BA}_{23}$ histogram, the precision of approximation has thus a significant influence on the effort to be spent for comparing two $\mathcal{BA}_{23}$ matrices.

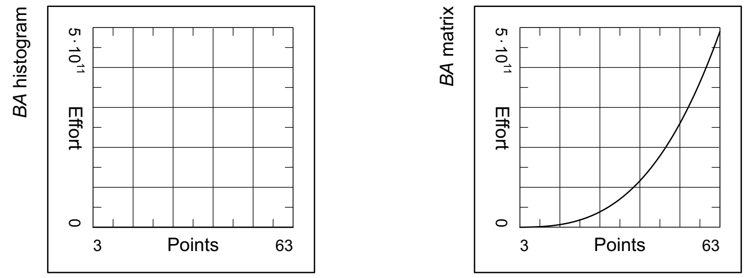


**Fig. 8.** The computational effort to be spent for the experiments conducted in relation to the degree of approximation. For the $\mathcal{BA}_{23}$ histogram, the computational effort does not depend on the number of polygon points (left). The asymptotic complexity of the $\mathcal{BA}_{23}$ matrix is $O(n^3)$ with $n$ being the number of polygon points.

The retrieval task can be parallelised by splitting up each query into individual comparisons of two shapes. These comparisons can then be delegated to different operating system threads and processors. All experiments have been conducted on a computer with eight dual-core AMD Opteron 8218 processors with 2.6 GHz each and 64 GB RAM in total. The times for conducting the experiments (i. e., 1,960,000 comparisons each) are as follows. For the $\mathcal{BA}_{23}$ histogram, an experiment took about 7.608 seconds. Depending on the number of polygon points, the time consumed by $\mathcal{BA}_{23}$ matrix experiments ranges from 8.040 seconds for triangles to 544.387 seconds for polygons with 63 points. This means that an individual query (i. e., 1,400 comparisons) took about 5.434 milliseconds for the $\mathcal{BA}_{23}$ histogram and between 5.743 and 388.848 milliseconds for $\mathcal{BA}_{23}$ matrices. Note that this relationship does not correspond to the theoretically expected values. This can be explained by the administrative overhead (e. g., thread scheduling and result management) that has more influence for compact descriptors and becomes neglectable for more complex ones.

## 5 Research Questions Revisited

Based on the comprehensive evaluation, it is now possible to answer the research questions raised in Section 1. Section 5.1 addresses the influence of the shape part-order while the appropriate degree of approximation is dealt with in Section 5.2. Finally, Section 5.3 compares the retrieval performance of the qualitative descriptions to existing ones.

### 5.1 Influence of the Shape Part-Order

The shape part-order has a significant influence on the retrieval performance of the investigated shape descriptor. The performance of the part-order preserving $\mathcal{BA}_{23}$ matrix in the reference test is 29 percentage points or 68 % better than that of the part-order abandoning $\mathcal{BA}_{23}$ histogram. This result is even clearer than that of an earlier experiment [15] with another pair of shape descriptions in which the improvement was only 17 percentage points or 38 %. Although the exact figures are clearly not transferable to other shape descriptors, there is a strong indication for the significance of the part-order for shape recognition.

However, there is a price to be paid for preserving the part-order, namely computational effort. Comparing two $\mathcal{BA}_{23}$ histograms is very cheap and does not depend on the number of polygon points. For the improved retrieval performance of the $\mathcal{BA}_{23}$ matrix, considerably higher costs arise. Therefore, it is important to identify an appropriate degree of polygonal approximation.

### 5.2 Influence of the Approximation Precision

Also, the precision of polygonal approximation has a significant influence on the shape retrieval performance. On the one hand, it is not surprising that an approximation that too coarse (e. g., a triangle) is not meaningful enough. On the other hand, the evaluation shows that the number of required polygon points is actually limited. Initially, the retrieval performance increases with an improved approximation. The gradual increase can be explained by the fact that the individual shape classes successively reach the minimum precision required by them. From about 30 points on, there is no further significant improvement. This finding is backed by the exemplary approximation of the beetle (which is one of the more detailed classes) in Figure 3. In this example, also about 30 equally-spaced points suffice to make the beetle visually recognisable.

The limit for the approximation precision is very valuable for taming the computational effort. For the $\mathcal{BA}_{23}$ matrix, the correlation of the computational complexity and the number of $n$ polygon points is $n^3$. Note, however, that as soon one has decided for a particular degree of approximation, the correlation of the computational complexity and the number of shapes in the search space is linear (of course, the effort for the matrix is higher than for the histogram). As explained in Section 4.2, the retrieval task can be highly parallelised. Therefore, it depends on the number of shapes to be compared and on the computer power available whether the descriptors are applicable for real-time applications, such as RoboCup vision or interactive systems.

A thorough examination of the precision-recall-approximation diagrams (Figures 6 and 7) reveals that neither the $\mathcal{BA}_{23}$ histogram nor the matrix can retrieve the identity shape for low degrees of approximation precision (i. e., the precision value for a recall value of 0.05 is not 1.0 as one might expect). This can be explained by the fact that the qualitative position relations are not sufficiently distinctive at this coarse level of approximation.

### 5.3 Significance of Qualitative Position Information

Finally, the question has to be answered how the retrieval performance of the $\mathcal{BA}_{23}$ histogram and the $\mathcal{BA}_{23}$ matrix relate to other approaches described in the literature. This can be accomplished based on the bullseye values.

The $\mathcal{BA}_{23}$ histogram pertains to the class of compact shape descriptions. Quantitative methods in this class are the numeric shape descriptors compactness, radius ratio, and aspect ratio from text books [2, 5] which achieve between 16.8 and 24.1 % [14]. The Hu moments [7] which can also be computed for polygons [16] achieve bullseye values of 34.1 % [14]. The $\mathcal{BA}_{23}$ histogram with its 41.8 % thus outperforms these quantitative methods. It is itself slightly outperformed by the qualitative scope histogram which achieves 45.5 % [14].

The $\mathcal{BA}_{23}$ matrix clearly outperforms all compact descriptions above. As a complex shape descriptor, it has to be compared to other approaches of its class. The curvature scale space approach [13] has a bullseye value of 75.4 % [10]. The correspondence of visual parts [8] achieves 76.5 % [10]. The Procrustes distance even achieves 79.2 % [11]. Although the 70.4 % of the $\mathcal{BA}_{23}$ matrix are slightly below those quantitative approaches, it is a remarkable result as it is achieved based only on positional-contrast. Future investigations will thoroughly examine which additional information has to be considered to improve these results.

## 6 Summary and Outlook

To summarise, this paper has systematically investigated influence factors for shape retrieval. In particular, it can be learnt that, on the one hand, preserving the part-order has a positive effect on the retrieval performance. On the other hand, it also requires additional computational effort which can be limited with an appropriate approximation precision. To this end, the recall-precision-approximation diagram is an important contribution because it helps identify an adequate degree of approximation. This type of diagram is introduced in this paper; usually, such extensive examinations were not made. Finally, the paper shows that remarkable retrieval results can be achieved even if only qualitative position information is incorporated. Therewith, this paper lays important foundations for the development of novel shape descriptions.

Directions for future research are threefold. The first question is whether a minimal approximation with which shapes can be recognised by humans can be identified based on the insights gained in this paper. The second question is whether more sophisticated algorithms for approximation improve the retrieval results (thereby potentially requiring more expensive matching methods). Finally, the retrieval performance of matrices and histograms of other qualitative spatial relations and distance measures has to be examined. More complex distance measures might, for instance, exploit the conceptual neighbourhood of the qualitative relations [4, 6].

## References

1. F. Attneave. Some Informational Aspects of Visual Perception. *Psychological Review*, 61(3):183–193, 1954.
2. R. O. Duda and P. E. Hart. *Pattern Classification and Scene Analysis*. John Wiley & Sons, New York, NY, USA, 1973.
3. A. Fabisch, T. Laue, and T. Röfer. Robot Recognition and Modeling in the RoboCup Standard Platform League. In *Humanoids 2010*, Nashville, TN, USA, 2010.
4. C. Freksa. Temporal Reasoning Based on Semi-Intervals. *Artificial Intelligence*, 54(1):199–227, 1992.
5. G. D. Garson and R. S. Biggs. *Analytic Mapping and Geographic Databases*. Sage Publications, Newbury Park, CA, 1992.
6. B. Gottfried. Qualitative Similarity Measures — The Case of Two-Dimensional Outlines. *Computer Vision and Image Understanding*, 110(1):117–133, 2008.
7. M.-K. Hu. Visual Pattern Recognition by Moment Invariants. *IRE Transactions on Information Theory*, 8(2):179–187, 1962.
8. L. J. Latecki and R. Lakämper. Shape Similarity Measure Based on Correspondence of Visual Parts. *PAMI*, 22(10):1185–1190, 2000.
9. L. J. Latecki and R. Lakämper. Polygon Evolution by Vertex Deletion. In *Scale-Space 1999*, pages 398–409, Corfu, Greece, 1999. Springer-Verlag.
10. L. J. Latecki, R. Lakämper, and U. Eckhardt. Shape Descriptors for Non-rigid Shapes with a Single Closed Contour. In *CVPR 2000*, pages 424–429, Hilton Head Island, SC, USA, 2000. IEEE Computer Society.
11. G. McNeill and S. Vijayakumar. 2D Shape Classification and Retrieval. In *IJCAI 2005*, pages 1483–1488, Edinburgh, UK, 2005. Professional Book Center.
12. D. A. Mitzias and B. G. Mertzios. Shape Recognition with a Neural Classifier Based on a Fast Polygon Approximation Technique. *Pattern Recognition*, 27:627–636, 1994.
13. F. Mokhtarian, S. Abbasi, and J. Kittler. Robust and Efficient Shape Indexing through Curvature Scale Space. In *BMVC 1996*, pages 53–62, Edinburgh, UK, 1996. British Machine Vision Association.
14. A. Schuldt, B. Gottfried, and O. Herzog. Retrieving Shapes Efficiently by a Qualitative Shape Descriptor: The Scope Histogram. In *CIVR 2006*, pages 261–270, Tempe, AZ, USA, 2006. Springer-Verlag.
15. A. Schuldt, B. Gottfried, O. Osterhagen, and O. Herzog. On the Importance of Preserving the Part-Order in Shape Retrieval. In *SIGIR 2007*, pages 771–772, Amsterdam, The Netherlands, 2007. ACM Press.
16. C. Steger. On the Calculation of Arbitrary Moments of Polygons. Technical Report FGBV-96-05, Informatik IX, Technische Universität München, 1996.
17. K. Zimmermann and C. Freksa. Qualitative Spatial Reasoning Using Orientation, Distance, and Path Knowledge. *Applied Intelligence*, 6:49–58, 1996.