Available online at www.sciencedirect.com

ScienceDirect

Computer Vision and Image Understanding 110 (2008) 117–133

Computer Vision and Image Understanding

www.elsevier.com/locate/cviu

Qualitative similarity measures—The case of two-dimensional outlines

Björn Gottfried

Centre for Computing Technologies, University of Bremen, P.O. Box 33 04 40, 28334 Bremen, Germany

Received 6 December 2006; accepted 14 May 2007
Available online 26 May 2007

Abstract

In this paper qualitative similarity measures are introduced. Depending on the underlying representation such similarity measures are based on specific qualitative distinctions which are frequently motivated by perceptual clear distinctions. Here, we discuss one such representation and show how it applies to different domains. In particular, qualitative methods are useful as soon as specific qualitative features can be defined for the purpose of characterising specific objects. Accordingly, we set two examples, namely for a domain of historical objects and for the geographic domain. Afterwards, however, we also demonstrate that our qualitative representation performs quite well when applied to a well-known test data set, without specifying any specific features. Instead, frequencies of qualitative relations are taken into account. The results indicate that qualitative measures not only relate to distinctions which can be easily comprehended by vision but that they are especially efficient in terms of runtime complexity, both issues being of particular importance in the case of image databases.

Keywords: Qualitative representations; Qualitative similarity measures

1. Introduction

In computer vision, similarity matching is frequently viewed as a matter of taking a number of measurements using quantitative reference systems: given two vectors (each one describing an image or a single object); we suppose that each vector consists of n attributes. Then, the question is how large their difference is. Similarity measures which are based on real valued attributes reflect how similar objects are along n quantitatively partitioned spectra, each of which comprises a metric on some particular granularity level. The similarity between objects, then, can be determined with corresponding precision. From what follows, we are concerned with the usual n-dimensional feature space.

One might ask whether precision is always that important; indeed, whether quantitative reference systems are to be taken at all; or, is there an alternative approach for the purpose of determining similarities in computer vision? The class of methods we advocate in this article, do in fact also pertain to methods at the feature space level—though, what distinguishes them is that they rely on differences in kind, rather than of measurement. According to Freksa [5], this means taking into account qualitative distinctions obtained by comparing features within the object domain rather than by measuring them in terms of some artificial external scale. In this sense, qualitative features are of a relative kind where the reference entity is a single value rather than a whole set of categories. For instance, in a scene we distinguish whether two points lie on the same side with respect to a line (the reference entity) or whether they lie on different sides: we focus on how features are altogether arranged, and we determine their ordering in the two-dimensional plane but omit any quantitative distinctions. As a consequence, a new category of appropriate similarity measures is required. From now on, we shall speak of measures only in the sense of those qualitative distinctions.

It is not our purpose to propose just any alternative class of similarity measures. Our concept of *qualitative similarity measures* distinguishes itself to be robust and efficient. However, we have to pay for it with imprecision. Interestingly, imprecise descriptions are frequently sufficient, and

E-mail address: bg@tzi.de

1077-3142/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.cviu.2007.05.002

in particular, frequently related to the human's visual system, who can comprehend such qualitative, imprecise features better than precise quantitative measurements (we are able to state whether two points lie on the same side of a line or not, but we are not able to state how far they are apart). This is where our method will be useful: either coarse descriptions are sufficient for our purposes, or features have to be provided which can be comprehended by the user.

In order to relate our method to the state of the art we shall finally compare it with a number of quantitative approaches. Since many methods have been devised in computer vision we have to select some of them; we do this by focusing on those methods which fall into the same complexity class (concerning the costs for the comparison of two-dimensional outlines). Also, there exist some qualitative methods which have been proposed over the last decade. These include Cohn [1], who describes a region-orientated technique that distinguishes different concave shapes by considering the notion of the connection of regions and their convex hulls; Schlieder [21], who introduced a point-oriented approach by describing how triples of vertices of a polygon are related relative to each other; Galton and Meathrel [6] proposed a representation of outlines by means of strings over an alphabet of seven qualitative curvature types. Eventually, the slope projection approach of Jungert [17] maps the vertices of polygons onto both the x-axis and the y-axis; depending on the ordering of vertices on these axes several features can be derived, for instance, whether a vertex forms a convex or concave part of a closed polygon. All these approaches characterise outlines qualitatively, as we will do below. They are detailed and discussed in the context of our method in [15]. It shows that opposed to quantitative approaches qualitative approaches can be easily compared on a conceptual level. This allows their representational expressiveness easily to be determined, and as a consequence, what features they are able to distinguish.

In Section 2 we shall motivate the employment of qualitative features by showing shape properties we are going to look at. Section 3 introduces our qualitative feature scheme. Sections 4.1 and 4.2 provide example applications which show the usefulness of qualitative similarity measures in different domains; Section 4.3 shows how the qualitative representation performs quite well in comparison to other well-known approaches. Section 5 analyses how stable the approach is, in particular when being faced with different degrees of precision, distortions, changes in viewpoint, and segmentation errors. Then, Section 6 defines the concept of qualitative similarity measures. We conclude with a discussion about granularity and complexity issues in Section 7.

2. Accessing qualitative features

The motivation for describing qualitative features of objects is as follows. Collections of objects d'art, historical tools, or natural objects such as in the geographic domain are some examples of collections to which experts want to access efficiently. Most notably, such collections show a number of objects from the same category, each exemplar showing specific details being typical for that exemplar. Frequently, the expert has the visual appearance of those specifics in mind. Therefore, a query-by-sketch system, such as that described in [14] which puts emphasis on specific shape qualities, would provide an appropriate starting point in searching for particular objects. In the following, two scenarios are described that motivate the employment of the methods we are going to present afterwards.

2.1. Scenario I

The *Bamberger Naturkundemuseum* owns a 200-year-old collection of malaceous and stone fruits (Fig. 1 shows some exemplars). These wax fruits were manufactured and sold by the *Landes-Industrie-Comptoir* of Bertuch (1747–1824). A collection such as this one in the pomological cabinet in Bamberg is of great interest to pomologists, who want to know what kinds of apples existed in the past. The collection is a considerable archive of old fruits; often common in their time, many of which have subsequently become extinct. For the purpose of identifying old fruits which are rediscovered in nature, the Bertuch-collection is an important, indeed indispensable, source. It is exclusively the visual appearance which is conserved by wax fruits and it is therefore only this to which we have access. Consequently, we have to specify visual properties of a fruit to identify it. Besides colour and texture which both fade gradually over the years, it is primarily the shape of an apple which distinguishes it from others, and it is therefore the shapes of objects in which we are interested in.

Images of the objects of the Bertuch-collection have been taken at the Bamberger Naturkundemuseum. They

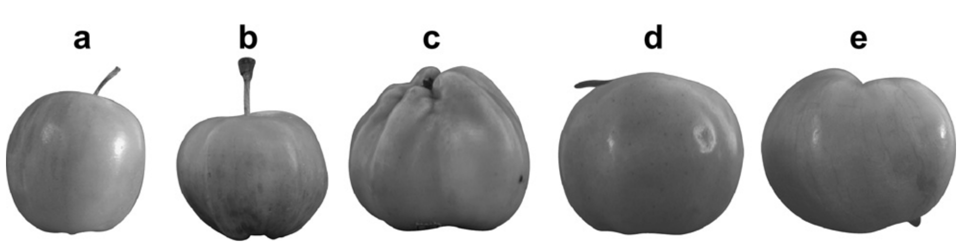

Fig. 1. Apples of the pomological cabinet, showing the expert differences in shape.

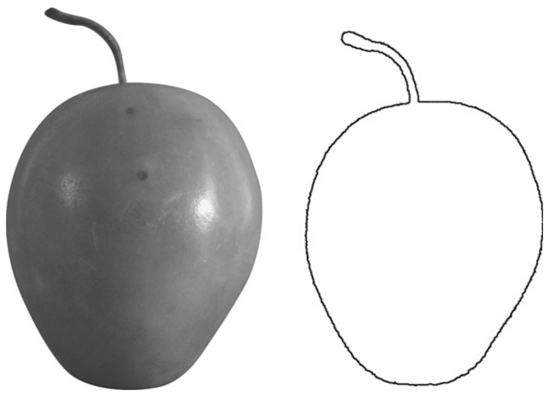

Fig. 2. An *Italienischer Weisser Rosmarienapfel* and its contour (right).

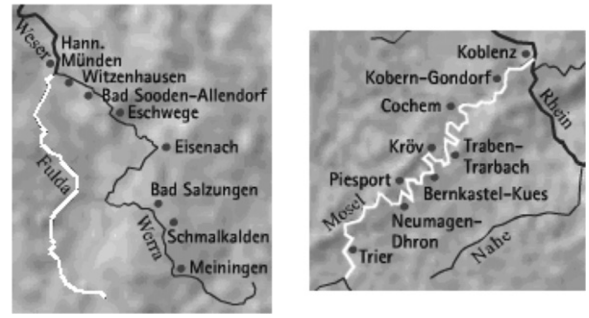

Fig. 3. Rivers differing regarding their meanderings.

provide a database which shows fruits pertaining to 27 different types of apples. Each fruit relief can be specified by the following features: (a) the overall shape is round (Fig. 1b) or vertical-shaped (Fig. 1a) or wide (Fig. 1e); (b) hills at the top (Fig. 1c) or bottom or both or neither; (c) dents at the top (Fig. 1e) or bottom or both or neither; (d) a straight (Fig. 1b) or bent stalk (Fig. 1d) or no stalk shown (Fig. 1c). These features combine to a configuration space with 144 classes. For example, the *Italienischer Weisser Rosmarienapfel* has a vertical-shaped body and its stalk is bent (see Fig. 2).

2.2. Scenario II

In geographical information systems topological relations between geographical objects are frequently used [4]. It may, for example, be crucial to a particular query that there is a forest and that there is a river which is not connected to the forest; but it does not matter at all what the boundary of the forest looks like, or how far the river is from the forest provided that they are not in contact; such geometrical relationships are not important when we are interested in those cases where only the given topological relationships hold. Precise correspondences would retrieve fewer results than there are actually in the database. But sometimes topological relationships do not sufficiently characterise the query. For example, it might be crucial to take the curve progression of linear objects into account, such as rivers, contour lines in topographic maps, coastlines, borders of countries and other regions, transportation networks, such as roads and railways, irrigation networks, and sewer systems. See as an example the rivers in Fig. 3, which are to be distinguished regarding their meander. Here, we address the problem of how to represent smooth curves by polygons in such a way that the curve progression of linear objects is made explicit.

One particular important field is spatial planning: for a town planner, for example, it is of interest how objects in geographic space spread out, how they determine the spatial layout of their environment, so that it becomes possible to state how other objects can be related to them. Then, looking for specific configurations in geographic space the town planner puts emphasis on how objects are shaped. Instead of precise shape descriptions, however, only a number of qualitative distinctions are of importance for him, whether a river meanders through a town in a twisty way or whether it smoothly takes its direction, for example. Fig. 3 shows two rivers which differ with respect to how they meander through their environment. While the first scenario shows how to deal with imprecise queries by qualitative features, this second scenario demonstrates how this very same description even allows specific Gestalt features to be dealt with. But instead of taking a classical approach for determining Gestalt features such as the curvature it is then possible to determine these features in a query-by-sketch system applied to imprecise shapes as shown in the first scenario. Other features can be defined similarly like meanders on the representation which will be introduced in the next section.

2.3. Object of interest

There are two main challenges a query-by-sketch system has to cope with. First, objects which belong to the same category (and which therefore have quite a similar appearance to each other) need to be distinguished. These objects will, however, obviously differ in some characteristics to which the expert attaches great importance (apples with a round body versus those with a wide body or rivers which are twisty versus those which are smooth). Second, these characteristics are to be specified in a simple but reliable way (both the apples' bodies and the rivers' curves will be drawn only roughly in a sketch). There is, necessarily, a trade-off between similarity and variety, the latter relating to those characteristics which are important only to the expert. These characteristics concern those properties that the expert is particularly looking for, and as a consequence which he will take care to specify graphically. Their qualitative description will be made possible by the qualitative feature scheme of polygonal outlines we will present in section 3. That is, the shapes will be approximated by polygons and the parts of single shapes are described by line segments. This is the reason why our method will be based

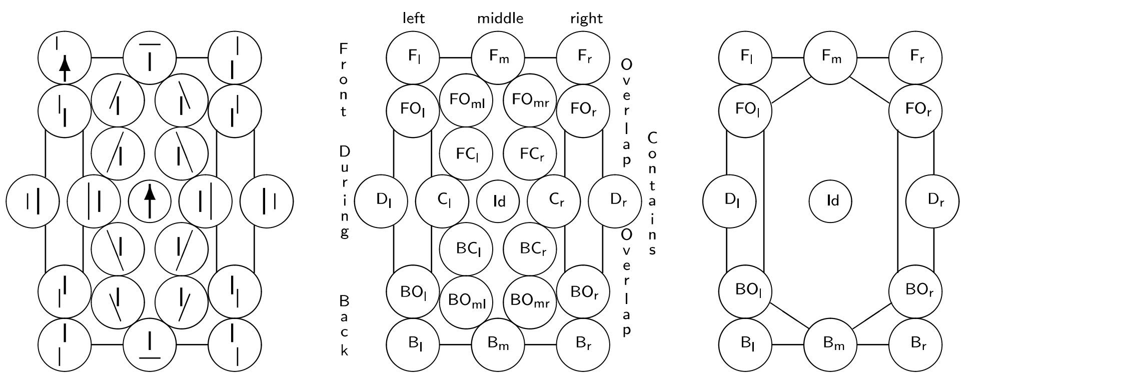

Fig. 4. Conceptual neighbourhood graphs of $\mathcal{BA}_{23}$ (middle) and $\mathcal{BA}_{13}$ (right); example instantiations of all relations (left), the bold vertical segment being the reference segment which is oriented upwards as indicated in two cases (for Id and F_l).

on line segments and how they can be arranged in the picture plane for defining specific shape features.

3. A qualitative feature scheme

In this section we will summarise previous work on a qualitative feature scheme. It serves as an example approach to which we shall apply the more general notion of qualitative similarity measures. For the purpose of providing a manageable overview on how qualitative representations are defined and on how they are employed, we focus on the description of linear objects (open as well as closed objects) which can be approximated by polygons. Accordingly, our feature scheme consists of a number of relations which describe arrangements between line segments (of polygons) in two dimensions. Such relations have been referred to as *bipartite arrangements*, $\mathcal{BA}$ for short [10]. Fig. 4 shows these relations. They relate parts of a contour to other parts of that same contour.

3.1. Course

If one line of a polygon is made the basis, the position of every other line can be described relative to it, using the relations of $\mathcal{BA}_{23}$ which are shown on the left-hand side of Fig. 4. In this way, the qualitative context of a polygonal line, x, is considered. Conceiving the polygon to be oriented anticlockwise each reference segment has an orientation which allows the positions of other segments to be described relative to it, those segments being at its front or back, lying towards its left or right, and so on. In comparison to the reference segment the orientations of all other segments are not considered, only their positions relative to the reference segment are taken into account. Consequently, the relations are defined by projecting both the endpoints of a segment orthogonally onto the line going through the reference segment. This enables one to distinguish whether those endpoints lie in the front of the reference segment (F), back to it (B), or somewhere in between (D). Additionally, this projection comes either from the left (l) or right (r), determining the second dimension (found in the indices of the relations). Those relations found on the left-hand side of Fig. 4 can be distinguished by this technique by combining both dimensions (resulting in for instance F_l), and by combining both endpoints (resulting in for instance F_l if both endpoints lie F_l of the reference segment, or in, for instance, C_l if one endpoint lies at F_l and the other one at B_l). Table 1 shows the definition of all relations in accordance to the regions defined in Fig. 5. Note that these relations are in particular sufficient for the purpose of characterising simple polygons in Section 4.1.4.

For a polygon with n lines we obtain a list of relations which we refer to as the *course* of reference segment x, in short C(x):

Definition 1. Course

x is line segment of a simple polygon P. Its course is denoted by C(x) and describes all $\mathcal{BA}_{23}$ relations between all lines of P and x:

$$C(x) \equiv (x_{y_1}, .., x_{y_n}), \quad x_{y_i} \in \mathcal{BA}_{23}, \quad i = 1, .., n \tag{1}$$

with x_{y_1} meaning that line segment y_1 is described with respect to reference segment x.[1] In particular, it holds that $x_x = \mathrm{Id}$, i.e. whenever a line segment is related to itself the identity relation holds. A subset $\mathcal{BA}_{13} \subset \mathcal{BA}_{23}$, shown on the right-hand side of Fig. 4, provides a set of atomic relations in the sense that all other relations can be obtained by combining $\mathcal{BA}_{13}$ relations which relate to adjacent regions (as shown on the left-hand side of Fig. 5). Any connected chain of relations according to $\mathcal{BA}_{13}$ forms a valid combination and allows any relation $\mathcal{BA}_{23} \setminus \mathcal{BA}_{13}$ to be represented. In particular some relations can be represented differently. For instance $C_l = BO_lD_lFO_l$ or $C_l = BO_lFO_l$ depending on how the line segments relate to the reference segment under consideration (cf. the right-hand side of Fig. 5). This shows that the similarity of such different arrangements is captured by the relation C_l which is more abstract than the atomic $\mathcal{BA}_{13}$ relations. Note that this mode of combination excludes cases of non-adjacent

[1] Instead of the common infix notation in a relational diction we use indices in order to be able to list many $\mathcal{BA}_{23}$ relations in a compact way.

Table 1
Definition of $\mathcal{BA}$ relations between some segment p and a reference segment q; p and q either coincide (it holds Id) or they are disjoint (and have no point in common); p_1 and p_2 simply denote both endpoints of line segment p

Name	Meaning	Definition
Id	Identity	$p_1 = q_1 \wedge p_2 = q_2$
F_l	Front left	$p_1 \in F_l \wedge p_2 \in F_l$
FO_l	Front-overlap left	$p_1 \in F_l \wedge p_2 \in D_l$
D_l	During left	$p_1 \in D_l \wedge p_2 \in D_l$
BO_l	Back-overlap left	$p_1 \in D_l \wedge p_2 \in B_l$
B_l	Back left	$p_1 \in B_l \wedge p_2 \in B_l$
B_m	Back middle	$p_1 \in B_l \wedge p_2 \in B_r$
B_r	Back right	$p_1 \in B_r \wedge p_2 \in B_r$
BO_r	Back-overlap right	$p_1 \in B_r \wedge p_2 \in D_r$
D_r	During right	$p_1 \in D_r \wedge p_2 \in D_r$
FO_r	Front-overlap right	$p_1 \in D_r \wedge p_2 \in F_r$
F_r	Front right	$p_1 \in F_r \wedge p_2 \in F_r$
F_m	Front middle	$p_1 \in F_r \wedge p_2 \in F_l$
FO_{ml}	Front-over. mid.-left	$p_1 \in F_r \wedge p_2 \in D_l$
FC_l	Front-contains left	$p_1 \in F_r \wedge p_2 \in B_l$
C_l	Contains left	$p_1 \in F_l \wedge p_2 \in B_l$
BC_l	Back-contains left	$p_1 \in F_l \wedge p_2 \in B_r$
BO_{ml}	Back-over. mid.-left	$p_1 \in D_l \wedge p_2 \in B_r$
BO_{mr}	Back-over. mid.-right	$p_1 \in B_l \wedge p_2 \in D_r$
BC_r	Back-contains right	$p_1 \in B_l \wedge p_2 \in F_r$
C_r	Contains right	$p_1 \in B_r \wedge p_2 \in F_r$
FC_r	Front-contains right	$p_1 \in B_r \wedge p_2 \in F_l$
FO_{mr}	Front-over. mid.-right	$p_1 \in D_r \wedge p_2 \in F_l$

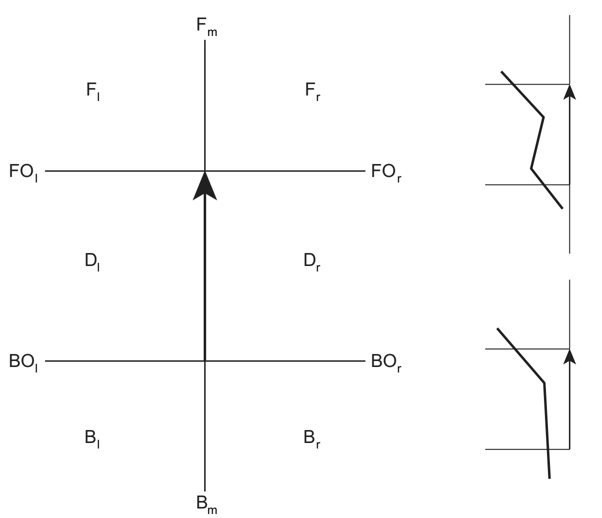

Fig. 5. (Left) The distinguished regions induced around some reference segment. (Right) Two different representations of the C_l relation ($BO_lD_lFO_l$ and BO_lFO_l).

$\mathcal{BA}_{13}$ relations as F_lF_r which do not form a connected path of relations around Id. In order to get from F_l to F_r in an outline shape one has to pass through F_m. This is the line which goes through the reference line and which divides the plane into left and right with respect to the reference line. Note that there might be the particular case of a line segment starting or ending precisely on the F_m position. In this case one passes at this endpoint through F_m.

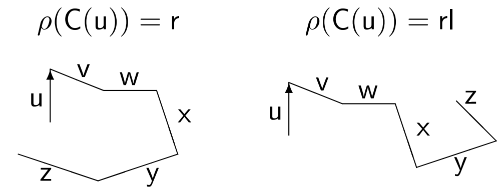

Fig. 6. Different circulation directions.

Having a polygon with n line segments there exist n courses, each one comprising n bipartite relations, i.e. such a feature scheme comprises a total of n^2 relations. Writing down all courses, one below the other, for a polygon with six lines the following matrix, M, is obtained (compare the polygon on the left-hand side of Fig. 6, and note that singular relations, such as between x and y, are dealt with in accordance to Gottfried [11]):

$$\begin{array}{cccccccccccccc}
\mathrm{Id} & u_v & u_w & u_x & u_y & u_z & : & \mathrm{Id} & D_r & D_r & BO_r & B_r & B_m \\
v_u & \mathrm{Id} & v_w & v_x & v_y & v_z & : & D_r & \mathrm{Id} & F_l & F_m & F_r & C_r \\
w_u & w_v & \mathrm{Id} & w_x & w_y & w_z & : & B_m & B_l & \mathrm{Id} & F_r & C_r & B_r \\
x_u & x_v & x_w & \mathrm{Id} & x_y & x_z & : & B_r & B_r & B_r & \mathrm{Id} & D_r & D_r \\
y_u & y_v & y_w & y_x & \mathrm{Id} & y_z & : & F_r & FO_r & D_r & D_r & \mathrm{Id} & F_r \\
z_u & z_v & z_w & z_x & z_y & \mathrm{Id} & : & D_r & D_r & BO_r & B_r & B_r & \mathrm{Id}
\end{array}$$

3.2. Scope, extent, and scope histogram

The entire range of relations (according to $\mathcal{BA}_{13}$) where C(x) runs along is called its scope, and the number of different $\mathcal{BA}_{13}$ relations involved is called the extent, $\eta(C(x))$ for short. The shortest extent is 0 in which case there is no other line segment than the reference segment x itself; the largest possible extent is 12 (which is $|\mathcal{BA}_{13} \setminus \{\mathrm{Id}\}|$) in which case the course runs completely around x. Scopes of such courses are also referred to as universal scopes since all relations of $\mathcal{BA}_{23}$ are realisable within these scopes. Eventually, there exist different scopes which have the same extent, for instance, one course may go from F_l to F_r and another one from B_l to B_r; both courses, however, have an extent of 3.

Definition 2. Scope of a simple polygon
x is line segment of a simple polygon and C(x) is its course. The entire range of regions where C(x) runs along is called the scope σ of C(x). This range of regions is characterised by a neighbourhood of $\mathcal{BA}_{12}$ relations, starting with $r_1 \in \mathcal{BA}_{12}$ and ending with $r_2 \in \mathcal{BA}_{12}$ ($\mathcal{BA}_{12} = \mathcal{BA}_{13} \setminus \mathrm{Id}$) :

$$\sigma(C(x)) \equiv [r_1 r_2]. \qquad (2)$$

The scope runs around the reference segment clockwise from r_1 to r_2. If the reference segment partitions the polygon such that the scope is interrupted, two scopes are to be given: that one before the reference segment and that one after it. For those two parts we assume that they are self-connected. Otherwise, for each separate self-connected part a scope would be needed.

For the polygon on the left-hand side of Fig. 6, we obtain the following scopes (beginning with the first row): $[D_rB_m]$, $[D_rD_r][F_lB_r]$, $[B_mB_l][F_r\ B_r]$, $[B_rB_r][D_rD_r]$, $[F_rD_r][F_r\ F_r]$, $[D_rB_r]$. Note that for cases such as the second line segment we first of all expand non $\mathcal{BA}_{12}$ relations to $\mathcal{BA}_{12}$ relations (in this case C_r is expanded as $F_rFO_rD_rBO_rB_r$). The length of a scope is called its extent (e.g. $\|[D_rB_m]\| = 4$, $\|[D_rD_r]\| = 1$, $\|[F_lB_r]\| = 7, \ldots$):

Definition 3. Extent

$\sigma(C(x))$ is the scope of course $C(x)$ of Polygon P. The distance between the start-relation and the end-relation of the scope $\sigma(C(x))$ is called the extent of the scope, or simply the extent of course $C(x)$, denoted by $\eta(C(x))$. The extent is determined with the relations of $\mathcal{BA}_{12}$, and it holds that:

$$\eta(C(x)) \in \{0, 1, \ldots, 12\}. \tag{3}$$

The extent of the polygon is the average extent of all courses:

$$\eta(P) = \frac{\sum_{i=1}^{|P|} \eta(C(x_i))}{|P|}. \tag{4}$$

According to Schuldt et al. [23], 86 scopes for simple, closed polygons are possible. Some example scopes are as follows: $[F_l\ F_l]$, $[F_lF_m]$, $[F_lF_r]$. That is, a polygon can be characterised independently of its number of line segments by its scope histogram with its 86 entries.

Definition 4. Scope histogram

M is the matrix of all n courses of a simple, closed polygon. $H(M)$ denotes the scope histogram of M. Each entry $H_i(M)$ with $i = 1, 2, \ldots, 86$ gives the number of scopes which exist in M.

That is, the number of different scopes (one for each course, i.e. n altogether for a polygon with n line segments) is simply counted and taken as a simple description for the polygon.

For polygon M on the left-hand side of Fig. 6, we obtain the following scope histogram: $H(M)_{[D_rB_m]} = 1$, $H(M)_{[D_rD_r]} = 2$, $H(M)_{[F_lB_r]} = 1$, $H(M)_{[B_mB_l]} = 1$ $H(M)_{[F_rB_r]} = 1$, $H(M)_{[B_rB_r]} = 1$, $H(M)_{[F_rD_r]} = 1$, $H(M)_{[D_rB_r]} = 1$. All other entries of $H(M)$ are 0.

3.3. Circulation direction

It can be distinguished whether a course circulates *left* around or *right* around the reference segment:

Definition 5. Circulation direction

x is line segment of a polygon. The circulation between two neighbouring line segments, y and y′, is left of x, i.e. $\rho(x_y, x_{y'}) = l$, if the path from the first relation x_y to the second relation $x_{y'}$ runs anticlockwise around Id; otherwise, it is right of x, in short r. If the direction does not change it holds that $\rho(x_y, x_{y'}) = \varepsilon$, as it does if y and y′ are not adjacent line segments.

$\rho(x_y, x_z) = \varepsilon$ denotes the case that it cannot be determined from the point of view of x whether the circulation direction from y to z is clockwise or anticlockwise with respect to x. This indeterminacy can either be compensated by another reference segment or y and z are not adjacent in which case their circulation direction cannot be derived from y and z alone. Fig. 6 shows two examples. On the left-hand side the course of x circulates entirely right of x. On the right-hand side it also runs right of x, but it then turns back, and as a consequence, after this turn it runs left of x.

Further examples clarify the meaning of the circulation direction[2]: $\rho(F_lFO_l) = l$, $\rho(FO_lF_l) = r$, $\rho(F_lF_m) = r$, $\rho(F_lFO_{mr}BO_r) = rr$, $\rho(F_lFO_{mr}FO_rFO_rD_r) = rlr$, and $\rho(FO_lF_lIdD_rBO_r) = rr$. Equal neighbouring directions can be omitted in order to obtain only the changes, i.e. changes between left and right or anticlockwise and clockwise, respectively. It then holds that the number of changes between left and right of $\rho(C(x))$ is less than or equal to the length of the course.

Taking each of the line segments of the polygon on the left-hand side of Fig. 6 as a reference segment, the circulation direction to each other line segment is always either r (clockwise) or ε (indeterminate). But the circulation directions of different reference segments of the same polygon are not always equal, as demonstrated by the polygon on the right-hand side of Fig. 6. Here, $\rho(u_w,\ u_x) = \rho(D_r, BO_r) = r$, whereas $\rho(z_w, z_x) = \rho(F_l, FO_l) = l$. The circulation direction depends on the position of the reference segment with respect to the other line segments of the same polygon.[3] Changing the circulation direction is what we will focus on in the next section.

3.4. Reversals

As soon as a course changes its circulation direction from left to right or right to left, the course includes a reversal, as indicated by its first derivative. For example, $C(x) = F_rFO_r\ D_rFO_rF_r$ comprises a reversal since the circulation direction is changed after D_r with the second FO_r relation.

Definition 6. Reversal

x is a line segment of a polygon and $C(x)$ is its course. If $C(x)$ comprises two sections which circulate in different directions around x it holds that $\varrho(C(x))$, saying that $C(x)$ contains a reversal.

[2] Note that in the following we simplify the notation and write $\rho(x_yx_z) = l$ instead of $\rho(x_y,\ x_z) = l$. Moreover, an arbitrary number of relations are allowed meaning that all according circulation directions are to be mentioned, such as $\rho(F_lFO_{mr}BO_r) = rr$.

[3] It depends also on its orientation relative to the other line segments. However, the current work solely relies on relative positions and analyses their relationships.

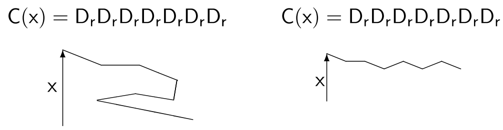

Fig. 7. Two polygons describing the same course with respect to line segment x.

Examples of courses without reversals from the viewpoint of line segment x are depicted in Fig. 8. Examples of courses with exactly one reversal with respect to x are depicted in Fig. 9. As the left-hand side of Fig. 7 shows, all reversals of a polygon cannot always be deduced from single courses. It is rather necessary to test all courses of a polygon in order to determine whether there are reversals in the polygon. An algorithm which determines reversals is proposed in [13]. Eventually, determining line segments at which reversals occur enables reversals to get distinguished according to $\mathcal{BA}_{23}$.

3.5. Summary

To summarise, at the finest level sequences of $\mathcal{BA}_{23}$ relations are considered (Definition 1); expanded as $\mathcal{BA}_{12}$ relations they form together scopes from the point of view of one reference segment (Definition 2), such a scope telling us where a polygon develops regarding this reference segment; the length of a scope is called its extent (Definition 3); scope histograms integrate the scopes of all courses of one polygon (Definition 4); following a list of $\mathcal{BA}_{23}$ relations a course either orbits clockwise or anticlockwise around the reference segment (Definition 5). Taking the first derivative of a course amounts to considering changes in direction, which we refer to as reversals (Definition 6). Positions at which such changes occur are analogous to local minima and maxima of functions and they can be referred to by $\mathcal{BA}_{23}$ relations. While scope and extent simply state *where* a polygon runs along according to the reference segments taken into account, circulation directions and reversals tell us *how* polygons take their way along their scopes.

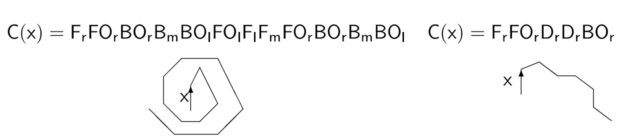

Fig. 8. Polygons without reversals from the point of view of reference segment x.

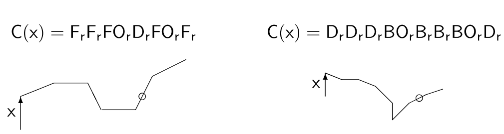

Fig. 9. Polygons with reversals; note that relations are put together when they are equal and adjacent; the circles indicate line segments at which reversals occur.

4. Applications

The following applications demonstrate:

- How to use the qualitative representation in a query-by-sketch system (using the *basis relations* and *scopes* and their *extent*, i.e. $\mathcal{BA}$ relations).
- How it shows to be useful in the geographic domain (employing the *circulation direction* and *reversals*, i.e. Gestalt features).
- How it performs regarding a well known test data set (applying *scope histograms*, i.e. frequency distributions).

All example applications especially show how well the qualitative description can be comprehended by the human operator, regardless of whether using just the basis relations, complex Gestalt features which are based on them, or simply frequency distributions of the qualitative relations. However, in all these applications we focus on describing the silhouettes of objects which can be done regarding this qualitative line segment based representation.

4.1. Scenario I

In the first scenario we consider the objects of the Bertuch-collection. The images in this collection are preprocessed and their features have been saved for efficient access. Both the apples of the collection and graphical queries (sketches) are analysed as follows. Gaps in the sketches are closed by morphological operations [25], the images are binarised (reduced to 1 bit per pixel), the contour is extracted, and eventually approximated by a closed polygon [20]. A granularity level can be defined on the basis of the maximum difference between the original contour and the approximating polygon. The larger this difference is, the coarser the granularity level. Fig. 10 shows an example. We consecutively show how a qualitative and how a quantitative method deals with measuring the similarity of objects (query and database image).

4.1.1. Qualitative approach

At first, the convex parts of an object are determined by walking anticlockwise around the polygon and taking the triangle orientation of three adjacent points: if each triangle orientation is anticlockwise the whole object is convex (in this case there are only left-of relations R_l, $R_l \in \{F_l, FO_l, D_l, C_l, BO_l, B_l\}$); otherwise concave parts are identified by obtaining triangle orientations which are oriented clockwise (then, right-of relations are found $\{F_r, FO_r, D_r, \ldots\}$). The length of the convex parts are used for distinguishing parts of the body (large convex portions) from those pertaining to the stalk (short portions). A number of changes of the triangle orientation along small parts at the bottom or top indicate the existence of single dents or hills.

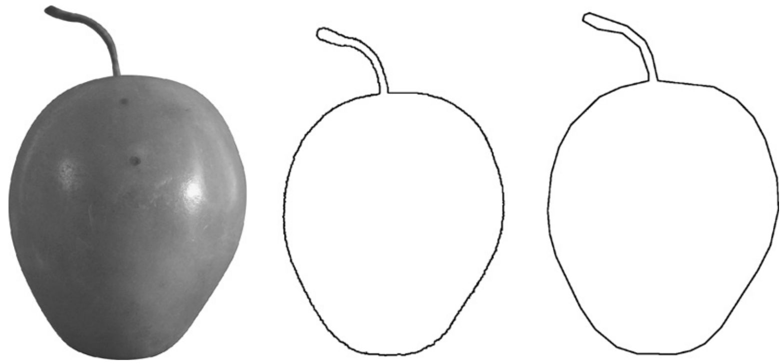

Fig. 10. An apple, its contour and its polygonal approximation.

Second, the qualitative matrix of a polygon is analysed. For the purpose of classifying the body of an apple, its overall shape is determined by considering the convex parts of the body (as described in the previous paragraph) and by looking for those columns in the qualitative matrix in which *overlap*- and *during*-relations (relations such as $\mathsf{D_l}$, $\mathsf{D_r}$, $\mathsf{FO_l}$, $\mathsf{FO_r}$, $\mathsf{BO_l}$, $\mathsf{BO_r}$) mount up along those convexities, showing deviations from a roundish shape and indicating whether these deviations result in a vertical or wide shape regarding the image plane. This is possible just because the representation allows different convex shapes to be distinguished as opposed to other qualitative approaches which are not capable of distinguishing different convexities [12].

In order to distinguish two kinds of stalks we measure the extent of those parts which have been identified as stalks (Definition 3). The extent tells us something about the complexity of a polygon's course (Definition 1), i.e. if line segments are placed equally or differently with respect to a reference segment, and to what extent they surround it. This allows us to distinguish stalks which are almost straight from those which are bent. Further (local) qualitative properties are defined in [8,9,12,14].

4.1.2. Quantitative approach

In order to show how our approach compares to others, it is useful to choose an alternative approach which is also based on polygons. The discussion about the performance of different approaches can then concentrate on the question of how these approaches work on the same representation of an object. In this sense it is equally important to compare our approach with an algorithm which is based on the same representation, at the same granularity level.

By describing and comparing polygons using a classic quantitative geometric approach (specifically, comparing the lengths of, and angles between corresponding line segments in two polygons), we can evaluate our method as an alternative. But before comparing polygons using this quantitative approach, the lengths of the line segments must be normalised with respect to the longest side; this allows polygons which are differently scaled to be compared. For two polygons with the same number of vertices, it is then possible to quantitatively calculate their distance from each other. Since there are n ways to match two polygons with n lines, all these matches are calculated, and the one giving the smallest distance is taken as the result. Two equal polygons have a distance of zero, and the more their angles and the lengths of sides differ, the higher their distance becomes.

4.1.3. Experimental procedure

Three students have participated on a voluntary basis. They had to concern themselves with the shapes of apples in order to get an idea of what shape properties typically exist. After the participants felt confident about the way one could discern different types of apple by shape properties, they had to imagine differently shaped apples which they had to draw. These sketches were then used for querying the Bertuch-collection. Reference sets for each of those query sketches have been made by an expert. The ordering of the result is determined in accordance to domain specific qualitative features; for the pomological domain the ranking of these features is described in [14].

4.1.4. Results

The first sketch is shown on the left-hand side in Fig. 11. The reference set consists of seven objects (shown in Table 2). Fig. 12 shows the precision-recall curves for that query-sketch. The right-hand side of Fig. 12 also shows the average precision for all six sketches which are shown in

Table 2
First sketch's precision and recall; *Pos.* refers to the position in the ranked result set

Pos.	Name	Prec.	Rec.
1	Gestreifter Winter Erdbeerapfel I	1.0	0.14
2	Doppelmontagne I	1.0	0.28
5	Sommer Zuckerapfel I	0.6	0.42
6	Veilchenapfel I	0.66	0.57
7	Veilchenapfel II	0.71	0.71
8	Weisser Winterkalvill I	0.75	0.85
20	Rother Wintercalville I	0.35	1.0

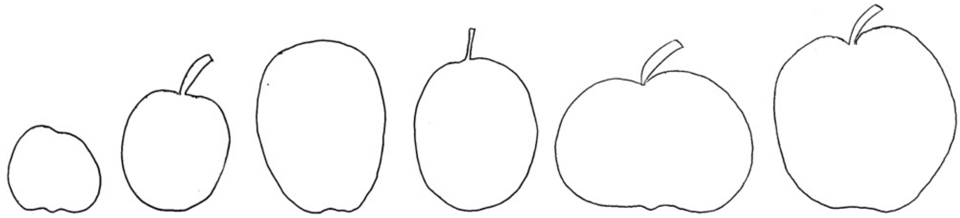

Fig. 11. Six sketched queries.

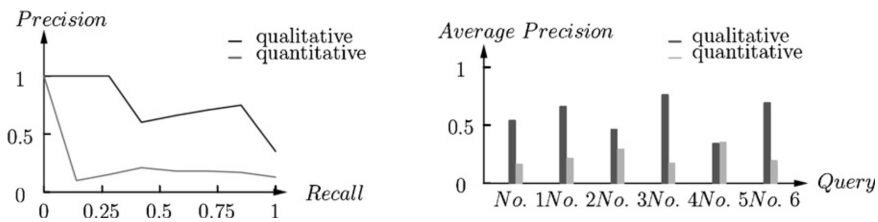

Fig. 12. (Left) The first precision-recall curve. (Right) Average precision of all requests.

Fig. 11. For each but one of the sketches the qualitative approach performs better than the quantitative one.

Since we obtain different polygons depending on how finely the fruits' reliefs are approximated, the sketches have been analysed at different granularity levels. The left-hand side of Fig. 13 shows the performance of both algorithms at a finer granularity level than Fig. 12. While there is almost no difference for the quantitative approach, the performance of the qualitative one is slightly lower, especially at higher recall rates. The right-hand side of Fig. 13 shows two further precision-recall curves for a coarser granularity level than Fig. 12. While the qualitative algorithm performs significantly better than the quantitative approach at the two finer granularity levels, it is only slightly better at the coarsest granularity level.

4.1.5. Discussion

Our results show that the qualitative algorithm generally performs much better than the quantitative one. While the performance of the quantitative one is almost independent of the recall rate, the qualitative algorithm shows a tendency to achieve higher precision at lower recall rates, with the precision decreasing slightly as the recall increases. It is obviously much simpler to sketch qualitative differences than precise quantitative distinctions.

How does our approach manage to identify the qualitative differences when dealing with the imprecise shapes of the queries? We shall have a closer look at the analysis of the stalk in order to demonstrate our method. If we zoom into the fifth sketch shown in Fig. 11, and show the stalk in close-up, we obtain the left-hand side of Fig. 14. Here we see how qualitative global features could allow us to distinguish whether a stalk is bent relative to its body or not. This can be detected by checking whether there are line segments at the tip of the stalk relative to which parts of the body are in relations such as $x_y = \mathsf{BO_r}$ and $x_z = \mathsf{FO_r}$—there would not be such overlap relations if the stalk points straight up, rather than being bent in this way.

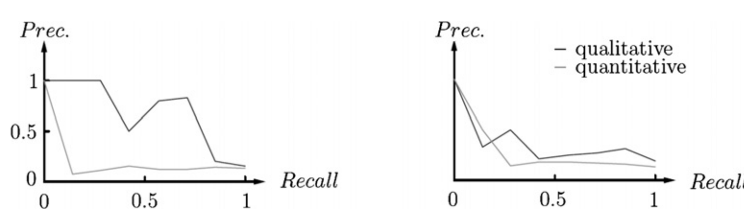

Fig. 13. Precision-recall curve at a fine granularity level (left), at a coarse one (right).

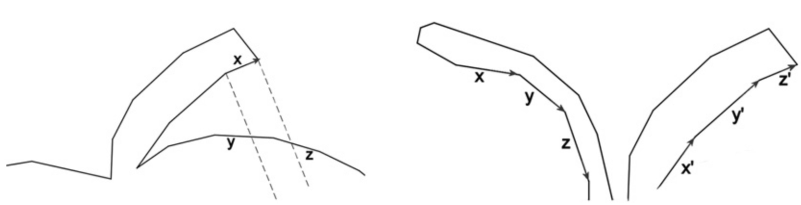

Fig. 14. (Left) A stalk and its relation to the body. (Right) Two differently bent stalks.

Another variation is shown on the right-hand side of Fig. 14. In this case, the stalks of two apples have been completely detached from their body, and we will demonstrate how their curvature can be determined without referring to the bodies of the apples. On the left is the stalk of the Italienischer Weisser Rosmarienapfel and on the right-hand side the one from the fifth sketched query. The extent of each polygon shows whether a stalk is bent or not. Consider the scopes (Definition 2) of the polygon on the left-hand side:

$$\sigma(C(x)) = [B_l F_r][F_r F_r] \wedge \eta(x) = 7$$
$$\sigma(C(y)) = [B_r F_r][F_r F_r] \wedge \eta(y) = 9$$
$$\sigma(C(z)) = [B_r F_l][F_r F_r] \wedge \eta(z) = 8.$$

By contrast, for the less bent stalk on the right-hand side it holds (for the concave line segments) that:

$$\sigma(C(x')) = [B_l F_r] \qquad \wedge \eta(x') = 7$$
$$\sigma(C(y')) = [B_l F_m][B_r B_r] \wedge \eta(y') = 7$$
$$\sigma(C(z')) = [B_r D_l][B_r B_r] \wedge \eta(z') = 5.$$

The scopes and their extent show some differences. In particular, for the stalk on the left-hand side the maximal extent is larger by two than the maximal extent of the stalk on the right-hand side. This allows different degrees in the curvature of the stalks to be distinguished.

4.2. Scenario II

The second scenario shows how a conceptual feature can be characterised within the geographic domain taking a single qualitative feature alone. In their case study [13] compared the similarity of the fourteen largest German rivers. For instance, while the Mosel (on the right-hand side of Fig. 3) has many twists and turns, the Fulda (on the left-hand side) is less curved but comprises also some windings. Does the concept of reversals as introduced in the qualitative framework account for those distinctions? In order to obtain polygons the rivers have been approximated by the polygonal approximation algorithm of Mitzias and Mertzios [20]. Using 1:5,000,000 small scale maps, the Mosel has been approximated by 51 line segments while the Fulda has been approximated by 22 line segments, allowing for the same error in both cases, namely not exceeding a cell in a raster representation of these maps. The reversals of these objects are simply computed according to Definitions 5 and 6.

The Fulda comprises 8 reversals (on average, 0.36 per line segment) and the Mosel comprises 94 reversals (on average, 1.84 reversals per line segment). This shows that the method identifies the Mosel to be five times more curved than the Fulda. Note that counting the number of reversals we do not suddenly turn to some quantitative measurement. The number of reversals is simply a quantitative measure of the incidence of some qualitative feature, namely of reversals which are still of a qualitative kind. Moreover, it is the ordering of the number of reversals of different rivers which matters, allowing the rivers to get ranked, this ordering also being of relative (and hence qualitative) nature.

In Fig. 15 there are fourteen of the largest rivers in Germany and their polygonal approximations. They have been ordered regarding their meanders, i.e. for each river the number of reversals has been determined and normalised with the number of line segments involved. The polygonal approximations have been printed outside the map in order to allow them to be compared more easily. The first river with the fewest reversals is the river Havel, that one with the most reversals is the river Mosel (last one in Fig. 15). The shown ordering can in fact be comprehended very well: beginning with the Havel the rivers getting more and more complex. It shows that both little meanders like in the case of the Mosel and larger meanders which extend larger portions of the river like in the case of the Elbe (third one, first row) and the Fulda (second one, second row) are equally identified by this method.

Looking not only for specific configurations of objects in geographic space (which would already be covered by a spatial query-by-sketch system as that one from Egenhofer [4]), the town planner in our example is also interested

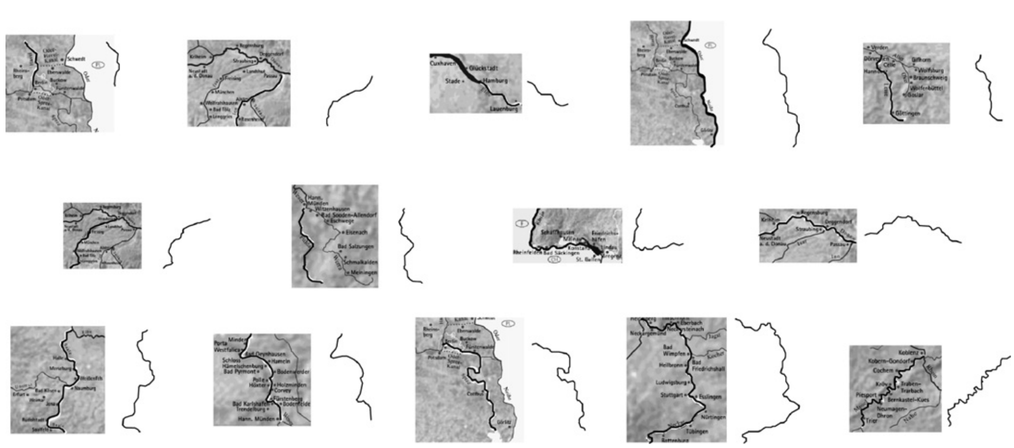

Fig. 15. German rivers and their meanderings, getting more complex from left to right and from top to down.

in what shapes look like. However, he is not looking for precise shape distinctions (as they are made by quantitative methods) but for conceptual features he will care to distinguish (as they are represented by features such as reversals). The river Neckar and the river Mosel, for example, are quite different when making precise distinctions. Taking into account a qualitative concept (like reversals), the rivers are classified as similar—in particular when comparing them with smooth rivers like the Fulda and the Isar. More important, the user of a query-by-sketch system will be able to draw quite twisty curves, somehow less twisty curves, and also almost smooth curves, but he will not care about the precise appearance of the curves; the results show that indeed the qualitative approach judges the Mosel and the Neckar to be similar with regard to this specific qualitative feature; by contrast, a quantitative approach that measures the difference between the Mosel and the Neckar would find quite a large difference. However, reversals form just one example demonstrating how qualitative features are both easily perceivable and readily sketched. Eventually, a further motivation for the approach in the geographic domain is the omnipresence of large amounts of polygonal data which require methods that can be directly applied to them.

4.3. Scenario III

The retrieval performance of our method has been compared with several other approaches in Schuldt et al. [23]. For this purpose, the well-known core experiment CE-Shape-1 for the MPEG-7 standard has been used which allows a comparison of approaches to be accomplished by taking into account only retrieval results [19]. Part B tests the similarity-based retrieval performance with a database of 1400 images: these images are semantically grouped into 70 classes of various shapes, each class containing 20 objects (some objects are shown in Fig. 16). Each image is used as a query, all other images in the database are ordered with respect to their similarity as it is measured by the approach under test. For each query the number of images which belong 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. As a consequence, the total number of correct matches for all 1400 queries is 28,000. The result of the test is the ratio of the number of found objects and the total number of correct matches. As explained by Latecki et al. [19], a retrieval rate of 100% is not possible when only using shapes, since some classes contain objects which are semantically similar but simultaneously significantly different 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 to mere chance.

Since the scope histogram offers constant time complexity for the comparison of two objects (note that the qualitative matrix as well as the histogram only have to be computed once off-line), the approach should be compared to others also having this property. This holds for the seven invariant moments proposed by Hu [16]. Applying Steger [26], these moments can directly be computed for polygons (as in our approach). Simpler are quantitative numeric features characterising polygons by a single numeric value. A prominent example is the compactness [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 [7] as well as the aspect ratio $\frac{H_r}{W_r}$ of the minimal enclosing rectangle [3]. Performing the MPEG test for those approaches leads to the results listed in Table 3: it shows that the numeric features, namely compactness, radius ratio, and aspect ratio, which characterise a shape by one single number gain results between approximately 16% and 24%. This is already significantly better than when ordering the shapes randomly (which is less than three percent correct matches). Better results can be achieved by using the seven Hu moments. Their results are at least 10 percentage points better as they retrieve about 34% of the total number of correct matches. The scope histogram outperforms all other examined approaches and retrieves about 46%. In sum, using the scope histogram (according to Definition 4) it is possible to achieve retrieval results which are about 16 times better than a random ordering. Furthermore, our approach outperforms the other examined approaches which also offer constant time complexity.

After having examined the retrieval results of the scope histogram, the question arises whether these results can be improved by combining the scope histogram with one or more of the other approaches with constant time complexity. Table 4 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%.

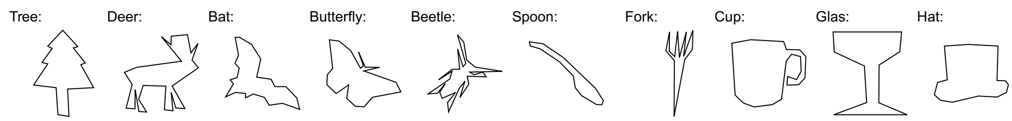

Fig. 16. Ten exemplars from 10 different categories of the MPEG-test data set.

Table 3
Classification results of compactness (CO), radius ratio (RR), aspect ratio (AR), Hu moments (HU), and scope histogram (SH) for CE-Shape-1/B

CO	RR	AR	HU	SH
21.86	16.82	24.12	34.13	45.52

Table 4
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

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 result with all features (AL) without our scope histogram (NH), shows that the scope histogram improves the results by about 10 percentage points. Eventually, it is worth mentioning that a retrieval result of about 64% is only about 12 percentage points less than the results achieved by the correspondence of visual parts of Latecki and Lakämper [18] (which is 76.45%), which has a significantly higher time complexity of $O(mn^3)$ for the comparison of two objects.

Yet another advantage of qualitative features, such as the scope histogram, is discussed in Schuldt et al. [24], 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 data set 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 5. 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.

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 s 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 GHz.

Table 5
The classification results of Table 3 and of the combination of all numeric features (NF) can be improved if a prototype is computed for each class of the MPEG test data set

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

5. Robustness

Having shown the application of the qualitative approach, the question arises as to how stable it is. This is important inasmuch deviations between polygons might exist although these polygons approximate the same object. This is due to noisy data or differently approximated outlines. In order to analyse how qualitative shape concepts behave under such circumstances we have to address a number of issues:

- The shapes are to be approximated with different degrees of precision for the purpose of showing how the features behave when applied to different granularity levels;
- The MPEG test data set is discussed in order to show that this data set is compiled so as to test distortions of non-rigid objects, changes of viewpoints, as well as other problems arising due to digitisation and segmentation noise;
- It will be discussed why the qualitative representation provides promising results even when faced with distorted and incorrect shape approximations.

5.1. Different granularity levels

Four quite different approximation levels have been tested which cover a wide range of granularities. In the most simple case each outline would be approximated by three points. This, however, means that each object would be represented by a triangle and consequently the objects could not be distinguished anymore. When taking the original outlines of the 1400 objects of the MPEG data set there are 619 contour points per object on average. It shows that about 10–15 points should be taken so that visually at least simply shaped categories can be distinguished. On the other extreme, it shows that even complex categories are approximated very well when taking into account about 50 contour points. Fig. 17 shows an example.

For the four levels of precision presented in Table 6 the complete MPEG test has been carried out in Schuldt [22]. In the first case only 2% of all contour points have been taken. On average this means that each object is approximated by only 13 points. Additionally to the scope histo-

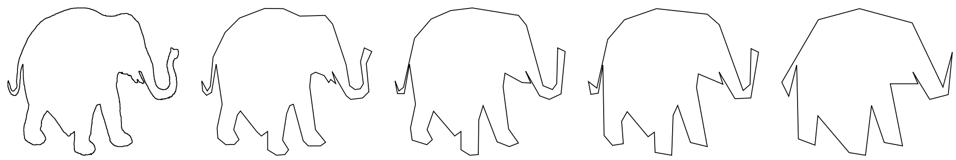

Fig. 17. An elephant approximated at different granularity levels.

Table 6
The MPEG test carried out on four different granularity levels for testing the scope histogram (SH), polygonal extent (PE), the development of $\mathcal{BA}$ relations (BA), compactness (CO), radius ratio (RR), and aspect ratio (AR)

Average number of points	Qualitative features			Quantitative features		
	SH	PE	BA	CO	RR	AR
13 (2%)	39.25	25.04	22.59	21.00	16.26	23.65
19 (3%)	43.67	24.97	22.62	21.86	16.82	24.12
29 (5%)	41.68	24.06	22.91	22.29	17.84	24.99
50 (8%)	42.45	22.55	22.19	22.52	17.76	25.54

gram (SH) the test has been carried out for the polygonal extent (PE, see Definition 3). Furthermore, for the purpose of comprehending how stable $\mathcal{BA}$ relations are the courses have been completely analysed ($\mathcal{BA}$, see Definition 1). This analysis considers all changes of $\mathcal{BA}$ relations from the point of view of each segment. These changes have been counted and averaged so as to show whether polygons similar develop around their reference segments. For this purpose each pair of segments in the qualitative matrix has been taken into account (see Section 3.1). For comparison the single numeric features are also analysed on those different granularity levels (CO, RR, and AR).

The results show that the differences between different granularity levels are rather small. This holds for the qualitative as well as the quantitative approaches, indicating that the features are robust with respect to precision. The variability of differences is only slightly higher for the qualitative approaches with the exception of BA that varies only within a range of 0.72%; while the other qualitative features vary in a range of about 2% and 3%, the quantitative approaches vary in a range of approximately 1.5% and 2%. This shows that the chosen approximation precision does only slightly influence the results. From Fig. 17 one can learn that although details get distorted on coarse granularity levels the overall relative frequency of existent $\mathcal{BA}$ relations stays similar (e.g. what is left of the elephant's trunk: two parts of the body (near the eyes and near the tail); what is right of it: nothing)—essentially the absolute number of relations gets smaller on coarse granularity levels; characteristic relations, however, remain.

5.2. The MPEG test data set

Having used the MPEG test data set a number of problems have been addressed implicitly, namely those related to changes in viewpoint, digitisation and segmentation errors, and distortions. This data set has been compiled so as to include object depictions which suffer from these problems. That is there are some outlines which are almost perfect while others have been made imperfect on purpose. We shall discuss those problems in the following by picking out examples that illustrate the complexity of the MPEG test data set.

5.2.1. Changes in viewpoint

Changing the viewpoint on an object generally entails a change in the object's outline shape. Fig. 18 demonstrates this for the fly category. Those depictions are quite different. Nevertheless, some properties keep visible even when changing the viewpoint from above to a side view: legs and feeler are always depicted as thin elongated parts while the wings are relatively large areas going somehow to the side and being closely attached to the body from which they cannot be completely separated. However, the precise shape of the wings, legs, feeler, and the body change. But still their overall arrangement stays similar when changing the viewpoint slightly, implying that positional-contrast (measured by $\mathcal{BA}$ relations) stays also similar.

5.2.2. Digitisation and segmentation errors

Fig. 19 gives some examples of digitisation and segmentation errors. They lead to small (the left of Fig. 19) or large (the right of Fig. 19) indentations in the contour; or those errors lead to a contour which becomes frayed (the middle of Fig. 19). The examples also show that exact sym-

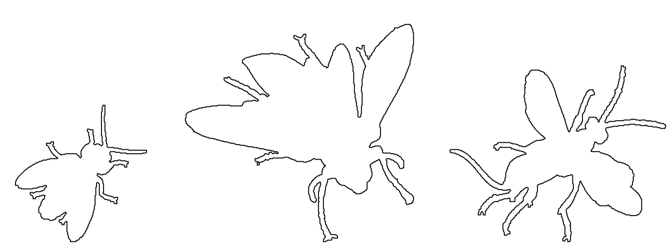

Fig. 18. Some exemplars of the fly category showing different viewpoints, i.e. from above (left), from front-above (middle), and from the side (right).

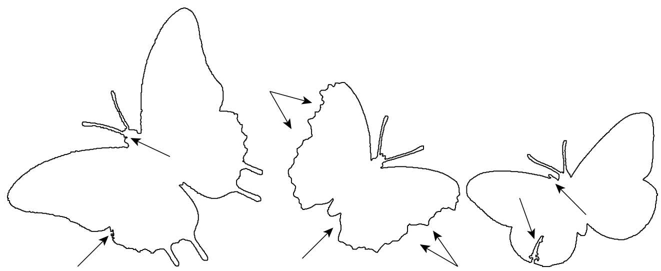

Fig. 19. Some exemplars of the butterfly category with few small errors (left), many small errors (middle), few large errors (right); the arrows point to specific faults.

metries of the butterflies get lost due to noisy data. Precise geometrical approaches suffer from those differences to the original outline since such differences sum up in quantitative similarity measures. Here again, the overall positional-contrast measured by $\mathcal{BA}$ relations stays similar.

5.2.3. Distortions of non-rigid objects

In the case of non-rigid objects such as animals the problem arises that outlines change while a running camel for example lifts its legs; and sometimes we cannot tell apart the two humps which might sometimes be close to each other while the camel walks and the viewpoint changes only a little bit (compare Fig. 20). $\mathcal{BA}$ relations change regarding those parts, just as any similarity measure would change in those cases (such as the Hausdorff distance). The frequencies of $\mathcal{BA}$ relations regarding other parts, however, remain the same.

Note that carrying out the MPEG test each of those distorted and imperfect shapes is used both as a query and as an object of the queried database.

5.3. The qualitative handling of imperfect shapes

Having shown the stability of qualitative features with regard to the MPEG test, a closer look reveals the reasons, i.e. how and why the qualitative representation manages the problems discussed in the previous sections.

The notion of positional-contrast [15] summarises what the qualitative approach does: to roughly measure how different parts of the outline relate to other parts of that same outline. Line segments are used as parts and their relative positions are measured according to the orientation grid (shown in Fig. 5), making the distinctions depicted in Fig. 4. Taking those $\mathcal{BA}$ relations in particular for line segments which are far apart those relations do not change, even if the primary segments are oriented differently; what matters is whether they lie left of or right of the reference segment. Clearly if a reference segment is oriented differently the description becomes different; but small distortions change the orientations only slightly or only for a few reference segments. Changes in orientation mainly influence the description of nearby line segments. That is the qualitative description is not that robust when relating adjacent line segments to each other ($2 * n$ out of n^2); but with respect to non-adjacent line segments the description is much more robust (and there are many more of them, namely $n^2 - 2 * n$ out of n^2).

Generally, taking the atomic $\mathcal{BA}$ relations alone the description is quite sensitive since distortions change parts of the qualitative matrix. Taking frequency distributions of those relations the description becomes more robust since the ordering of relations is abandoned. Additionally, little changes entail either no changes of $\mathcal{BA}$ relations at all (the tolerance range of single qualitative relations is quite high) or the histogram is slightly changed. This however, Section 4.3 has shown, still allows the scope histogram to get ranked higher than other approaches pertaining to the same complexity class (like compactness, radius ratio, aspect ratio, and Hu moments).

Eventually taking qualitative Gestalt features things become even more stable. Reversals for instance occur regardless of whether dealing with only a few line segments on a coarse approximation level or whether being faced with many small segments (see Section 3.4). The main problem with reversals is that they disappear if the approximation level is too coarse. But this is because the corresponding shape features, then, disappear, and consequently, other shape descriptors would fail for the same reason. Fig. 17 provides a good example which demonstrates the stability of reversals: taking a part (i.e. a line segment) as a reference there are almost always equally many reversals on all granularity levels with respect to that same part (e.g. part of a leg, of the body's back or the top of the trunk—for this purpose recall Section 3.4).

6. Qualitative similarity measures

What is the distinction between our concept of qualitative similarity measures (as applied in the previous sections) and the classical approach? Many methods have been devised most of which rely on similarity measures which are defined in some quantity space, as on the Cartesian coordinate system. Such metric-driven approaches use external reference systems which define an artificial scale relative to which objects are described [5]. By contrast, a qualitative approach allows objects to get compared directly instead of using some intermediate reference system. Opposed to metric-driven methods, we refer to the lat-

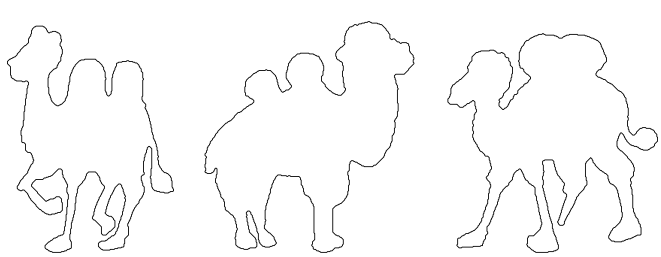

Fig. 20. Some exemplars of the camel category showing distortions due to non-rigid object motion: lifting two legs (left), front legs cannot be separated (middle), the humps cannot be separated (right).

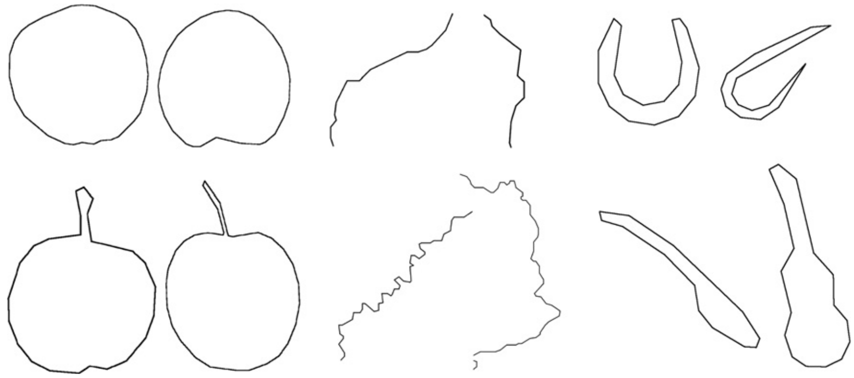

Fig. 21. $\mathcal{BA}$-similar objects grouped together: apples, rivers, and MPEG objects.

ter approach as category-driven: perceptual features which define *salient object specific spatial structures* [8] directly determine the similarity of objects at a categorical level. Their closeness to specific concepts warrant calling them conceptual features.

Given two objects which are to be compared. While a common metric-driven approach measures, for instance, the transformation necessary in order to map one object onto the other one (looking for precise differences of two apples), category-driven methods confine themselves to compare conceptual features (looking at whether both apples have a stalk). The latter approach is in particular of advantage whenever the application does not call for precise mappings (how do the stalks precisely differ), but for similarities on the conceptual level (straight versus bent stalk). In fact, conceptually similar objects might be quite dissimilar with respect to the Cartesian coordinate system. In these cases, qualitative similarity measures become even necessary.

A main difference between the metric-driven and the category-driven approach concerns precision. If precision is necessary, a metric-driven approach is required. By contrast, if conceptual properties matter (which might be quite different from the precise metrical point of view), a qualitative approach might be preferable. We summarise these observations by putting them into the following Definitions, in which the notion of *perceptual attributes* refers to distinctions that are easily obtainable by vision (without requiring sophisticated measurement tools). Some of the most prominent perceptual attributes are defined as qualitative representations in [2].

Definition 7. Qualitative representation

A qualitative representation consists of a number of jointly exhaustive and pairwise disjoint relations w.r.t. a perceptual attribute that can be defined on a metric.

$\mathcal{BA}_{23}$ relations form a number of such relations. The attribute which is encoded by $\mathcal{BA}_{23}$ concerns relative positions among line segments. This attribute is a perceptual attribute in that all relations can be reliably distinguished in vision (compare Fig. 4). The $\mathcal{BA}_{23}$ relations are jointly exhaustive in the sense that they define (on a specific level of accuracy) all possible relations among two straight line segments which are free of intersections. That they are pairwise disjoint simply means that each configuration on an arbitrary fine metric maps to exactly one (qualitative) relation. Qualitative similarity measures, then, are defined as follows:

Definition 8. Qualitative similarity measure Given a two-place function Δ which maps two objects onto a similarity value. If Δ is defined on a qualitative representation according to Definition 7, then Δ is called a qualitative similarity measure.

The apples and rivers, and also the objects of the MPEG-test are all ordered in accordance to how the underlying qualitative representation of $\mathcal{BA}$ relations rank them. As a consequence, the similarity values we are concerned with are of an ordinal kind, i.e. we simply determine how a number of objects are ranked. The domain of our similarity measure is a qualitative representation while its co-domain is of an ordinal kind. Two objects which are near each other in the list of ranked objects are perceptually more similar than those which are rather far apart. Fig. 21 shows pairs of examples for each of our domains: apples (with or without stalks and specific dents), rivers (which are either smoothly curved or twisty), and eventually object categories (which comprise some bent interior versus categories which comprise a handle and a body).

7. Discussion

In the first scenario it was sufficient and not difficult to find a number of appropriate properties to make the distinctions required, but this may well be more difficult in other domains. On the other hand, having modelled a domain by a number of such qualitative properties, according to our results, they form a robust set of features appropriate in a query-by-sketch system. This is similar in the geographic domain. However, the MPEG-test additionally shows that even the renunciation of specific domain properties leads to satisfactory results with this method. Here, only frequencies of qualitative relations have been considered.

7.1. Granularity

A crucial factor in the performance is the chosen granularity level. In the case of the apples, for example, a dent at the top might disappear at a coarse granularity level, showing why the algorithm does not perform so well at granularity levels which are so coarse that details disappear. Such details will probably also disappear for some of the objects in the collection and, as a consequence, objects which have different descriptions at finer granularity levels become more similar at coarser granularity levels and eventually get ranked equally in the result set. However, a single granularity level appears to be sufficient for all objects. We do not need to decide separately for each object how fine it should be approximated in order to achieve satisfactory results, at least in this scenario. But also regarding the many different and complex object categories in the MPEG test data set, Section 5.1 has shown that the approach performs robustly with respect to different levels of granularity.

A useful granularity level can easily be determined with regard to the application at hand. Examples are given in the geographic domain: for the purpose of spatial planning the town planner might decide on a precision which is about a meter; for the purpose of deciding difficulties a river presents for a shipping company it suffices to take into account an even lower level of precision. In this way, the choice of granularity levels is domain dependent. On the other hand, a multi-scale approach can easily be applied if necessary by looking for qualitative distinctions at several granularity levels. In this case, however, an algorithm is required that integrates the rankings which have been obtained at different granularity levels.

7.2. Complexity

Similarity measures do not go without complexity issues: how useful is a similarity measure which performs poorly in terms of run time complexity, especially when faced with large image databases?

In the first scenario we have compared two search algorithms, one of them based on qualitative features the other one on quantitative features, in terms of their precision-recall behaviour. The runtime complexity is in fact different for both approaches. Taking the number of lines involved, n, the quantitative approach needs to try n matches, from which it takes the best (i.e. closest) one. Each match involves linear time complexity, since only the angles of adjacent lines and their lengths need to be computed. As a consequence, the run time complexity of the quantitative approach is $O(n * n) = O(n^2)$.

The qualitative approach on the other hand needs only to compare a handful of characteristics, which requires constant time. But the qualitative approach also requires the computation of a polygon's course (qualitative matrix), which is $O(n^2)$ since the relative position of each line segment is considered in relation to each other line segment, increasing the runtime. However, while the quantitative approach has to be computed every time a comparison between two objects is made, the polygon's qualitative matrix and probably a number of further qualitative features only need to be computed once for each object; the qualitative features of the polygon can be stored and directly referred to whenever its similarity to another polygon needs to be determined; and this is done within constant time. Clearly, the same holds also for the other two scenarios, the runtime complexity being independent of the chosen domain.

Using qualitative features, there is another advantage which relates to the organisation of image databases. Images or (single objects) which are characterised by a number of qualitative features can be stored in a tree structure, allowing objects to be indexed in logarithmic time. This is of advantage in comparison to those methods which require to check how similar two objects are in terms of how well they can be (geometrically) mapped onto each other. However, we do not claim that qualitative representations are to be taken in every case. The question of whether to employ qualitative or quantitative features strongly depends on both the domain (which features are to be distinguished) and the application (query-by-sketch of imprecise sketches versus query-by-example of precise object depictions, for instance).

7.3. Ongoing work

A number of challenges remain which will further improve the described qualitative framework. General issues are, how robust qualitative features are with respect to different approximation algorithms. Additionally, which kinds of further qualitative features can be defined which characterise more Gestalt features of outlines than the notions of the circulation direction and reversals? This relates to the application of the method to other domains which probably require the definition of other features. At last, it is of interest how the proposed representation can be extended to allow for interior shapes and configurations of objects.

7.4. Summary

The qualitative representation introduced in this paper enables a concise characterisation of shape features. While the limitations of this approach concerns precision, qualitative features closely match perceptual distinctions since they rely on simple spatial relations which can be unambiguously determined in vision and which are therefore robustly obtainable. This is also the reason why qualitative features closely relate to category specific properties. As a consequence, they can be used as a basis on which the domain expert defines those category properties which the expert will take care to distinguish. Both the pomological scenario and the geographical scenario provide evidence for this, showing how the field of computer vision and image understanding provides promising means to improve methods in other disciplines; in this case for efficiently representing, describing, comparing, and retrieving images.

Acknowledgments

I thank Arne Schuldt for his contributions to the qualitative shape framework and for carrying out the MPEG-test. I should also like to thank the Bamberger Naturkundemuseum for making available the historical fruit collection and Mona Hess, Bamberg University, for taking the pictures of the fruits. Prof. Otthein Herzog gave me the opportunity to carry out the research. Eventually, I am grateful to the anonymous referees for a number of useful suggestions for improving this article.

References

[1] A.G. Cohn, A hierarchical representation of qualitative shape based on connection and convexity, in: A.M. Frank, W. Kuhn (Eds.), COSIT 1995, Spatial Information Theory, LNCS, Springer, Berlin, 1995, pp. 311–326.

[2] A.G. Cohn, S.M. Hazarika, Qualitative spatial representation and reasoning: an overview, Fundamenta Informaticae 43 (2001) 2–32.

[3] R.O. Duda, P.E. Hart, Pattern Classification and Scene Analysis, John Wiley and Sons, Inc., New York, NY, USA, 1973.

[4] M.J. Egenhofer, Query processing in spatial-query-by-sketch, Journal of Visual Language and Computing 8 (4) (1997) 403–424.

[5] C. Freksa, Using orientation information for qualitative spatial reasoning, in: A.M. Frank, I. Campari, U. Formentini (Eds.), Theories and Methods of Spatio-Temporal Reasoning in Geographic Space, Springer, Berlin, 1992.

[6] A. Galton, R.C. Meathrel, Qualitative outline theory, in: T. Dean (Ed.), Proceedings of 16th IJCAI-99, Morgan Kaufmann, Inc., Stockholm, Sweden, 1999, pp. 1061–1066.

[7] G.D. Garson, R.S. Biggs, Analytic Mapping and Geographic Databases, Sage Publications, Newbury Park, CA, 1992.

[8] B. Gottfried, Tripartite line tracks, in: K. Wojciechowski (Ed.), International Conference on Computer Vision and Graphics, Zakopane, Poland, 2002, pp. 288–293.

[9] B. Gottfried, Tripartite line tracks, qualitative curvature information, in: W. Kuhn, M. Worboys, S. Timpf (Eds.), Spatial Information Theory: Foundations of Geographic Information Science, International Conference COSIT 2003, LNCS (2825), Springer, Berlin, 2003, pp. 101–117.

[10] B. Gottfried, Reasoning about intervals in two dimensions, in: International Conference on Systems, Man and Cybernetics, IEEE Press, piscataway, NJ, USA, 2004, pp. 5324–5332.

[11] B. Gottfried, Singularities in qualitative reasoning, in: B. Buchberger, J. Campbell (Eds.), AISC 2004, Seventh International Conference on Artificial Intelligence and Symbolic Mathematical Computation LNAI (3249), Springer, Berlin, 2004, pp. 276–280.

[12] B. Gottfried, Global feature schemes for qualitative shape descriptions, in: H.-W. Guesgen, C. Freksa, G. Ligozat (Eds.), IJCAI-05 WS on Spatial and Temporal Reasoning, 2005.

[13] B. Gottfried, Characterising meanders qualitatively, in: M. Raubal, H.J. Miller, A.U. Frank, M.F. Goodchild (Eds.), The Fourth International Conference on GIScience, LNCS (4197), Springer, Berlin, 2006, pp. 112–127.

[14] B. Gottfried, Querying for silhouettes by qualitative feature schemes, in: DIAL'06, Second International Conference on Document Image Analysis for Libraries, IEEE Press, 2006, pp. 190–281.

[15] B. Gottfried, Shape from Positional-Contrast—Characterising Sketches with Qualitative Line Arrangements, Deutscher Universitätsverlag, Wiesbaden, 2007.

[16] M.-K. Hu, Visual pattern recognition by moment invariants, IRE Transactions on Information Theory 8 (2) (1962) 179–187.

[17] E. Jungert, Symbolic spatial reasoning on object shapes for qualitative matching, in: A.U. Frank, L. Campari (Eds.), COSIT 1993, Spatial Information Theory: A Theoretical Basis for GIS, LNCS (716), Springer, Berlin, 1993, pp. 444–462.

[18] L.J. Latecki, R. Lakämper, Shape similarity measure based on correspondence of visual parts, IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (10) (2000) 1185–1190.

[19] L.J. Latecki, R. Lakämper, U. Eckhardt, Shape descriptors for non-rigid shapes with a single closed contour, in: CVPR, Conference on Computer Vision and Pattern Recognition, 13–15 June, Hilton Head, SC, USA, IEEE Computer Society, 2000, pp. 1424–1429.

[20] D.A. Mitzias, B.G. Mertzios, Shape recognition with a neural classifier based on a fast polygon approximation technique, Pattern Recognition 27 (1994) 627–636.

[21] C. Schlieder, Qualitative shape representation, in: P. Burrough, A.M. Frank (Eds.), Geographic Objects with Indeterminate Boundaries, Taylor & Francis, London, 1996, pp. 123–140.

[22] A. Schuldt, Klassifikation von Formen durch Lage-Kontrast, Master thesis in German, Universität Bremen, 2005.

[23] A. Schuldt, B. Gottfried, O. Herzog, Retrieving shapes efficiently by a qualitative shape descriptor: the scope histogram, in: H. Sundaram, M.R. Naphade, J.R. Smith, Y. Rui (Eds.), CIVR, LNCS (4071), Springer, Berlin, 2006, pp. 261–270.

[24] A. Schuldt, B. Gottfried, O. Herzog, Towards the visualisation of shape features: the scope histogram, in: Twenty-ninth Annual German Conference on Artificial Intelligence. LNAI (4314), Springer, Berlin, 2006, pp. 289–298.

[25] J. Serra, Image Analysis and Mathematical Morphology, Academic Press, New York, 1983.

[26] C. Steger, On the Calculation of Arbitrary Moments of Polygons, Techincal Report FGBV-96-05, Informatik IX, Technische Universität München, 1996.