Towards the Visualisation of Shape Features
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. Classifying objects in computer vision, we are faced with a great many features one can use. This paper argues that diagrammatic representations help to comprehend properties of features. This is important for the purpose of deciding which features should be used for a given classification task. We introduce such a diagrammatic representation for a shape feature and show how it enables one to decide whether this feature helps to distinguish some categories given. Additionally, it shows that the proposed feature keeps up with other features falling into the same complexity class.

1 Introduction

In computer vision several features have been devised in order to describe objects. While some of them are more intuitive than others, even the intuitive ones raise the question whether they are appropriate given a number of categories which are to be distinguished. For instance, a feature like the compactness obviously allows two object categories to get distinguished which are clearly different regarding their roundness. However, determining this feature for a number of categories it is quite difficult to comprehend its meaning. What does a difference, lets say, of 0.1 tell us about the shapes of two object categories? This question is especially to be asked from the point of view of the expert, who possesses particular skills about his domain but not about some methods he needs to apply to this domain.

It is frequently difficult or even impossible to determine which object properties, and as a consequence, which categories can be distinguished by specific features without thoroughly analysing them regarding that feature. Furthermore, it is even more difficult to comprehend how a number of different features combine. The problem we are faced with is more general and concerns the difficulty to choose among a number of methods. Normally, the appropriateness of the chosen method is not revealed before having used it. Here, we argue in favour of visualisation techniques which relate to diagrammatic representations. Such techniques enable us to assess the meaning of some method or feature by visualising some of its properties. Thereby, its inherent structure is made explicit, allowing attention to be guided in a specific way.

From the point of view of diagrammatic reasoning, [8] point out that attention mechanisms are relevant when interpreting diagrams. In particular, the ease

of recognition is strongly affected by what information is explicit in a representation. The idea is to make use of these findings by devising diagrams which relate a feature's semantics to its range of values. In this paper we shall demonstrate how this works for a representation which comprises a structure that enables a specific feature defined on it to be visualised, so that the relationship between semantics and structure is made visible. As a consequence, it can intuitively be decided how well a number of categories can be distinguished by this feature.

In Sect. 2 previous work on a qualitative shape description is reviewed. At the same time, a diagrammatic representation for a shape feature is introduced which is based on this shape description. Applying this feature to shapes of the MPEG test dataset, Sect. 3 shows that this feature is in fact more expressive than others which pertain to the same class of features in terms of the complexity required for comparisons. But more importantly, it also shows that the feature's meaning can be comprehended very well through its diagrammatic representation. Section 4 discusses for a number of pairwise similar as well as pairwise different categories how their diagrammatic representation enables one to decide in favour or against the application of our feature. Finally, Sect. 5 summarises our conclusions.

2   Scope Histograms

For the purpose of dealing with shape data we choose polygons to be the underlying representation for our approach. Polygons can easily be extracted from binary raster images containing the silhouette of an object. Additionally, they allow the creation of a more compact representation through the application of polygonal simplification algorithms — even with only little influence on the perception of shape [2]. We apply especially the method proposed in [11], choosing an approximation error of one percent of a polygon's perimeter, in this way being invariant regarding the scale.

Consider the 13 qualitative relations between one-dimensional time intervals [1]. Their generalisation to two dimensions define the notion of positional-contrast [7] which allows intervals in the two-dimensional plane to be characterised. Using the orientation grid of [12], [6] introduces 23 $\mathcal{BA}_{23}$ relations that can occur between two line segments (Fig. 1 left and centre). This description is not only restricted to single lines. Instead, using an ordered sequence of $\mathcal{BA}_{23}$ relations, it is also possible to describe the course of a polygon w. r. t. one of its segments. This description has linear space complexity, since each of the polygon's segments is described by one $\mathcal{BA}_{23}$ relation. Indeed, an even more compact description is possible. This can be accomplished by characterising the polygon by its position as a whole, instead of describing the positions of its segments individually.

2.1   Representing Polygons by Their Scope

While $\mathcal{BA}_{23}$ relations are suitable in order to describe arrangements between line segments, they are not devised for the purpose of describing polygons as a
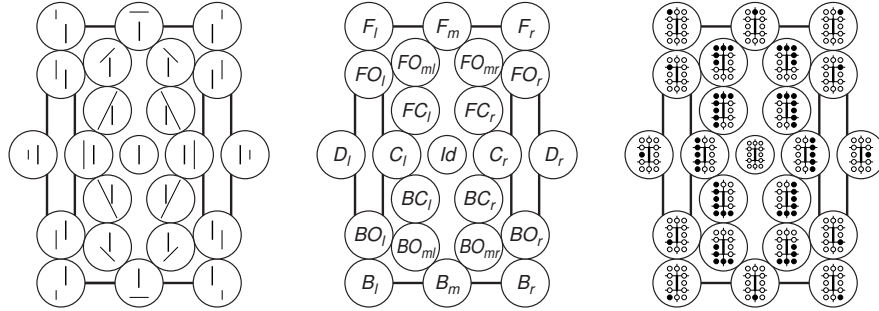
Fig. 1. The 23 $\mathcal{BA}_{23}$ relations that can be distinguished between two line segments in the two-dimensional plane. Left: Example arrangements. Centre: Mnemonic labels. Right: Iconic representation of the respective scopes

whole. This is due to the fact that, in contrast to single line segments, polygons are not restricted to being straight. On this account, there exist polygons that cannot be characterised by single $\mathcal{BA}_{23}$ relations, namely those with a course circulating one or more times completely around the reference segment. $\mathcal{BA}_{12}$, which is a subset of $\mathcal{BA}_{23}$, can be used instead. It contains only the relations B_l, BO_l, D_l, FO_l, F_l, F_m, F_r, FO_r, D_r, BO_r, B_r, and B_m. These relations are referred to as atomic $\mathcal{BA}$ relations. They have the property of being located either in only one of the orientation grid's sectors or of passing exactly one of the orientation grid's singularities connecting two adjacent sectors.

Applying the notion of a *scope* [7], which is nothing other than a set of connected $\mathcal{BA}_{12}$ relations, it is possible to describe the location of any other segment or even the whole course w.r.t. the reference segment. As depicted on the right hand side of Fig. 1 the scopes of the atomic relations contain only the respective relation itself. In contrast, the scopes of the non-atomic $\mathcal{BA}_{23}$ relations are obtained by simply computing the union of the atomic relations comprising their positions. Starting from this, the description of a course is straightforward by simply aggregating $\mathcal{BA}_{12}$ relations which together make up this course.

2.2 Conceptual Neighbourhoods of Scopes

Applying the scope representation described above, the characterisation of a polygon's course w.r.t. one of its segments can be accomplished with constant space complexity. This is due to the fact that we are always dealing with a set of at most twelve atomic relations. In theory, $2^{12} = 4096$ different scopes can be distinguished. When dealing with simple, closed polygons, we can confine ourselves to those scopes, which do not have gaps between their atomic relations. The application of this definition restricts the number of realisable scopes to 133.

After having defined the set of realisable scopes, the question arises as to how they relate concerning their similarity. As an example, the scope of relation F_l may be considered to be more similar to F_m than it is to B_r. As suggested in

[4], it is useful to connect qualitative relations explicitly by defining a conceptual neighbourhood structure. This does not only allow coarse knowledge to be dealt with, as described in [4], but also defines a distance between scopes. Starting from the scope's representation, we define two scopes to be conceptual neighbours if they can directly be transformed into one another by shortening or elongating them with exactly one atomic relation. Since the underlying reference system is circular, the visualisation of the conceptual neighbourhood structure is circular, too. The scopes of the $\mathcal{BA}_{12}$ relations are located at the outmost positions. The more atomic relations a scope contains the shorter is its distance to the centre, where the universal scope is located which contains all atomic relations. The top left quadrant of this neighbourhood structure is depicted in Fig. 2. For instance, in accordance to the neighbourhood structure the distance of the scopes of F_l and F_m is two. In contrast, the scopes of F_l and B_r are located at converse locations within the neighbourhood structure, which corresponds to a distance of twelve edges within this graph.

2.3 Computing Scope Histograms

The scope representation introduced in Sect. 2.1 has the advantage of allowing a polygon to be described w. r. t. one of its segments with constant space complexity. That is to say, we always deal with a set of at most twelve atomic relations regardless of how many line segments are required to represent a specific polygon. Since this is a very coarse description, it is necessary to characterise a polygon not only w. r. t. one of its segments, but w. r. t. to all of its segments. In doing so, we gain a description with linear space complexity, $O(n)$. This approach has the disadvantage that its time complexity for the comparison of two polygons is accordingly higher. This is due to the ordered sequences of scopes which are circularly permutable and due to the fact that different polygons do not have the same number of scopes, from what follows, that time complexity for the comparison of two scope sequences is $O(mn^3)$. Since efficient retrieval algorithms are generally desired, it is worth analysing how performance can be improved. Abandoning the scopes' order, time complexity can be reduced to $O(1)$ by computing a histogram of the occurring scopes. This also leads to constant space complexity for the description, since only the frequencies of occurring scopes have to be described.

It is worth mentioning that when dealing with simple, closed polygons not all of the 133 scopes introduced in Sect. 2.2 actually occur. By confining ourselves to polygons with a mathematically positive order only 86 out of the 133 distinguishable scopes can be realised (Fig. 3 left). That is, only the frequencies of this reduced set of scopes have to be determined.

2.4 Visualising Scope Histograms

As mentioned before, an important property of scope histograms is that they allow two polygons to be described and to be compared with constant complexity. Another important property is the fact that a simple visualisation for scope
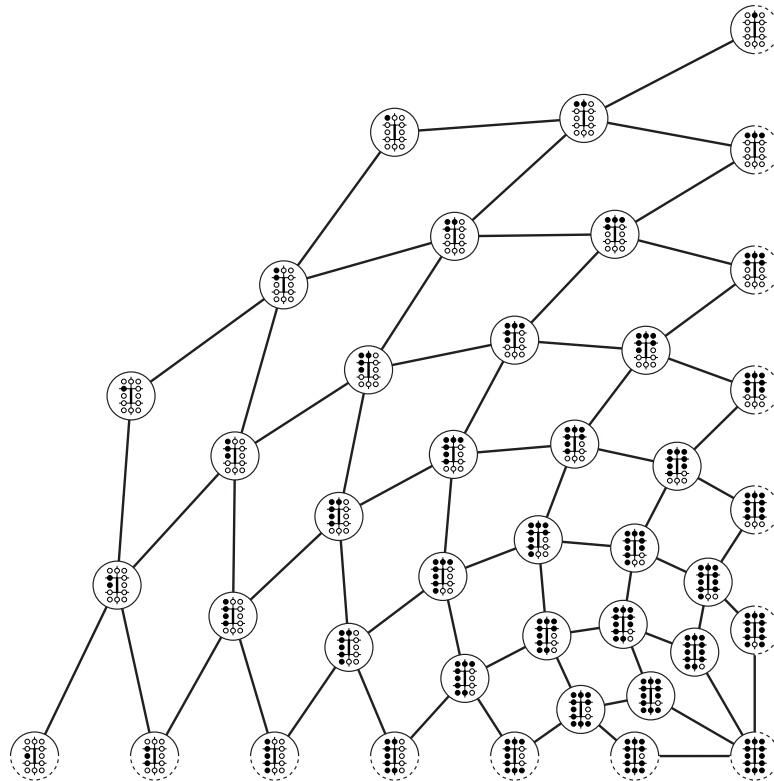
Fig. 2. The top left quadrant of the scope's conceptual neighbourhood structure. The missing quadrants can be created by reflecting the image along the bottom axis and the right axis. Two adjacent scopes can be transformed into one another by shortening or elongating one scope by exactly one atomic relation

histograms exists. For this visualisation we arrange the scopes accordingly to the conceptual neighbourhood structure (Fig. 2). As depicted on the right hand side of Fig. 3 only the position of the 86 realisable scopes have to be taken into consideration. The size of each entry in the histogram's visualisation depends on how often the scope appears for a given polygon. Fig. 4 shows two examples.

Since our visualisation arranges the histogram's entries in accordance to their position in the conceptual neighbourhood structure, even the non expert is able to estimate how similar two entries are. This similarity can directly be derived from their distance in the visualisation. Extending this observation to histograms as a whole, it is even possible to judge the similarity of two polygons solely on the basis of their histograms' visualisations. Another property of this visualisation is that it can easily be determined whether two scopes are mirror images of each other. This is the case if their entries in the histogram can be mapped onto each other by reflecting them along the middle axis. As depicted in Fig. 4, polygons
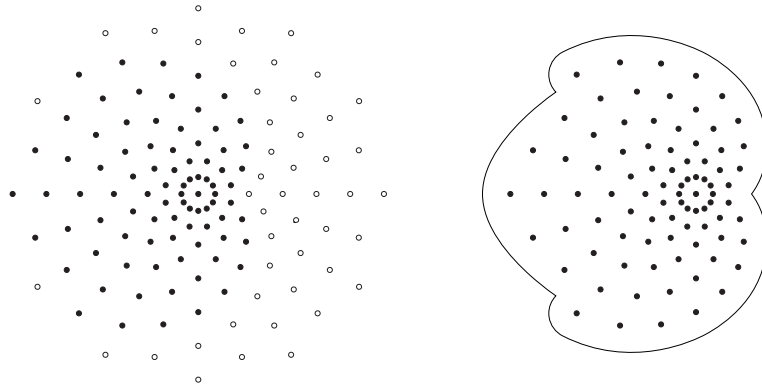
Fig. 3. Left: Each of the 133 distinguishable scopes is depicted by a point which is located accordingly to its position in the conceptual neighbourhood structure. The 86 scopes that are realisable by simple, closed polygons are depicted opaque, the others transparent. Right: Scope histograms can be visualised by only taking into account realisable scopes
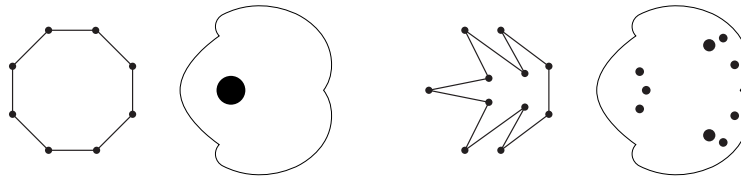

Fig. 4. Two polygons and their respective scope histograms. The size of an entry's depiction correlates with its frequency. Left: Since the polygon has the same position w. r. t. each of its segments, there exists only one (particularly large) entry in the histogram. Right: This polygon has different positions w. r. t. its segments. As a consequence, its scope histogram also contains several different entries

that have an axis of symmetry have a histogram which is symmetric w. r. t. its middle axis. In contrast, not every scope histogram having this property describes a reflection-symmetric polygon. With some experience, it is to some extent possible to identify scopes only by their position in the histogram. It is at least possible to get a coarse impression of the properties of a scope, relating to the correspondence of structure (e.g. large entry for the universal scope) and semantics (parts of the polygon circulating around other parts).

3   Categorising Objects

After having introduced the scope histogram in Sect. 2 we will now concentrate on the evaluation of its performance before we discuss its visualisation in the following Sect. 4.

A well-known method for the comparison of shape classification approaches is the core experiment CE-Shape-1 [10] for the MPEG-7 standard. This experiment compares approaches only by their retrieval results, instead of comparing the underlying methods directly. We apply especially part B which tests the capability of similarity-based retrieval techniques with a database of 1400 images. These are semantically grouped into 70 shape classes, each one consisting of 20 object instances. During the test all 1400 images are used as a query one after another. For each query all images in the database are ordered w. r. t. their similarity by the approach under consideration. Subsequently, the images belonging to the same class as the query are counted within the first 40 results. Since each class contains 20 instances, for each query up to 20 correct matches can be found. For all 1400 queries the total number of correct matches is 28000. The result of the test is the ratio of the number of found objects and the total number of correct matches.

In order to measure the performance of the scope histogram, we apply features that pertain to the same class of complexity required for comparisons, i. e. that they also allow two shapes to be compared with constant time complexity. Especially we use the compactness [3], which corresponds to the ratio $\frac{4\pi A}{P^2}$ of area and perimeter, the radius ratio $\frac{R_{min}}{R_{max}}$ of the minimum enclosing circle and the maximal contained circle [5], and the aspect ratio $\frac{H_r}{W_r}$ of the minimal enclosing rectangle [3]. Each of these features characterises the shape of an object by a single number and therefore compares two objects with constant time complexity. The performance of these features as well as that of the scope histogram is listed in Table 1. The results show that the numeric features, namely compactness, radius ratio, and aspect ratio gain results between about 16% and 24%. In contrast, our scope histogram achieves results of about 46% and therefore outperforms the other approaches. Better results can be achieved by combining features. The results show that a combination of the three numeric features already gains about 52% correct matches. Combining these features with the scope histogram almost 64% can be achieved. Eventually, it is worth mentioning that this retrieval result is only about twelve percentage points less than the 76.46% achieved by the correspondence of visual parts of [9]. However, this latter approach has a significantly higher time complexity of $O(mn^3)$ for the comparison of two objects, while the 64% are achieved with constant time complexity.

Table 1. Classification results of compactness (CO), radius ratio (RR), aspect ratio (AR), and scope histogram (SH) for CE-Shape-1 Part B. Furthermore, the classification has been evaluated for the combination of all numeric features (NF) as well as their combination with the scope histogram (NS)

CO	RR	AR	SH	NF	NS
21.86	16.82	24.12	45.52	51.58	**63.75**

There is yet another advantage of scope histograms, namely that they allow prototypes of categories to be defined. This is useful since a common categorisation technique is based on the definition of clusters which define classes by training examples. In order to analyse whether the scope histogram qualifies itself as such a clustering method, we shall define clusters upon the MPEG dataset and rerun our evaluation on this basis. For this purpose the average of the values of all features is taken for each class. For the scope histogram we determine the average of the corresponding entries. In proceeding this way we define a number of 70 prototypes, one for each class. Using these prototypes we achieve the results listed in Table 2. It shows that the classification results of the three numeric features do not significantly change when they are solely applied in a clustering scenario. By contrast, the scope histogram's results can be improved by 20 percentage points to 66%. The scope histogram now even outperforms the combination of the three numeric features, which achieve together about 63%. A combination of the numeric features and the scope histogram leads to almost 83% correct matches.

Table 2. The classification results of Table 1 can be improved if a prototype is computed for each class of the MPEG test dataset

CO	RR	AR	SH	NF	NS
22.14	15.43	24.43	65.57	62.57	**82.93**

Summarising these results, the scope histogram combined with other features offering constant time complexity achieves about 64% in the MPEG test. By computing a prototype for each class the retrieval results for the MPEG test dataset can even be improved to almost 83%. As elaborated above, this result reflects the scope histogram's qualification as a clustering method.

4 Discussion

In this section we will take a closer look at some selected scope histograms. On this basis we will discuss how the visualisation introduced in Sect. 2.4 supports decisions on whether the application of our approach is useful or not, given a number of classes which are to be distinguished. For this purpose we analyse the classes from the MPEG test dataset and especially their prototypes. The prototypes' scope histograms as well as one example instance from each class are depicted in Figs. 5 to 7.

Comparing the polygons it shows that the prototype of the "Spring" class can be distinguished from all other prototypes by the universal scope which frequently occurs in this class; this scope consists of all twelve atomic $\mathcal{BA}_{12}$ relations, and it represents those line segments which are completely circulated by their polygon's course. In the case of the spring class, the frequency of the

universal scope can be attributed to those segments which are located within its ends. This is a particular salient feature of this class in that the universal scope appears in no other class so frequently. Analysing all classes the universal scope is found rather seldom. Further remarkable occurrences can be determined within the body of the "Sea Snake", in the inner half of the "Horseshoe", at the inner side of the tail of the "Lizzard", and within the handle of the "Cup" class.

Another notable scope is that one of the $\mathcal{BA}_{23}$ relation C_l (Fig. 1), i. e. the relation containing all atomic relations located on the left side of the orientation grid. It does not only appear if line segments are in relation C_l but whenever the polygon is located completely left w. r. t. the reference segment. Since this scope describes a reference segment that is convex w. r. t. its polygon (i. e. the segment is part of the polygon's convex hull), it is not surprising that it appears much more frequently than the universal scope. As an example, it has remarkable occurrences in the "Bottle", "Cellular Phone", "Face", and the "Pencil" classes. It also appears quite often in the scopes of the "Apple" and the "Pocket" class. As objects pertaining to these classes show, apples and pockets have in common that they have a big round body with a significantly smaller part on the top. This is reflected in their scope histograms, which show that there are five entries in the case of the apple and similar frequencies at the very same entries in the scope histogram of the pocket. It is worth mentioning, that due to the fact, that we do not consider the order in which the scopes appear, there are yet other classes which have similar scope histograms. This holds in particular for the "Heart" class and the "Jar" class.

Apparently, the prototypes of some classes are described by very simple scope histograms, i. e. they are made up of only a few different scopes. This can be explained by the fact, that all instances in these classes have very similar scope histograms. On the one hand this brings in the advantage that distances between two instances of this class are small. On the other hand this poses a problem if the dominant scopes are very common and if their combination appears in other classes, too. Examples for classes that do not contain dominant scopes and which have similar scope histograms are the "Guitar", the "Key", as well as the "Spoon" class.

Another problem arises from classes containing instances with completely different shapes. This is due to the fact that the classes for the MPEG test dataset have been compiled semantically, so that it is impossible to classify shapes correctly using solely shape information. Examples for classes of this kind are the "Elephant" class as well as the various "Device" classes.

Coming to a conclusion, two classes of shapes can be distinguished very well by our method if their scope histograms have frequencies which are both dominant and different. It has been shown that the visualisations of the scope histograms enable us to directly judge how promising their application is. Defining those visualisations only on the basis of the size of the spots depicting frequencies in the histograms, such diagrams are not difficult to interpret. Moreover, in contrast to other histograms which are used to show, for example, the colour distribution of images, the entries of the scope histogram have been arranged

in a specific way. By this means relations between semantics and structure are made explicit in that the position of a spot in the diagram reflects the complexity of the polygon w. r. t. specific reference segments as well as to specific shape properties such as the convexity for specific parts.

5 Summary

We introduced a new approach for the description and comparison of objects by their shape. Due to its intuitive diagrammatic visualisation, it can be easily estimated whether the application of our approach is promising or not. Besides, our approach offers a constant time complexity for the comparison of two objects. It shows that by combining it with other features offering the same time complexity retrieval results of about 64% can be achieved in the MPEG test. Furthermore, retrieval performance for the MPEG test dataset can even be improved to almost 83% by the application of prototype shapes.

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. 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. Reasoning about Intervals in Two Dimensions. In W. Thissen, P. Wieringa, M. Pantic, and M. Ludema, editors, *IEEE International Conference on Systems, Man and Cybernetics*, pages 5324–5332, The Hague, 2004. IEEE Press.
7. B. Gottfried. *Shape from Positional-Contrast — Characterising Sketches with Qualitative Line Arrangements*. Deutscher Universitäts-Verlag, 2007.
8. J. H. Larkin and H. A. Simon. Why a Diagram is (Sometimes) Worth Ten Thousand Words. *Cognitive Science*, 11:65–99, 1987.
9. L. J. Latecki and R. Lakämper. Shape Similarity Measure Based on Correspondence of Visual Parts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 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 International Conference on Computer Vision and Pattern Recognition*, pages 424–429, 2000.
11. 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.
12. K. Zimmermann and C. Freksa. Qualitative Spatial Reasoning Using Orientation, Distance, and Path Knowledge. *Applied Intelligence*, 6:49–58, 1996.
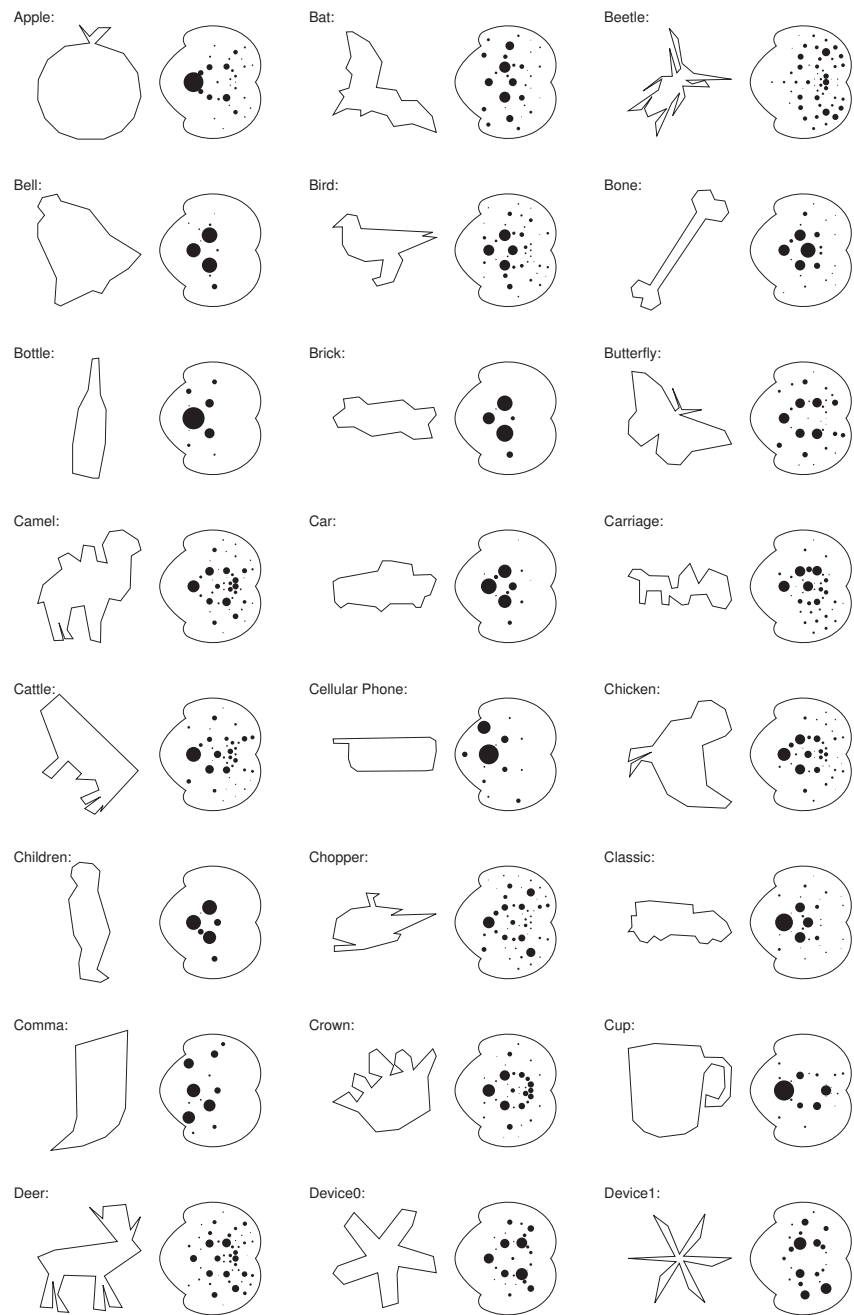
Apple:

Bat:

Beetle:

Bell:

Bird:

Bone:

Bottle:

Brick:

Butterfly:

Camel:

Car:

Carriage:

Cattle:

Cellular Phone:

Chicken:

Children:

Chopper:

Classic:
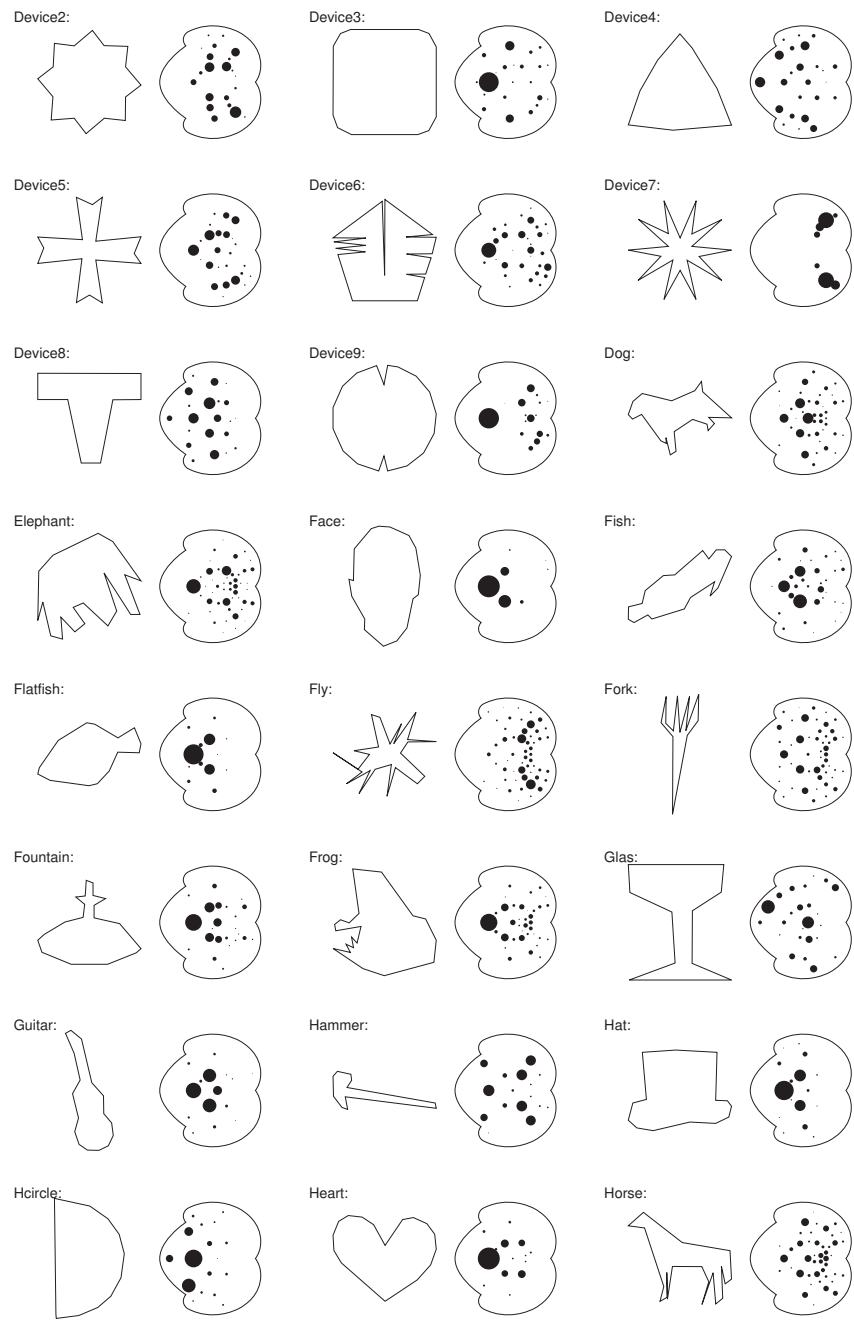
Comma:

Crown:

Cup:

Deer:

Device0:

Device1:

Fig. 5. Histogram prototype and example instance for each MPEG class (Part I)

Fig. 6. Histogram prototype and example instance for each MPEG class (Part II)

Horseshoe:

Jar:

Key:

Lizzard:

Lmfish:

Misk:

Octopus:

Pencil:

Personal Car:

Pocket:

Rat:

Ray:

Sea Snake:

Shoe:

Spoon:

Spring:

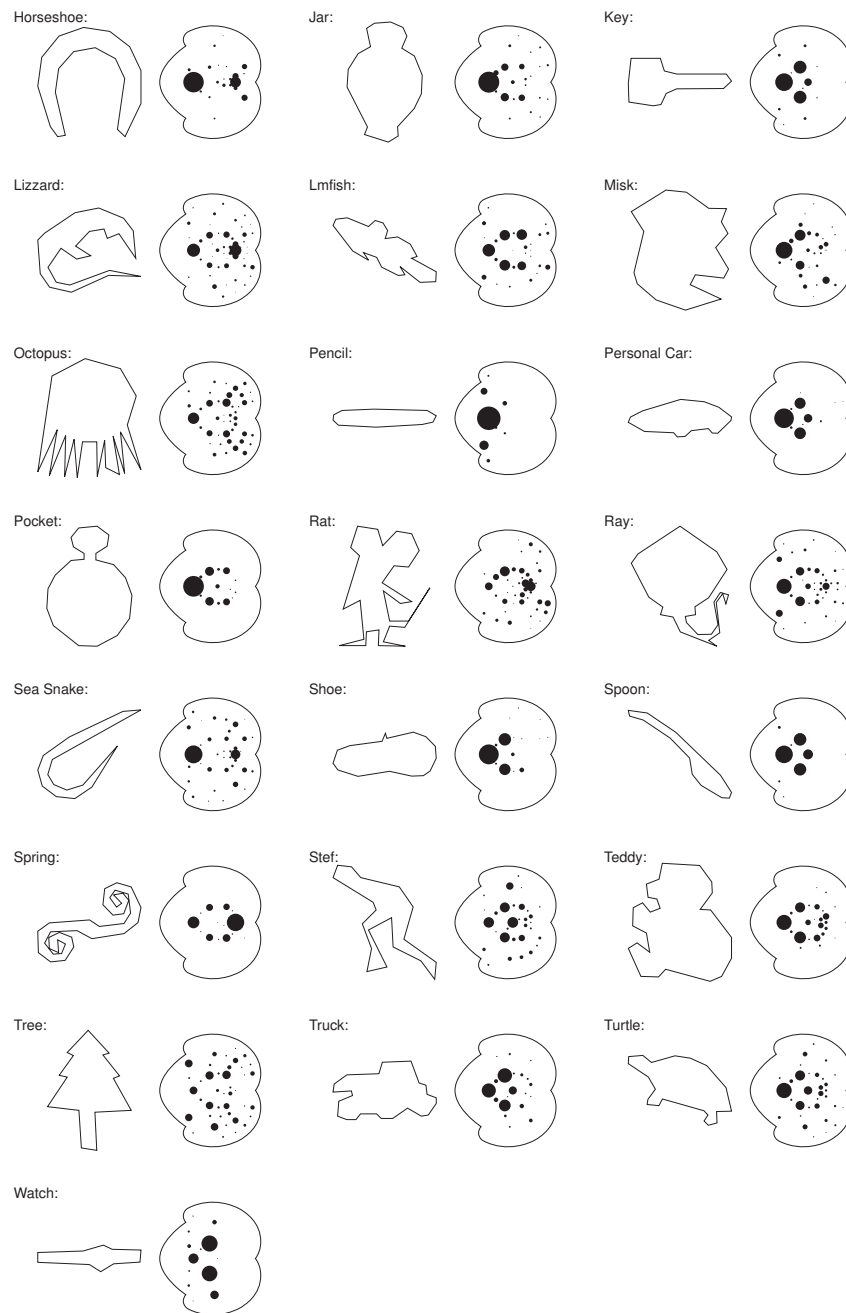Stef:

Teddy:

Tree:

Truck:

Turtle:

Watch:

Fig. 7. Histogram prototype and example instance for each MPEG class (Part III)