Pattern matching using similarity measures

Patroonvergelijking met behulp van gelijkenismaten
(met een samenvatting in het Nederlands)

PROEFSCHRIFT

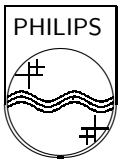
ter verkrijging van de graad van
doctor aan de Universiteit Utrecht
op gezag van Rector Magnificus, Prof. Dr. H. O. Voorma,
ingevolge het besluit van het College voor Promoties
in het openbaar te verdedigen
op maandag 18 september 2000 des morgens te 10:30 uur

door

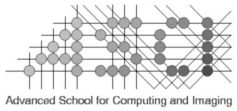
Michiel Hagedoorn
geboren op 13 juli 1972, te Renkum

promotor:	Prof. Dr. M. H. Overmars Faculteit Wiskunde & Informatica
co-promotor:	Dr. R. C. Veltkamp Faculteit Wiskunde & Informatica

ISBN 90-393-2460-3



The research described in this thesis has been made possible by financial support from Philips Research Laboratories.



The work in this thesis has been carried out in the graduate school ASCI.

Contents

Chapter 1

Introduction

1.1 Pattern matching

There are two types of pattern matching: combinatorial pattern matching and spatial pattern matching. Examples of combinatorial pattern matching are string matching [74], DNA pattern matching [26], tree pattern matching [35], and edit distance computation [34]. Spatial pattern matching is the problem of finding a match between two given intensity images or geometric models. Here, a match may be a correspondence or a geometric transformation.

The remainder of this section only considers spatial pattern matching; the term pattern matching is used as a shorthand for spatial pattern matching.

The following is a simple example of pattern matching. Given are two input patterns A and B, shown in Figures 1.1 and 1.2, respectively. The problem is finding a transformation g for which $g(A)$ is similar to B. Here, g is allowed to be a combination of scaling, rotation and translation. The output of a good pattern matching algorithm should be a transformation g like the one depicted in Figure 1.3 where $g(A)$ is superimposed on B.

In the general case, the input of a pattern matching algorithm is a pair of patterns, and the output is a transformation. Pattern matching problems can be categorised by three components: the collection of patterns, the class of transformations, and the criterion used to select a transformation. Examples of pattern collections are colour images, grey scale images, CAD models, and vector based graphics. Examples of transformation classes are translations, Euclidean isometries (which preserve the Euclidean distance between each pair of points), affine transformations (which are compositions of linear transformations and translations), elastic transformations (which do not necessarily preserve straight lines) and correspondences (which are relations between par-

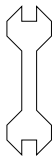


Figure 1.1: Pattern A

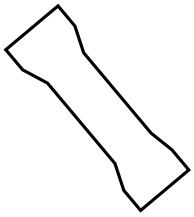


Figure 1.2: Pattern B

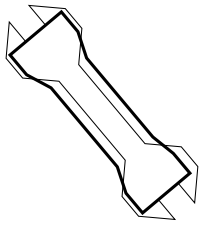


Figure 1.3: Match

ticular features in both patterns). An example of a selection criterion is that a transformation must minimise the value of a similarity measure (which is a function that assigns a nonnegative real number to each pair of patterns). The selection criterion is not always made explicit the description of a method.

Photometric and geometric pattern matching

Pattern matching can be subdivided into two categories, depending on the type of input patterns: photometric[1] pattern matching and geometric pattern matching. Photometric methods work directly on images which are considered as arrays of intensity values or real valued functions. Geometric methods work on geometric data such as finite point sets or polygons. This geometric data may be obtained directly from vector based object representations. It may also be obtained from images using feature extraction techniques. In this context, geometric pattern matching may be called feature based pattern matching. Sometimes photometric methods and geometric methods are used in combination. For example, when features are used to generate candidate matches and photometry is used to verify the candidate matches [28].

Total and partial pattern matching

There is a distinction between total pattern matching and partial pattern matching. In total pattern matching, a pattern is matched with another pattern as a whole. In partial pattern matching, a pattern is matched with part of another pattern.

Partial pattern matching can be seen as the geometric analogue of substring matching. Care is necessary in the formulation of partial pattern matching problems. For example, define partial pattern matching as transforming one

[1]Photometry stands for the intensity information measured by cameras, scanners, MRI, CT, etc.

pattern so that it becomes a subset of another pattern. Such a definition may lead to unwanted results if it is implemented, especially if transformations that include scaling are considered.

There are various ways to define partial pattern matching in terms of total pattern matching. An example is identifying the subpatterns of a pattern that correspond to distinct objects; each of these subpatterns gives rise to a total pattern matching problem. Another example is repeatedly applying total pattern matching to a pattern and the subpattern of another pattern that is contained in a window whose position varies.

This thesis deals primarily with total pattern matching. However, much of the theory and many of the algorithms discussed in this thesis can be applied to partial pattern matching problems.

Exact and approximate pattern matching

Perhaps the adjectives exact and approximate should be avoided in relation with pattern matching since they may cause confusion. However, since these terms are often used, a short discussion of them is appropriate.

Exact pattern matching usually means finding a transformation under which one pattern becomes identical to (part of) another pattern. Approximate pattern matching means finding the transformation under which a pattern becomes similar (but not necessarily identical) to another pattern. The confusion lies in the fact that the adjective approximate is also used in relation with optimisation algorithms; many types of pattern matching can be seen as optimisation problems. An approximation algorithm computes a solution that is not necessarily the optimal one. In this context, the approximation has nothing to do with the type of pattern matching. The word exact is often used to denote algorithms that compute an optimal solution. From here on, the terms exact and approximate are avoided in relation with pattern matching.

Specialised and general purpose pattern matching

Pattern matching methods are specialised in varying degrees. Some pattern matching methods are specialised to detect or recognise a restricted class of shapes such as circles and squares. Another class of highly specialised methods is found in optical character recognition (OCR). General purpose (or context independent) pattern matching methods are not specialised for any specific task and work for unrestricted classes of images or geometric shapes. These methods do not incorporate application domain dependent assumptions; they work purely on photometry or geometry.

In principle, a pattern matching algorithm that is specialised for a given application domain is more efficient and reliable than a general purpose pattern

matching algorithm, if applied to that domain. It is difficult to incorporate domain knowledge elegantly in an algorithm. Attempts to incorporate such knowledge may result in "messy" or "ad-hoc" methods. To keep the discussion clean, this thesis focuses on the subject of general purpose pattern matching. This is on itself an interesting and challenging problem. No attempts are made to specialise the algorithms for specific application domains.

Problems related to pattern matching

Object recognition is similar to pattern matching. However, there is no consensus about the exact meaning of the word object recognition. There are three distinct uses of the term object recognition.

First, the term object recognition is used for region segmentation. Region segmentation is the problem of determining the regions in an image that correspond to distinct physical objects.

Second, the term object recognition is used for a form of partial pattern matching. Here, the problem is to detect occurrences of a given object in an image.

Third, the term object recognition is used for shape recognition. Shape recognition itself can be considered in a narrow and a broad sense.

In the narrow sense, shape recognition is finding the pattern in a finite database whose shape most resembles that of a query pattern. This is known as model based object recognition. Content based image retrieval can be seen as a form of model based object recognition.

In the broad sense, shape recognition is determining what shape a pattern has. This does not necessarily involve the use of model patterns. Shape classification is a form of shape recognition in which model patterns are not always used. The shape classification problem is determining to which category a pattern belongs (where a fixed set of categories is defined).

The approaches to shape recognition vary. The methods used for shape recognition include neural network based techniques [104], statistical methods [50, 57], syntactical methods [61] and indexing methods [33, 41, 108, 117, 124, 129].

1.2 Applications

Pattern matching is a tool of major importance in many applications. To get some feeling for the range of applications, a few examples are discussed below.

Automatic telescope guidance

Cox, de Jager and Warner [43] address the task of automatically guiding a telescope so that it faces a desired star or constellation. Normally, a telescope has to be brought into position manually by an operator who references star charts. A pattern matching algorithm allows this process to be automated.

The technique matches the CCD (charge coupled device) image produced by a telescope with a star finder chart. Both the CCD image from the telescope and the star finder chart are turned into finite point sets (where the points correspond to stars). Then, a (partial) correspondence between the two point sets is found. This is done using techniques that are invariant under similarity transformations (which are compositions of translation, rotation and scaling). The resulting correspondence is used to compute a similarity transformation between the point sets. This transformation is translated into a signal to the motors of the telescope, bringing it into the correct orientation.

Depth reconstruction from images

Images of the same physical scene that are obtained from two (or more) different camera viewpoints can be used to reconstruct depth information. This task is called depth reconstruction. Depth reconstruction can be performed using the specialised form of pattern matching that is known as stereo matching [37, 49, 63, 103, 125, 126, 130]. The correspondence between the images found by a stereo matching algorithm allows the computation of the distances of the visible objects in the scene to the viewpoints.

A typical approach is the following. First, features such as corners and edges are extracted from the images. Then, a correspondence is found between the two feature sets using a stereo matching algorithm. This correspondence is then extended to a full correspondence between the pixels of both images. Finally, using the extended correspondence, the distances of the visible objects to the camera can be approximated for each pixel in each image.

Image compression

When digital images contain multiple occurrences of the same pattern, the amount of memory needed to store the image may be reduced. Occurrences of a repeated pattern can be detected by means of pattern matching [16, 18, 92]. The idea behind the compression is that each pattern is stored only once; multiple occurrences are replaced by pointers to the stored pattern.

Video compression

Digital video basically consists of a sequence of still images. However, to physically store video as a sequence of images takes a lot of space. Using the fact that successive images in the video sequence tend to be similar, the amount of memory needed to store the video can be significantly reduced without much loss in quality. This is an example of video compression. For the purpose of video compression, a pattern matching algorithm can be used to determine the similarities between the still images in a digital video recording [13]. Thus, patterns occurring in multiple images are detected and replaced with pointers to a single pattern.

Optical character recognition

Optical character recognition (OCR) is the task of automatically reading handwritten or printed text [45]. OCR is a highly specialised form of pattern matching. For example, special OCR methods exist for the recognition of Chinese characters [80]. OCR algorithms convert the scanned image of a text page into character strings. Most OCR techniques are designed specifically for character recognition and can not be generalised for recognition of more complex geometric shapes.

Vehicle tracking

Vehicle tracking is the problem of maintaining the position of a moving vehicle. This is useful in driver assistance systems (DAS). An example is the automatic prevention of collisions with vehicles on a road. Vehicle tracking can be done by means of the global positioning system (GPS) or visually by means of a camera. For the latter type of tracking, pattern matching is applied [85]. A camera delivers a sequence of images. If the position of a vehicle in an image is known, pattern matching is used to find back the vehicle in the next image. This way the position of the vehicle relative to the camera is updated through time.

Medical image registration

Image registration is the process of matching two images so that corresponding points in the two images correspond to the same physical region of the scene being imaged. Medical image registration is the alignment of medical images that are taken from different views or obtained using different input devices. Accuracy is of primary importance in registration. The images may be two or three-dimensional, or a combination of both. An important example of medical image registration is multimodality matching [75, 93, 119]. In

multimodality matching, the images are obtained using different types of input devices. For example, one image might be a CT (computed tomography) scan while the other is an MRI (magnetic resonance imaging) image. In multimodality matching there is a clear division between photometric methods and geometric methods

Satellite image registration

Another type of image registration is satellite image registration [56]. As in multimodality matching, satellite images are obtained using different types of sensors. The images may be obtained from different wavelengths and from different satellites. An example of satellite image registration is the registration of SPOT images and TM (thematic mapper) images.

Other applications

The previously discussed applications are just a few examples. There are still other applications in which pattern matching plays an important role. These include pose determination, pharmacore identification, computer aided design, robot vision, remote sensing, automatic part inspection, automatic circuit inspection and content based image retrieval.

1.3 Obtaining geometric patterns

This thesis deals with geometric pattern matching algorithms. Such algorithms work on geometric patterns. A geometric pattern is a subset of some base space (which is usually $\mathbb{R}^2$ or $\mathbb{R}^3$). From here on, the word pattern is used as an abbreviation for geometric pattern.

Each geometric pattern matching algorithm works on a specific pattern collection. Particular pattern collections frequently occur. An important example is the collection of all finite subsets of the plane. Other examples are the collection of all polylines in the plane and the collection of all polygons in the plane. For each algorithm considered in this thesis, the pattern collection consists of finite unions of geometric primitives. Examples of geometric primitives in two dimensions are line segments, curve segments and triangles. Examples of geometric primitives in general dimension are points, balls, hypersurface patches and simplices.

Patterns are obtained from two sources: images and geometric models. Images are usually stored as arrays of intensity values. A geometric model may be stored as a list of geometric primitives, a boundary representation or an hierarchical representation such as constructive solid geometry (CSG). Pattern

matching algorithms usually need simple representations as input, for example, lists of geometric primitives. For images and geometric models, different steps must be taken to obtain pattern representations that are suitable as input for a pattern matching algorithm.

To obtain patterns from geometric models, representation conversions may be necessary. If the geometric model is stored as a list of geometric primitives or a boundary representation, the conversion tends to be straightforward. However, if the geometric model is stored as an hierarchical representation, the conversion can be non-trivial.

To obtain patterns from images, feature extraction is necessary. As a preliminary step to feature extraction, image processing techniques are applied to enhance the quality of the image. For some feature extraction algorithms, the output is a binary valued (black and white) image. In these cases, an additional vectorisation step is necessary to obtain geometric primitives.

Feature extraction is a field of research on itself. It is related with disciplines such as signal processing, image processing, computer vision, and artificial intelligence. Examples of feature extraction are corner detection, edge detection and region detection.

Corner detection is the task of finding corners of physical objects in objects. Corners are points on the contour of an object at which the curvature is high.

Edge detection is the process of finding the boundaries between regions in an image that correspond to distinct physical objects. The simplest approach compute differences in intensities between neighbouring pixels. These differences are then thresholded, resulting in edges.

Region detection is the problem of determining the regions in an image that correspond to a single physical object. The most primitive region detection methods apply pixelwise thresholding on colour or intensity. More advanced techniques work on different levels of detail.

Another example of feature extraction is fitting instances of geometric primitives to an image [24, 116]. Specialised methods exist for various types of geometric primitives such as segments, curves and circles. After a collection of geometric primitives has been fit to an image, the union of these primitives can be used as an input pattern in a pattern matching algorithm.

1.4 Paradigms in geometric pattern matching

Two main paradigms can be identified in geometric pattern matching: correspondence methods and alignment methods. Correspondence methods match patterns by fitting pairs of geometric primitives. Alignment methods match patterns by fitting the unions of geometric primitives. The two paradigms are not mutually exclusive. For example, bottleneck matching, discussed in Sec-

tion 1.5, can be considered both a correspondence method and an alignment method.

Correspondence methods

Correspondence methods combine the geometric primitives that make up the input patterns. An example is matching polylines in the plane under similarity transformations by pairing line segments of both polylines. Each combination of line segments results in exactly two similarity transformations. Other examples of correspondence methods are graph matching and geometric hashing, which are discussed shortly.

The behaviour of correspondence methods depends more on the representation and the topology of patterns than on the geometry of patterns. For example, in the above example of matching two similar polylines, the resulting match can only expected to be good if the two polylines have similar distributions of vertices and edges. This behaviour is acceptable (or even desirable) in matching CAD models. However, if patterns are obtained from images by means of feature extraction, assumptions on the topology and the representation of the patterns are usually unfounded. An advantage of correspondence methods is that they may produce an explicit relation between the geometric primitives in both patterns.

Graph matching

Graph matching is a correspondence method. The structure of a pattern is described as a graph. An example of such a description is the aspect graph [59]. After the graphs have been constructed, matching is performed between the graphs.

Matches between the graphs may be found by searching for graph isomorphisms or subgraph isomorphisms. For general graphs, there are no polynomial time algorithms for both these problems. It is not known if the graph isomorphism problem is NP-complete. However, the subgraph isomorphism problem is known to be NP-complete. For special classes of graphs, more efficient algorithms exist [4, 12, 21, 118].

Due to its discrete nature, graph matching is not very robust for errors: the success of the technique depends on the correct extraction of the graphs from the input. Another limitation of graph matching is a lack of discernment: large classes of patterns share the same graph.

Geometric hashing

Geometric hashing [127, 128] is another class of correspondence methods. In geometric hashing, the geometric primitives that make up a pattern are used to generate a normalised description of the pattern as a whole. For example, for finite point sets in the plane, an affine invariant description is generated when the point set is expressed in the coordinate systems formed by each point triple. All normalised descriptions are used to build a hash table. Suppose that occurrences of a fixed pattern A as a subset of various patterns B, need to be detected efficiently. In that case, a hashing structure is built for A, and normalised descriptions of B are inserted into the hashing structure.

Geometric hashing is robust for missing points and occlusion. Geometric hashing is (time) efficient if a fixed set of patterns has to be recognised; after the data structure has been built for these patterns the technique is fast. A drawback is that this data structure can become quite large, depending on the transformation group. Geometric hashing works fine on finite point set patterns. The geometric hashing approach is very dependent on the representation and topology of patterns, which is a drawback if the input patterns are obtained from images by means of feature extraction.

Alignment methods

Alignment methods search for a match between the two input patterns as sets, ignoring how the patterns is represented. That is, no matter how two given patterns are represented in terms of geometric primitives, the outcome of an alignment method will remain the same. Examples of alignment methods are global methods and similarity measure based pattern matching, which are described shortly.

In contrast with many correspondence methods, alignment methods are independent of representation. This fact makes alignment methods more suitable for situations in which patterns are obtained from images using feature extraction.

Global methods

Global methods are alignment methods that use global geometric information to find matches. An example is mapping the bounding box of one pattern onto the bounding box of another pattern. Another example of global matching is translating the centroid of a pattern onto the centroid of another pattern. Moment based techniques [90, 36] generalise the latter example. Moment based techniques can be used to match under Euclidean isometries and similarity transformations.

Global methods can be applied most successfully if the input patterns are almost equal (up to a transformation group under consideration). Because of their simplicity, these methods are very fast. Global methods are not well suited for finding plausible matches between two patterns that differ much. In addition, global methods are not well suited for partial pattern matching.

Similarity measure based pattern matching

Similarity measure based pattern matching is the main theme of this thesis. A similarity measure is a function that assigns a nonnegative real number to each pair of patterns. A typical example of similarity measure based pattern matching is finding a geometric transformation that minimises a similarity measure. In this example, the precise problems is, given patterns A and B, finding a transformation g (from a specific class of transformations) that minimises the value of the similarity measure on $g(A)$ and B.

As long as the used similarity measure is defined on patterns as sets (instead of their representations), similarity measure based methods are independent of representation. The performance of these methods depends much on the choice of similarity measure. In contrast with global methods, similarity measure based methods can be used to find plausible matches if the patterns differ much. In addition, the similarity measure based approach can be used in partial pattern matching. On the other hand, similarity measure based pattern matching tends to be slower than global matching.

Similarity measure based pattern matching is discussed in more detail in the next section, where several examples of similarity measures are given.

1.5 Similarity measure based pattern matching

This section discusses several pattern matching methods that are (implicitly or explicitly) based on similarity measures. First, the fundamental problem of exact congruence matching is discussed. This type of pattern matching is implicitly based on the discrete metric. After that, methods that are explicitly based on similarity measures are discussed. These similarity measures include the bottleneck distance, the Fréchet distance, the Hausdorff metric, and the volume of symmetric difference.

Exact congruence matching methods

The exact congruence matching problem is the following: given two patterns A and B, find a transformation g in G under which the image $g(A)$ equals B. Typically, A and B are finite subsets of $\mathbb{R}^k$, and the group G consists of

Euclidean isometries in $\mathbb{R}^k$. The discrete metric, denoted by c, is defined as

$$\mathsf{c}(A,B) = \begin{cases} 0 & \text{if } A = B, \\ 1 & \text{otherwise.} \end{cases} \tag{1.1}$$

Exact congruence matching of A and B under G is equivalent to computing the minimum over $g \in G$ of $\mathsf{c}(g(A), B)$.

Exact congruence matching algorithms perform the following normalised steps. Both patterns are translated so that their centroids lie on the origin. The resulting set is then projected onto the unit sphere in $\mathbb{R}^k$. Each projected point is labelled with the distance of the original point to the origin. The exact congruence matching algorithms proceed by matching the labelled point sets. In two dimensions, the problem of matching the labelled point sets reduced to string matching [20, 74]. In three dimensions it is reduced to a test for graph isomorphism [21, 12, 4]. In higher dimensions, dimension reduction techniques are used [12, 5, 31].

Exact congruence matching algorithms only find matches if the patterns are equal up to transformation. Therefore, the practical use of these algorithms is limited. Exact congruence matching algorithms do not lend themselves for generalisation to other similarity measures than the discrete metric. Exact congruence matching is a fundamental problem; it can be formally reduced to the problem of minimising a metric on patterns (where any metric is allowed).

No tight lower bound for the worst case complexity of exact congruence matching is known for general dimension. This is still an open problem. Because of the connection with metric based pattern matching, a lower bound for exact congruence matching is also a lower bound for all other metric based pattern matching problems.

Bottleneck matching

The bottleneck distance between two finite point sets of equal cardinality is the minimum over all bijections between the sets over the maximum distance between each two points that are related in a bijection [54, 53, 118]. The formal definition is as follows. Let A and B be finite subsets of a space X with metric ρ. Assume that A and B have the same cardinality. Let $F(A,B)$ be the set of all bijections from A onto B. Then, the ρ based bottleneck distance b_ρ is defined as

$$\mathsf{b}_\rho(A,B) = \min_{f \in F(A,B)} \max_{a \in A} \rho(f(a), a). \tag{1.2}$$

Bottleneck matching is closely related to what is called approximate congruence matching (Alt et al. [12]).

Let X be $\mathbb{R}^2$, and let ρ be the Euclidean metric. Efrat and Itai [53] show that for point sets of cardinality n, computation of the bottleneck distance takes time $O(n^{3/2} \log n)$, and minimisation of the bottleneck distance under translation can be done in time $O(n^5 \log^2 n)$.

Bottleneck matching may be useful in applications where it is known that there is a bijection between the points of both patterns. An example is stereo matching (which is discussed in Section 1.2).

The Fréchet distance

The Fréchet distance is a similarity measure on curves which is independent of parametrisation [10, 60]. Formally, a curve is a continuous function from the interval $[0, 1]$ into a topological space X. The definition of the Fréchet distance is as follows. Let X be a space with metric ρ and let $\mathrm{Hom}([0, 1])$ be the set of homeomorphisms from $[0, 1]$ onto itself. The ρ based Fréchet distance between two curves f and g is defined as

$$\mathsf{f}_\rho(f, g) = \inf_{\alpha,\beta \in \mathrm{Hom}([0,1])} \max_{t \in [0,1]} \rho(f(\alpha(t)), g(\beta(t))). \tag{1.3}$$

The Fréchet distance is a pseudometric on each collection of curves that shares a common range X.

Alt and Godeau [10] consider the following variations on the Frechet metric. The nonmonotone Fréchet distance, is obtained by replacing $\mathrm{Hom}([0, 1])$ in Equation 1.3 with the set of all continuous functions from $[0, 1]$ onto itself. A closed curve is a continuous function from the unit circle S^1 to X. The Fréchet distance on closed curves is obtained by replacing each occurrence of $[0, 1]$ in Equation 1.3 with S^1. The partial Fréchet distance between curves f and g is the infimum of $\mathsf{f}_\rho(h, g)$ over all "subcurves" h of f.

Let X be $\mathbb{R}^2$, and let ρ be the Euclidean metric. Alt and Godeau [10] give the following results for polylines. Let n and m be the numbers of vertices in the polylines. Computation of the Fréchet distance can be done in $O(nm \log(nm))$ time. An algorithm with the same time bound exists for computation of the nonmonotone variant. The Fréchet distance for closed curves can be computed in $O(nm \log^2(nm))$ time. The partial Fréchet distance can also be computed in time $O(nm \log^2(nm))$.

Fréchet distance based matching algorithms can be used only if the input consists of curves or closed curves. This limits the practical applicability of these algorithms.

The Hausdorff metric

The Hausdorff metric is the most studied similarity measure in computational geometry. There are several equivalent definitions. For computation, the def-

inition given below is often convenient. The ρ based directed (or one sided) Hausdorff distance $\mathsf{h}_\rho^\triangleright$ on compact subsets A and B of a space X with metric ρ is given by

$$\mathsf{h}_\rho^\triangleright(A,B) = \max_{a\in A}\min_{b\in B}\rho(a,b). \tag{1.4}$$

The ρ based Hausdorff metric is expressed as the maximum of two directed Hausdorff distances by

$$\mathsf{h}_\rho = \max\{\,\mathsf{h}_\rho^\triangleright(A,B), \mathsf{h}_\rho^\triangleright(B,A)\,\}. \tag{1.5}$$

Section 2.4 discusses several alternative but equivalent definitions of the Hausdorff metric.

Most combinatorial and computational results have been found for Hausdorff metrics that are based on $X = \mathbb{R}^k$ endowed with metrics $\rho = l_p$ which are parametrised with a positive integer p. The metrics l_p are defined as follows. For each $p \geq 1$, the norm $\|\cdot\|_p$ is defined by

$$\|x\|_p = \sqrt[p]{\sum_{i=1}^{k} |x_i|^p}. \tag{1.6}$$

The special case $p = 2$, the Euclidean norm, is denoted by $\|\cdot\|$. The norm $\|\cdot\|_\infty$ is given by

$$\|x\|_\infty = \max_{i=1,\ldots,k} |x_i|\,. \tag{1.7}$$

Using these norms the metric l_p on $\mathbb{R}^k$, for each $p = 1,\ldots,\infty$, is given by

$$l_p(x,y) = \|x-y\|_p. \tag{1.8}$$

The case $p = 2$ is known as the Euclidean metric or usual metric.

Results for computation and minimisation of the Hausdorff metric have been obtained for $X = \mathbb{R}^k$ and $\rho = l_p$. Alt, Behrends and Blömer [7] show that for two finite point sets in $\mathbb{R}^2$, the l_2 based Hausdorff metric can be computed in $O((n+m)\log(n+m))$ time. The results for minimisation of the l_2 based Hausdorff metric in $\mathbb{R}^2$ are the following. Huttenlocher, Kedem and Sharir [82] describe an $O(nm(n+m)\log(nm))$ time algorithm for minimisation under translation. Chew et al. [39] give an $O((n+m)^5\log^2(nm))$ time algorithm for minimisation under Euclidean isometries. Results for several other choices of l_p and $\mathbb{R}^k$ are summarised in Section 3.3. Still other algorithms exploit input assumptions, make constant-factor approximations, or are designed to work well in practice [109, 19, 14, 6, 25, 83, 86, 84].

As stated above, the Hausdorff metric is the most studied similarity measure in computational geometry. Perhaps the reason for this is that it is defined

on a very general pattern collection: the set of all compact subsets of $\mathbb{R}^k$. For correspondence based distances such as the bottleneck distance and parametrisation based distances such as the Fréchet distance, the pattern collection is much more restricted. In the context of geometric pattern matching, a serious drawback of the Hausdorff metric is its sensitivity for noise: the value of the Hausdorff distance is determined by the "worst matching" point. Section 2.4 mentions a number of variants of the Hausdorff metric that are designed to overcome this drawback.

The volume of symmetric difference

The (volume of) symmetric difference was first considered as a similarity measure in pattern matching by Alt et al. [8, 9]. The volume of symmetric difference is defined on the collection of Lebesgue measurable subsets of $\mathbb{R}^k$. The symmetric difference of two sets A and B, denoted $A \triangle B$, is the union of the two set differences $A - B$ and $B - A$. The k-dimensional volume of a subset P of $\mathbb{R}^k$ is denoted $\mathrm{vol}\,(P)$. Now, the volume of symmetric difference is written as

$$\mathsf{s}(A, B) = \mathrm{vol}\,(A \triangle B). \tag{1.9}$$

On the collection of compact subsets of $\mathbb{R}^k$ that are equal to the closure of their interior, s is a metric. The volume of symmetric difference has a discrete analogue called the Hamming distance (also called signal distance). For two binary strings of equal length the signal distance equals the number of positions in which the strings have different values. Section 2.5 deals with mathematical properties of the volume of symmetric difference.

For polygons (or more generally, finite unions of triangles) in $\mathbb{R}^2$, the volume of symmetric difference can be computed in $O((n+m)^2)$ time (using methods discussed in Section 3.4). Under volume preserving transformations, minimising the volume of the symmetric difference is equivalent to maximising the volume of the intersection. The following algorithms maximise the volume of intersection for polygons in $\mathbb{R}^2$ De Berg et al. [46] describe an algorithm that finds the maximising translation for convex polygons in $O((n+m)\log(n+m))$ time. The same paper proves that a factor $9/25$ approximation of the maximum can be achieved by the translation that maps the centroids of the convex polygons onto each other. Mount, Silverman and Wu [99] show that a representation of the volume of the intersection as a function of translations for (possibly nonconvex) simple polygons in $\mathbb{R}^2$ can be computed in $O(nm\log(nm)+c)$ time, where c is the complexity of the representation. Section 3.4 presents new results for the volume of symmetric difference function for general dimension and for several transformation groups.

The symmetric difference is a robust similarity measure; adding small noise regions only proportionally affects its value. This makes it a suitable similar-

ity measure for matching solid patterns (regions). For the convex case, the minimisation algorithms for the volume of symmetric difference are efficient. However, the convexity assumption severely limits the use of such algorithms in pattern matching. Without the convexity assumption, the complexity of the minimisation problem increases dramatically. This will become apparent in Section 3.4.

Section 2.5 introduces a normalised version of the volume of symmetric difference. It is proven to be a metric on the collection of compact subsets of $\mathbb{R}^k$ that equal the closure of their interior. Most important, the normalised version is invariant for all transformations that preserve the ratio of volumes (which includes the group of affine transformations).

Other similarity measures

Still other similarity measures are used in pattern matching[11, 120, 121]. Examples are the Monge-Kantovoric metric [115] and similarity measures that are defined in terms of Fourier descriptors [112].

1.6 Overview of this thesis

This thesis deals with the mathematical and algorithmic aspects of similarity measures. The first aspect, the subject of Chapter 2, concerns the question how the similarity between patterns should be measured. The second aspect, the subject of Chapters 3 and 4, concerns the computation of a similarity measure, and the minimisation of a similarity measure under the action of a transformation group.

In Chapter 2, I present a new theory for the analysis of similarity measures. The focus lies on similarity measures that are pseudometrics on a collection of subsets of a space. "Big-oh" notation is used to formulate meaningful statements concerning the efficiency of algorithms. The theory of Chapter 2 can be used to formulate meaningful statements concerning the quality of similarity measures. The new theory is applied in the analysis of both known similarity measures and new similarity measures that are introduced in this thesis. The existing similarity measures include the Hausdorff metric and the volume of symmetric difference. The new similarity measures presented in this thesis include the normalised volume of symmetric difference and the reflection visibility distance (and its normalised variant).

In Section 2.1 I discuss the theory of general pseudometric spaces. The discussion includes the transformation group under which a pseudometric space is invariant, the topology of a pseudometric space, and the operations that can be applied to pseudometric spaces. In addition, I show how the minimisation

of a pseudometric under a transformation group leads to a new pseudometric that is transformation independent. I also prove that each pseudometric can be extended with a new element such that the resulting pseudometric is consistent with the original one. This simple new extension technique can be applied to extend a similarity measure so that its domain includes the empty set.

In Section 2.2 I introduce the new concept of a pseudometric pattern space. This mathematical notion has a richer structure than general pseudometric spaces. In particular, pseudometric pattern spaces allow the formalisation of various notions of robustness. Many authors acknowledge the importance of robustness properties for a similarity measure. However, as far as I know, the robustness properties of similarity measures have not been made precise until now. In Section 2.2 I present new axioms, which express four types of robustness. These forms of robustness are called deformation robustness, blur robustness, crack robustness, and noise robustness. I prove that the axioms behave nicely under the application of various standard operations on pseudometric pattern spaces. The robustness axioms are used in the analysis of similarity measures in Section 2.4, 2.5 and 2.6.

In Section 2.3 I present a new construction method which is used to define pseudometric pattern spaces. The method is based on the assignment of real valued functions to patterns. I prove that simple conditions on this assignment guarantee invariance under a given transformation group. The new similarity measures presented in Section 2.5 and 2.6 are defined using the construction method.

In Section 2.4 I analyse the Hausdorff metric rigorously. In this section several new results for the Hausdorff metric are presented. I show at which levels of generality the distinct definitions of the Hausdorff metric are equivalent. I extend a result by Matheron [94] by showing that the $\varnothing$-extended Hausdorff metric induces exactly the "myope" topology on the collection of compact subsets of any metric "base space" X. Next, I use the myope topology in providing simple proofs of deformation, blur, and crack robustness for the Hausdorff metric. I also prove that the Hausdorff metric is not noise robust for a large class of pattern collections.

In Section 2.5, an analysis of the volume of symmetric difference, the following new results appear. I show that the metric topologies of the Hausdorff metric and the volume of symmetric difference are incomparable. A normalisation of the volume of symmetric difference is introduced. In addition, the maximal sets of diffeomorphisms under which the volume of symmetric difference and its normalised version are invariant are identified. It is proven that the four types of robustness hold for the volume of symmetric difference and its normalised version.

In Section 2.6, I introduce a new similarity measure which is called the re-

flection visibility distance. The definition is based on the construction method of Section 2.3. Like the volume of symmetric difference, the reflection visibility distance has a normalised version. It is shown that the new reflection visibility distance and its normalisation are metrics on a large collection of $(k-1)$-dimensional patterns in $\mathbb{R}^k$. Furthermore, it is shown that the reflection visibility distance is invariant for volume-preserving affine transformations, while its normalisation is invariant for all affine transformations. Each of the four robustness axioms is proven for both metrics.

Section 2.7 summarises the results of the previous sections. The properties of similarity measures and their restrictions to various pattern collections are compared in a table.

Section 2.8 contains experimental results. I implemented algorithms that compute the Hausdorff metric and (normalised) reflection visibility distance In the experiments, the implementations are used to compare the behaviour of the Hausdorff metric and the normalised reflection visibility distance. The results agree with the theoretical results in Sections 2.4 and 2.6, confirming that the normalised reflection visibility distance is more robust for noise than the Hausdorff metric.

Chapter 3 studies the combinatorial and computational aspects of similarity measures. This includes the number of minima that a similarity measure can have, if seen as a function of transformations. These types of results indicate how hard pattern matching under similarity measures is, seen from the viewpoint of computational geometry. Furthermore, Chapter 3 discusses existing and new algorithms for the minimisation of similarity measures under transformations and the techniques behind such algorithms. The treatment includes the discrete metric, the Hausdorff metric, the volume of the symmetric difference, and the reflection visibility distance.

In Section 3.1, I discusses general techniques for the minimisation of similarity measures under transformation groups. It is discussed how a similarity measure as a function of transformations gives rise to an arrangement using which all minima can be found.

Section 3.2 is about exact congruence matching. First, I present a formalisation for the notion of normalisation. Then, I prove that the worst case number of congruences and symmetries under any transformation group coincide. Subsequently, I present new asymptotic bounds on the number of symmetries under rotations in arbitary dimension.

Section 2.4 discusses existing combinatorial and computational results for the Hausdorff metric.

In Section 3.4 the following new combinatorial results are presented for the volume of symmetric difference. Using constructions, in tight lower bounds are derived for the worst case complexity of the volume of intersection function

of connected polygons in two dimensions. In higher dimensions, tight bounds are given for the volume of intersection of finite unions of simplices under transformations. Furthermore, a local description of the volume of intersection as a function of affine transformations is derived. It is shown that for finite unions of axis-parallel rectangles this function is piecewise rational which is a quotient of multivariate polynomials that have bounded degrees.

Section 2.6 presents the first algorithm that computes reflection visibility distance. Here, randomised techniques are applied. The computation depends on the construction of structures which are called reflection visibility partitions. These structures are closely related to visibility partitions. Tight lower and upper bounds are given on the complexity of reflection visibility partitions.

In Chapter 4, I present new algorithms using which the infimum of the values of a similarity measure under the action of a transformation group can be computed with any desired accuracy. First, the "GBB algorithm" is presented. After that I present a theorem which states the conditions under which an instance of the GBB algorithm terminates for all its inputs. Two instances of the GBB algorithm are presented: the traces approach and the partition combination approach. The development of the methods is motivated by three shortcomings in existing methods. First, for any transformation group of dimension 3 or higher, the existing minimisation algorithms are too inefficient to be of any practical use. Second, most of the existing minimisation techniques are hard to implement. Third, the flow of control of existing minimisation algorithms relies more on the representation of the patterns as finite constructions of geometric primitives than on the geometry of the patterns as subsets of a space. That is, the flow of control of the algorithm is the same if different representations of the same pattern are used.

In Section 4.1, I start with presenting the geometric branch-and-bound (GBB) algorithm in its most general form. This minimum approximation technique is defined in terms of three fundamental operations which depend on the function that has to be minimised. A theorem is presented which states that if these operations satisfy a uniform convergence condition, the GBB algorithm always terminates. Two different applications of the GBB algorithm are presented. Each of them has a different set of fundamental operations; in both applications proofs of uniform convergence are provided.

In Section 4.2, I present the first instance of the GBB algorithm which is a new method called the traces approach. The traces approach bounds the minimum distance over a set of transformations by sweeping patterns under the transformations. An advantage of this approach is that it applies to a variety of similarity measures, including the Hausdorff metric and the volume of the symmetric difference.

In Section 4.3, I present the second instance of the GBB algorithm which

is a new method called the partition combination approach. The partition combination approach works for the normalised volume of the symmetric difference. In this technique, the fundamental operations are defined only in terms of the volumes of patterns within rectangular regions, resulting in a representation independent flow of control. The partition combination approach uses simplification trees of both patterns to control a descent to the minima in transformation space.

Section 4.4 presents experimental results obtained using my implementations of the GBB algorithm following the traces approach. This section contains results for two different similarity measures.

In the appendices, many of the mathematical prerequisites for this thesis are dealt with. Appendix A discusses point set topology. Appendix B discusses topological transformation groups.

Chapter 2

A theory of similarity measures

A similarity measure is a function that assigns a nonnegative real number to each pair of patterns, defining a notion of resemblance. Similarity measures form the basis for many pattern matching algorithms. Besides that, similarly measures are often used in shape recognition tasks such as shape classification and content based image retrieval. For each of these tasks, particular properties of similarity measures are desirable. This chapter formalises such properties. The result is a theory for similarity measures. that are defined on geometric patterns (that is, subsets of a space). This theory is applied in the analysis of existing similarity measures and the definition of new similarity measures.

This chapter mainly deals with similarity measures that are pseudometrics or metrics. A lot can be said about pseudometrics and metrics in general. This is the subject of Section 2.1. More can be said if pseudometrics are defined on a pattern collection, that is, a collection of subsets of a space. For this purpose, Section 2.2 introduces a new mathematical structure called a pseudometric pattern space. A method for constructing pseudometric pattern spaces is presented in Section 2.3. Sections 2.4, 2.5, and 2.6 apply the developed theory. Section 2.4 is a detailed analysis of the Hausdorff metric in its most general form. Section 2.5 discusses the volume of symmetric difference. Section 2.6 presents a new pseudometric which is called the reflection visibility distance. Section 2.7 compares the theoretical results of Sections 2.4, 2.5 and 2.6. Section 2.8 compares the Hausdorff metric and the (normalised) reflection visibility distance using practical experiments.

2.1 Pseudometric spaces

This section summarises important results for metrics and pseudometrics in general. First, the definitions of metric and pseudometric spaces are given. After that, the topology of pseudometric spaces is discussed. Then, it is shown how transformation independent pseudometrics can be formed using minimisation. It is also shown that each pseudometric can be turned into a metric by identifying elements with zero distance. Finally, the behaviour of pseudometric spaces under various operations is discussed.

Pseudometric spaces

Formally, a similarity measure d on a set S is a nonnegative valued function $d : S \times S \to \mathbb{R}$. The axioms given below describe properties for similarity measure d on arbitrary sets S. A subset of these axioms leads to the definition of pseudometrics and metrics.

Self-identity is the property which says that the distance between identical objects is zero. This translates to the following self-identity axiom:

Axiom 2.1.1. *For all x in S, $d(x, x) = 0$.*

Positivity is the property which says that distinct objects have a nonzero distance:

Axiom 2.1.2. *For all $x \neq y$ in S, $d(x, y) > 0$.*

Symmetry says that the order of two elements does not matter for the distance between them:

Axiom 2.1.3. *For all x and y in S, $d(x, y) = d(y, x)$.*

The *triangle inequality* says that the distance between y and z does not exceed the sum of the distance between y and x and the distance between x and z:

Axiom 2.1.4. *For all $x, y, z \in S$, $d(y, z) \leq d(y, x) + d(x, z)$.*

Fagin and Stockmeyer [55] introduce a relaxed triangle inequality which may be formalised as:

Axiom 2.1.5. *There is an $\alpha > 1$ such that for all x, y and z in S, $d(y, z) \leq \alpha \left(d(y, x) + d(x, z) \right)$.*

A stronger version of the triangle inequality is the following:

Axiom 2.1.6. *For all x, y, and z in S, $d(y, z) \leq \max\{ d(y, x), d(x, z) \}$.*

A pseudometric is a function d that satisfies Axioms 2.1.1, 2.1.3, and 2.1.4. A metric is a pseudometric satisfying Axiom 2.1.2. A metric satisfying Axiom 2.1.6 is called an ultrametric.

Still other combinations of the axioms are studied [27]. If a similarity measure only satisfies Axioms 2.1.1, 2.1.2 and 2.1.3, then it is called a semimetric. A similarity measure that satisfies only Axioms 2.1.1 and 2.1.3 is called a semipseudometric. Care is needed because some authors use the word semimetric as a synonym for pseudometric.

Copson [42] uses a simpler but equivalent definition of a metric based on the following *alternative triangle inequality* which is not equivalent to Axiom 2.1.4:

Axiom 2.1.7. *For all x, y and z in S, $d(y,z) \leq d(x,y) + d(x,z)$.*

Together with self-identity, this alternative triangle inequality implies symmetry. Thus, a similarity measure is a pseudometric if and only if it satisfies Axioms 2.1.1 and 2.1.7.

Example 2.1.1. The most trivial pseudometric is the indiscrete pseudometric, denoted by z. For each set S, and each x and y, it is simply given by $\mathsf{z}(x,y) = 0$.

Example 2.1.2. A trivial example of a metric is the discrete metric, denoted by c. For an arbitrary set S, this metric is defined as $\mathsf{c}(x,x) = 0$ and $\mathsf{c}(x,y) = 1$ for $x \neq y$ in S. The discrete metric is an example on an ultrametric.

Example 2.1.3. The Euclidean metric in $\mathbb{R}^k$, denoted by l_2, is defined as $l_2(x,y) = \|x - y\|$.

A pseudometric space (S,d) is a set S together with a designated pseudometric d on S. Similarly, (S,d) is called a metric space if d is a metric on S.

Example 2.1.4. Consider the collection of all closed k-simplices in $\mathbb{R}^k$. This is a metric space under the volume of symmetric difference $\mathsf{s}(A,B) = \mathrm{vol}\,(A \triangle B)$.

Example 2.1.5. Consider the collection of compact subsets of $\mathbb{R}^k$. Defined on this collection, the volume of symmetric difference is a pseudometric but not a metric.

Example 2.1.6. On the Euclidean plane, the distance between the first coordinates of two points, $d(x,y) = |x_1 - y_1|$, is a pseudometric but not a metric.

The topology of pseudometric spaces

A pseudometric provides the underlying set with structure. Pseudometric spaces are instances of more general spaces for which much is known. For example, each pseudometric space is also a uniform space and a topological space [87, 89]. The corresponding uniformity and topology are called the pseudometric uniformity and the pseudometric topology, respectively. Different

pseudometric spaces may have the same uniformity or topology. It is also possible that different pseudometric spaces have the same topology but different uniformities. Here, only the topological properties that are induced by pseudometrics are dealt with. In Appendix A the fundamentals of point set topology are discussed.

The topology that is induced by a pseudometric is defined as follows. In a pseudometric space (S, d), the open ball with centre x and radius ϵ, denoted by $\mathrm{B}_d(x, \epsilon)$, is the set of all y in S for which $d(x, y) < \epsilon$. The collection of all open balls is a basis for the pseudometric topology of (S, d). That is, a subset U of S is open in the pseudometric topology if and only if U can be expressed as a union of open balls.

Example 2.1.7. The indiscrete pseudometric (of Example 2.1.1) has value 0 for each two points in the space S on which it is defined. The topology that this induces is called the indiscrete topology. In this topology, S has only two open subsets, namely $\varnothing$ and S itself.

Example 2.1.8. Consider a set S endowed with the discrete metric (of Example 2.1.2). For each point x in S, there exists an open ball containing only x, for example the ball centred at x with radius 1. Therefore, in the metric topology, each subset of S is open. This is known as the discrete topology.

Example 2.1.9. The Euclidean metric (of Example 2.1.3) induces what is called the usual topology on $\mathbb{R}^k$. A set is open in this topology if and only if it can be expressed as a union of Euclidean balls. This is the same topology that is generated by the basis consisting of all open k-dimensional intervals (which are rectangles in two dimensions). Whenever the topology of $\mathbb{R}^k$ is referred to, without mention of a specific topology or metric, the usual topology is meant.

The topologies induced on a set by two different pseudometrics can be compared using the open balls. Let d and d' be pseudometrics on the set S. The topology of (S, d) is finer than the topology of (S, d') if and only if for each $x \in S$ and each $\epsilon > 0$, there exists a $\delta > 0$ such that $\mathrm{B}_d(x, \delta)$ is a subset of $\mathrm{B}_{d'}(x, \epsilon)$. If the topology of (S, d) is finer than the topology of (S, d'), then it is also said that the topology of (S, d') is coarser than the topology of (S, d). If the topology of (S, d) is both finer and coarser than the topology of (S, d'), then the topologies of (S, d) are (S, d) equal. It is possible that there is no finer-than or coarser-than relation between two topologies. In that case, it is said that the two topologies are incomparable.

Example 2.1.10. Consider a set S with a pseudometric d. Define an alternative pseudometric on S by $d'(x, y) = \min\{\, 1, d(x, y) \,\}$. Then, the topology induced by d is finer than the topology induced by d'. In addition, the topology induced by d' is finer than the topology induced by d. Thus, the pseudometric spaces (S, d) and (S, d') have the same pseudometric topology.

Finally, it is important to observe that there is the following distinction between pseudometrics and metrics: the topology induced by a metric is always Hausdorff (Appendix A), while the topology induced by a pseudometric need not be. The indiscrete pseudometric (of Example 2.1.1) shows that the topology generated by a pseudometric is not necessarily Hausdorff. Ultrametric spaces have a special topological structure [77].

Transformations in pseudometric spaces

Below, the relation between transformation groups and pseudometric spaces is studied. Appendix B discusses the fundamentals of transformation groups, including examples.

The combination of a transformation group and a pseudometric space leads to a notion called invariance. A pseudometric is invariant for a transformation if the transformation preserves all distances. More precisely, a pseudometric space (S, d) is invariant for a function g from G to itself if

$$d(g(x), g(y)) = d(x, y) \tag{2.1}$$

for each x and each y in S. A pseudometric space (S, d) is invariant for a transformation group G on S if it is invariant for each member of G. If (S, d) is invariant for g it is also said that g is an isometry for (S, d).

The orbit of an element is the set of all images of the element under the members of the transformation group. Formally, given a transformation group G for S, and an object x in S, the orbit of G passing through x, denoted by $G(x)$, is defined as

$$G(x) = \{\, g(x) \mid g \in G \,\}. \tag{2.2}$$

By the definition of a transformation group, each two orbits are either equal or disjoint. The collection of all orbits is called the orbit collection, which is denoted by S/G. The orbit collection forms a partition of S.

Given a pseudometric d on S and a transformation group G on S, the similarity measure d^G on S is defined as

$$d^G(x, y) = \inf\{\, d(g(x), y) \mid g \in G \,\}. \tag{2.3}$$

The following theorem shows that d^G is a pseudometric if d is invariant for G.

Theorem 2.1.1. *Let G be a transformation group for S and let d be a pseudometric on S. If (S, d) is invariant for G, then d^G is a pseudometric on S, and for all g and h in G,*

$$d^G(g(x), h(y)) = d^G(x, y). \tag{2.4}$$

Proof. The first claim is that d^G is a pseudometric. This involves proving Axioms 2.1.1 and 2.1.7.

Axiom 2.1.1: By definition of a transformation group, the identity transformation, denoted by e_G, is in G. This implies $d^G(x,x) \leq d(e_G(x),x) = 0$.

Axiom 2.1.7: That fact that d is a pseudometric that is invariant for G implies

$$\begin{aligned} d(kh^{-1}(y),z) &= d(h^{-1}(y),k^{-1}(z)) \\ &\leq d(x,h^{-1}(y)) + d(x,k^{-1}(z)) \\ &= d(h(x),y) + d(k(x),z). \end{aligned}$$

Axiom 2.1.7 follows from this inequality by

$$\begin{aligned} d^G(y,z) &= \inf\{\, d(g(y),z) \mid g \in G \,\} \\ &\leq \inf\{\, d(h(x),y) + d(k(x),z) \mid h,k \in G \,\} \\ &= \inf\{\, d(h(x),y) \mid h \in G \,\} + \inf\{\, d(k(x),z) \mid k \in G \,\} \\ &= d^G(x,y) + d^G(x,z). \end{aligned}$$

This finishes the proof of the first claim.

The second claim is that Equation 2.4 holds for d^G. That is, the value of d^G is independent of the representatives chosen from two orbits. Applying that G is a group, and d is invariant for G, an equivalent definition of d^G is found by

$$\begin{aligned} d^G(x,y) &= \inf\{\, d(g(x),y) \mid g \in G \,\} \\ &= \inf\{\, d(h^{-1}g(x),y) \mid g \in G, h \in G \,\} \\ &= \inf\{\, d(g(x),h(y)) \mid g \in G, h \in G \,\}. \end{aligned} \tag{2.5}$$

Substituting $G = Gg$ and $G = Gh$ in the previous definition results in Equation 2.4. This finishes the proof of the second claim. □

If the condition of the theorem is met, then $d^G(x',y')$ is called the orbit pseudometric of d and G. This pseudometric has the same distance $d^G(x',y')$ for each choice of $x' \in G(x)$ and $y' \in G(y)$. In fact, d^G can be seen as a pseudometric on the orbit collection.

Forming metrics from pseudometrics

Consider a set S with a similarity measure d. Suppose that S is partitioned into subsets of elements that are considered identical. This partition gives rise to a new similarity measure which is defined on subsets of S instead of individual elements of S. Grouping together elements with zero distance results in a

metric, if the original similarity measure is a pseudometric (see Kelley [89], for example). Below, this idea is formalised.

A partition of S is a collection of disjoint subsets whose union equals S. Let $\mathcal{S}$ be a partition of S not containing the empty set. A pseudometric d on S naturally leads to the similarity measure $d^{\mathcal{S}}$ defined on $P, Q \in \mathcal{S}$ by

$$d^{\mathcal{S}}(P, Q) = \inf\{\, d(p, q) \mid p \in P \text{ and } q \in Q \,\}. \tag{2.6}$$

It is not true that $d^{\mathcal{S}}$ is a pseudometric for each partition $\mathcal{S}$. This is demonstrated by the following example in which the triangle inequality fails to hold.

Example 2.1.11. Consider the Euclidean metric d on the real line $\mathbb{R}$. Consider the partition $\mathcal{S}$ that consists of the three sets $A = (-\infty, 0]$, $B = (0, 1)$, and $C = [1, \infty)$. Clearly, $d^{\mathcal{S}}(A, B) = 0$, $d^{\mathcal{S}}(B, C) = 0$, and $d^{\mathcal{S}}(A, C) = 1$. This means the triangle inequality does not hold for $d^{\mathcal{S}}$.

Partitions are closely related to equivalence relations. A relation on S is a subset of $S \times S$. A relation R is an equivalence relation on S if it satisfies the following axioms:

1. For each $x \in S$, $(x, x) \in R$.
2. If $(x, y) \in R$, then $(y, x) \in R$.
3. If $(x, y) \in R$ and $(y, z) \in R$, then $(x, z) \in R$.

Given an equivalence relation R, the equivalence class $[x]_R$ determined by an element $x \in S$ is the set of all $y \in S$ satisfying $(x, y) \in R$. A well known result in set theory is that the collection of all equivalence classes is a partition consisting of nonempty sets.

The identification of points with zero distance results in an equivalence relation. Given a pseudometric d on S, the corresponding equivalence relation I_d consists of all $(x, y) \in S \times S$ with $d(x, y) = 0$. The partition of S induced by I_d is denoted by $\mathcal{I}_S$. The following theorem states that if d is a pseudometric, then the similarity measure $d^{\mathcal{I}_S}$ is a metric on $\mathcal{I}_S$. In this case, $d^{\mathcal{I}_S}$ is called the quotient metric of d.

Theorem 2.1.2. *If d is a pseudometric on S, then $d^{\mathcal{I}_S}$ is a metric on $\mathcal{I}_S$.*

Proof. Axiom 2.1.1: Let $P \in \mathcal{I}_S$. Choosing an arbitrary p in P, it follows that $0 \leq d^{\mathcal{I}_S}(P, P) \leq d(p, p) = 0$.

Axiom 2.1.2: Let $P \in \mathcal{I}_S$ and $Q \in \mathcal{I}_S$ where $P \neq Q$. Select arbitrary p in P and q in Q. Because $d(p, q) > 0$, it follows from the triangle inequality of d that $d^{\mathcal{I}_S}(P, Q) = d(p, q) > 0$.

Axiom 2.1.7: Let P, Q and R be in $\mathcal{I}_S$. For p and p' in P, q in Q and r in R it follows that

$$d(p,q) + d(p,r) \leq d(p,q) + d(p',p) + d(p',r) = d(p,q) + d(p',r).$$

The substitution of this inequality in the definition of results in

$$\begin{aligned} d^{\mathcal{I}_S}(Q,R) &= \inf\{\, d(q,r) \mid q \in Q \text{ and } r \in R \,\} \\ &\leq \inf\{\, d(p,q) + d(p,r) \mid p \in P,\, q \in Q \text{ and } r \in R \,\} \\ &\leq \inf\{\, d(p,q) + d(p',r) \mid p, p' \in P,\, q \in Q \text{ and } r \in R \,\} \\ &= \inf\{\, d(p,q) \mid p \in P \text{ and } q \in Q \,\} + \inf\{\, d(p,r) \mid p \in P \text{ and } r \in R \,\} \\ &= d^{\mathcal{I}_S}(P,Q) + d^{\mathcal{I}_S}(P,R). \end{aligned}$$

□

Kelley [89, page 123] observes the following two facts for $d^{\mathcal{I}_S}$. The metric topology of $(\mathcal{I}_S, d^{\mathcal{I}_S})$ equals the quotient topology for $d^{\mathcal{I}_S}$. The mapping from each element of S to its equivalence class in $\mathcal{I}_S$.

Pseudometric spaces under operations

Below, it is investigated how simple operations affect the properties of a pseudometric space. These operations include remapping the range of a pseudometric, restriction of the underlying set to a subset, and extension of the underlying set with a single element. The results are very simple, if not trivial. However, because the results are important in the sections that follow, they are presented in the form of theorems.

The distance values of a pseudometric can be remapped without changing many of the essential properties of the pseudometric (Baddeley [22], Kaplansky [88]). Such remappings can be achieved using subadditive functions A function f is subadditive if $f(\alpha + \beta) \leq f(\alpha) + f(\beta)$. A function f is monotone if $\alpha \leq \beta$ implies $f(\alpha) \leq f(\beta)$. The following theorem states that applying a subadditive, monotone function to the range of a pseudometric results in a new pseudometric.

Theorem 2.1.3. *Let f be a subadditive, monotone function such that $f(0) = 0$. If d is a pseudometric, then $f \circ d$ is a pseudometric. If $f(\alpha) > 0$ for each $\alpha > 0$, and d is a metric, then $f \circ d$ is a metric.*

Proof. To prove the first part of the claim, it is shown that Axioms 2.1.1 and 2.1.7 hold for $f \circ d$ if d is a pseudometric.

Axiom 2.1.1: From $d(x,x) = 0$ it follows that $f \circ d(x,x) = f(0) = 0$.

Axiom 2.1.7: From the inequality $d(y,z) \leq d(x,y) + d(x,z)$, it follows by monotonicity of f that $f(d(y,z)) \leq f(d(x,y) + d(x,z))$ By subadditivity of f, the right hand of the latter inequality is smaller than $f(d(x,y)) + f(d(x,z))$. This implies Axiom 2.1.7 for $f \circ d$.

This leaves the second part of the claim. It follows because Axiom 2.1.2 holds for $f \circ d$ if it holds for d and $f(\alpha) > 0$ for each $\alpha > 0$. □

Remapping the range does not influence the essential properties of a pseudometric, such as the pseudometric topology and the invariance under transformation groups. However, the previous theorem may facilitate proving the triangle inequality for some pseudometrics. In addition, remapping the range is useful because it can be applied to transform a similarity measure that is not a pseudometric, for example a semipseudometric, into a pseudometric.

Example 2.1.12. Mappings satisfying the conditions of Theorem 2.1.3 include $f(x) = \min\{\, 1, x \,\}$, $f(x) = x/(1+x)$, and $f(x) = \sqrt[p]{(x)}$, where $p \geq 1$.

The restriction of a similarity measure d to a subset R of S, denoted by $d|_{R\times R}$, is the restriction of d to $R \times R$ as a function. Invariance is a hereditary property of pseudometric spaces. That is, invariance is preserved under restriction. This is stated by the following theorem.

Theorem 2.1.4. *Let (S,d) be a pseudometric space and let R be a subset of S. If the pseudometric space (S,d) is invariant for a transformation g from S onto itself, then $(R, d|_{R\times R})$ is also invariant for g.*

Proof. The proof is trivial: if Equation 2.1 holds for all x and y in S, then it also holds for all x and y in the subset R of S. □

Let (S,d) be a pseudometric space. Suppose S is to be extended with a new element o, resulting in a new set denoted by $S^o = S \cup \{\, o \,\}$. An extended pseudometric d^o on S^o is defined as

$$d^o(x,y) = \begin{cases} d(x,y)/(1+d(x,y)) & \text{if } x \text{ and } y \text{ are in } S, \\ 0 & \text{if both } x \text{ and } y \text{ equal } \varnothing, \\ 1 & \text{otherwise.} \end{cases} \tag{2.7}$$

If d is a metric, then d^o is also a metric. Furthermore, the restriction of d^o to $S \times S$ equals d. The following result states that invariance in (S,d) and (S^o,d^o) is the same.

Theorem 2.1.5. *A pseudometric space (S,d) is invariant for a transformation g from S onto itself if and only if (S^o,d^o) is invariant for g.*

Proof. The proof is split into two transitions

$$(S, d) \leftrightarrow (S, d^o|_{S\times S}) \leftrightarrow (S^o, d^o).$$

The values of d and $d^o|_{S\times S}$ are related by the homeomorphism $\alpha \mapsto \alpha/(1+\alpha)$ from $[0, \infty)$ onto $[0, 1)$. This implies that the invariance group, and satisfaction of the axioms are unchanged in the first transition.

This leaves proving equivalence in the second transition. This transition adds a number of constraints that involve the new element o. It is sufficient to prove that these constraints hold by definition.

Recall that $d^o|_{S\times S}$ is invariant for g if and only if it satisfies Equation 2.1 for all x and y in S: Extension with the new element o adds three new instances of this equation. Each of these equations is shown to be true using the definition of d^o:

1. If $x = o$ and $y = o$, then $d^o(g(o), g(o)) = d^o(o, o)$.

2. If $x = o$ and $y \in S$, then $d^o(g(o), g(y)) = 1 = d^o(o, y)$.

3. The case $x \in S$ and $y = o$ is analogous to the previous case.

□

In the next section, the extension method is used to include the empty set in the domain of a similarity measure that is defined on a collection of subsets of a space.

2.2 Pseudometric pattern spaces

This section presents new theory for the special case of pseudometrics that are defined on collections of subsets of some base space. If the base space is taken into consideration, more properties of pseudometrics can be studied than for pseudometric spaces in general. This leads to the definition of pseudometric pattern spaces and metric pattern spaces, which are discussed at the beginning of this section. After this definition, it is observed that the notion of invariance has an interesting interpretation for pseudometric pattern spaces. Subsequently, axioms that reflect useful properties of pseudometric pattern spaces are presented. In addition, it is shown that the axioms behave nicely if particular operations are applied to a pseudometric pattern space. Finally, if the base space is Euclidean, some concepts can be simplified.

Pseudometric pattern spaces

The definition below presents a special type of pseudometric space in which the elements are subsets of a given space.

Definition 2.2.1. A *pseudometric pattern space* is a structure $(X, \mathcal{P}, d)$, where X is a topological space, $\mathcal{P}$ is a collection of subsets of X, and d is a pseudometric. If d is a metric, then $(X, \mathcal{P}, d)$ is called a *metric pattern space*.

The collection $\mathcal{P}$ is called the pattern collection. The elements of $\mathcal{P}$ are called patterns. The set X is called the base space. If X has a metric defined on it, then it is called the base metric. In the general case, the base metric is denoted by the symbol ρ.

Example 2.2.1. Let $\mathcal{P}$ be the collection of compact subsets of $\mathbb{R}^k$. A pseudometric s is given by $\mathsf{s}(A, B) = \mathrm{vol}\,(A \triangle B)$. The combined structure $(\mathbb{R}^k, \mathcal{P}, d)$ is a pseudometric pattern space.

Example 2.2.2. Let (X, ρ) be some metric space, and let $\mathcal{P}$ be the collection of nonempty closed bounded subsets of X. Let $\mathrm{N}_\rho(P, \epsilon)$ be the union of all open balls with radius ϵ centred at points of P (where balls are defined relative to the metric ρ). The Hausdorff metric h_ρ on $\mathcal{P}$ with base metric ρ is given by

$$\mathsf{h}_\rho(A, B) = \inf\{\, \epsilon > 0 \mid \ A \subseteq \mathrm{N}_\rho(B, \epsilon) \text{ and } B \subseteq \mathrm{N}_\rho(A, \epsilon) \ \}.$$

The structure, $(X, \mathcal{P}, \mathsf{h}_\rho)$ is a metric pattern space.

Transformations in pseudometric pattern spaces

A topological base space X has a "maximal" topological transformation group associated with it, namely, the class of homeomorphisms $\mathrm{Hom}(X)$. The group $\mathrm{Hom}(X)$ can also be seen as a transformation group on the collection of subsets of X. In this case, the result of applying a transformation t in $\mathrm{Hom}(X)$ on a subset P of X, denoted by $t(P)$, is defined as the image set $\{\, t(p) \mid p \in P \,\}$. This is how transformation groups on the base space X lead to transformation groups on the pattern collection in $\mathcal{P}$.

A set of transformations for X is only a transformation group for $\mathcal{P}$ if it maps elements of $\mathcal{P}$ to elements of $\mathcal{P}$. The collection of patterns $\mathcal{P}$ uniquely determines a maximal subgroup $\mathrm{Clos}(X, \mathcal{P})$ of $\mathrm{Hom}(X)$, called the closure group, under which $\mathcal{P}$ is closed. It consists of all t in $\mathrm{Hom}(X)$ such that for each $P \in \mathcal{P}$ both the image $t(P)$ and the inverse image $t^{-1}(P)$ are elements of $\mathcal{P}$. Throughout, it is assumed that $\mathrm{Clos}(X, \mathcal{P})$ has the relative compact-open topology (defined in Appendix A).

Section 2.1 defines orbits and orbit sets for general pseudometric spaces. A pseudometric pattern space is a special type of pseudometric space; consequently orbits and orbit sets have special interpretations. With respect to a

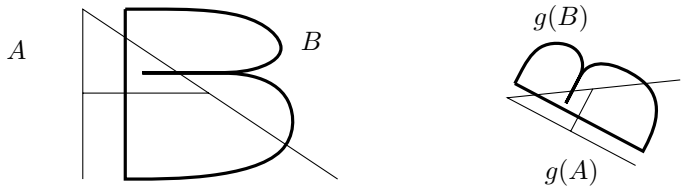


Figure 2.1: Invariance

given transformation group, an orbit can be interpreted as the shape of a pattern, and the orbit set can be seen as a class of shapes. This is a generalisation of the notion of shape used by Small [114], who defines shape as the orbit set of a pattern under the group of similarity transformations.

Example 2.2.3. Under the group of affine transformations, the set of all triangles, and the set of all ellipses are shapes.

Example 2.2.4. For projective transformations, the set of all quadrangles is a shape.

Figure 2.1 shows two planar patterns A and B, and their images $g(A)$ and $g(B)$ under an affine transformation g. The invariance of a similarity measure under affine transformations, implies that the distance between the two patterns on the left equals the distance between the two transformed patterns on the right. In the context of pattern matching and shape recognition, affine invariance of a similarity measure is a useful property. It means that the similarity is measured independent of coordinate system.

Under the condition of Theorem 2.1.1, the pseudometric resulting from Equation 2.3 can be seen as a pseudometric on shapes. If the invariance condition of Theorem 2.1.1 is not satisfied, the resulting similarity measure may not behave "nicely".

Example 2.2.5. Consider the Hausdorff metric for the base space $\mathbb{R}^k$ that is equipped with the Euclidean metric. This is the Euclidean Hausdorff metric, denoted by h. This metric is not invariant for the group of affine transformations Af^k. A simple example shows that $\mathsf{h}^{\mathrm{Af}^k}$ (defined as in Equation 2.3) is not symmetric. Consider a subset A of $\mathbb{R}^k$ consisting of two distinct points with distance 1 in $\mathbb{R}^k$, and a subset B of $\mathbb{R}^k$ consisting of a single point. Now $\mathsf{h}^{\mathrm{Af}^k}(A, B) = 0$ while $\mathsf{h}^{\mathrm{Af}^k}(B, A) = 1/2$.

Axioms for pseudometric pattern spaces

Below, robustness axioms are defined for pseudometric pattern spaces. The axioms express the following principle: if the "difference" between two pat-

terns is sufficiently "limited", then the distance to the original pattern will be "small". Each of the axioms corresponds to a distinct type of difference. Faulty feature extraction or imprecise geometric modelling may result in four types of differences.

First, there is an amount of "deformation" in a pattern, for example, as the result of measure errors and discretisation errors.

Second, patterns obtained by feature detection may contain "blur". For edge detection methods, blur may occur in the form of double edges. For region detection methods, blur may occur in the form of spots near the boundary of an pattern.

Third, there may be a "crack" between parts of pattern that correspond to a single object. This may be the result of edge detection or region detection. Alternatively, this type of error may occur in boundary representations.

Fourth, a pattern obtained by feature detection may contain "noise" in the form of outliers.

The four types of differences result in separate axioms, corresponding to deformation, blur, crack, and noise. Each of the axioms has a form that is analogous to the definition of continuity: "for each $\epsilon > 0$, the differences can be limited so that each resulting pattern has a distance no more than ϵ from the original pattern." Saying that the distance of B to A is at most ϵ under a pseudometric d is equivalent to saying that B is contained in the open ball centred at A with radius ϵ, denoted by $\mathrm{B}_d(A, \epsilon)$.

First, the deformation robustness axiom is defined. Deformations are chosen to be elements of the closure group $\mathrm{Clos}(X, \mathcal{P})$. The deformation robustness axiom states that if a pattern is transformed by a deformation "sufficiently close" to the identity transformation, then the image will have an "arbitrarily small" distance to the original pattern. Figure 2.2 shows a two-dimensional pattern A and its image under a transformation t that is "close" to the identity transformation. The notion of closeness between deformations is defined by giving $\mathrm{Clos}(X, \mathcal{P})$ the compact-open topology (defined in Appendix A). The pseudometric pattern space $(X, \mathcal{P}, d)$ is called *deformation robust* if it satisfies the following axiom.

Axiom 2.2.1. *For each A in $\mathcal{P}$ and each $\epsilon > 0$, an open neighbourhood I of the identity in $\mathrm{Clos}(X, \mathcal{P})$ exists such that $t(A) \in \mathrm{B}_d(A, \epsilon)$ for each t in I.*

An equivalent formulation of deformation robustness is that for each pattern $P \in \mathcal{P}$, the function $t \mapsto t(P)$ from $\mathrm{Clos}(X, \mathcal{P})$ to $\mathcal{P}$ is continuous.

The blur robustness axiom states that adding new pieces of boundary of a pattern "sufficiently close" to the original boundary results in a pattern that has an "arbitrarily small" distance to the original pattern. Closeness to the boundary is expressed using open neighbourhoods U of the boundary. The fact that the differences between patterns A and B are contained within a set U

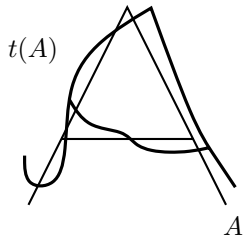


Figure 2.2: Deformation

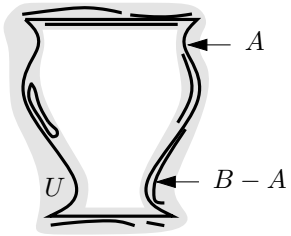


Figure 2.3: Blur

is expressed using the equation $A - U = B - U$. In words, this equation says that each point of A that lies outside of U is also a point of B and vice versa. Figure 2.3 shows a neighbourhood U of $\mathrm{Bd}\,(A)$ in which parts of $\mathrm{Bd}\,(B)$ occur that are not in $\mathrm{Bd}\,(A)$. A pseudometric pattern space $(X, \mathcal{P}, d)$ is called *blur robust* if it satisfies the following axiom.

Axiom 2.2.2. *For each A in $\mathcal{P}$ and each $\epsilon > 0$, an open neighbourhood U of $\mathrm{Bd}\,(A)$ exists such that $B \in \mathrm{B}_d(A, \epsilon)$ for each B in $\mathcal{P}$ satisfying $B - U = A - U$ and $\mathrm{Bd}\,(B) \supseteq \mathrm{Bd}\,(A)$.*

The crack robustness axiom says that changes in a pattern that are "sufficiently close" to a "crack" result in a pattern with an "arbitrarily small" distance to the original pattern. The crack robustness axiom uses the following definition of a crack.

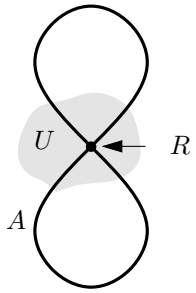


Figure 2.4: Crack

Definition 2.2.2. Let A be a subset of a topological space X. A subset R of $\mathrm{Bd}\,(A)$ is called a *crack* if R is a subset of $\mathrm{Cl}\,(\mathrm{Bd}\,(A) - R)$ and and R is homeomorphic to a closed ball in some finite-dimensional Euclidean space.

Example 2.2.6. If x is a point on a line $L \subseteq \mathbb{R}^2$. then the singleton x is a crack of L (which is homeomorphic to a closed ball in $\mathbb{R}^0$).

Example 2.2.7. If L is a line segment lying in a plane $P \subseteq \mathbb{R}^3$, then L is a crack of P (which is homeomorphic to a closed ball in $\mathbb{R}^1$).

Example 2.2.8. For the unit cube $[0,1]^3$ lying in $\mathbb{R}^3$, the set $[0,1] \times [0,1] \times \{\, 1/2 \,\}$ is a crack (which is homeomorphic to a closed ball in $\mathbb{R}^2$).

A pseudometric pattern space $(X, \mathcal{P}, d)$ is called *crack robust* if it satisfies the following axiom.

Axiom 2.2.3. *For each A in $\mathcal{P}$, each crack R of A, and each $\epsilon > 0$, an open neighbourhood U of R exists such that $B \in \mathrm{B}_d(A, \epsilon)$ for each B in $\mathcal{P}$ satisfying $B - U = A - U$.*

Figure 2.4 shows a pretzel, consisting of two topological 1-spheres glued together at a point x. The singleton set $R = \{\, x \,\}$ is a crack of A. In this example, crack robustness of a metric d means that the pretzel A may be cut into two pieces by modifying it in a small neighbourhood of x, resulting in a topologically different pattern which is close to A under the metric d.

The noise robustness axiom says that for each point in the base space, applying changes in a "sufficiently small" open neighbourhood of that point results in a pattern with an "arbitrarily small" distance to the original pattern. The definition is much like that of crack robustness. A pseudometric pattern space $(X, \mathcal{P}, d)$ is called *noise robust* if it satisfies the following axiom.

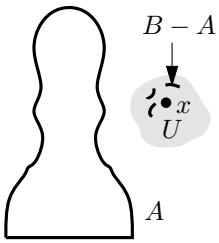


Figure 2.5: Noise

Axiom 2.2.4. *For each A in $\mathcal{P}$, each x in X, and each $\epsilon > 0$, an open neighbourhood U of x exists such that $B \in \mathrm{B}_d(A, \epsilon)$ for each B in $\mathcal{P}$ satisfying $B - U = A - U$.*

Figure 2.5 shows a pattern A and a point x. Addition of noise $B - A$ within a neighbourhood U of x results in a new pattern B. Axiom 2.2.4 says that U can be chosen so that the distance between A and B becomes as small as desired.

The following generalisation of a pseudometric pattern space is convenient in robustness proofs. A topological pattern space may be defined as a triple $(X, \mathcal{P}, \mathfrak{T})$, where X is a topological space, $\mathcal{P}$ is a collection of subsets of X, and $\mathfrak{T}$ is a topology on $\mathcal{P}$. Axioms 2.2.1, 2.2.2, 2.2.3, and 2.2.4 can be generalised for topological pattern spaces. The generalisation consists of replacing the open ball $\mathrm{B}_d(A, \epsilon)$ with an open neighbourhood $\mathcal{N}$ in the definition of Axioms 2.2.1, 2.2.2, 2.2.3, and 2.2.4. This means, the phrase "for each $\epsilon > 0$" becomes "for each open neighbourhood $\mathcal{N}$ of A in $\mathcal{P}$". If the topology $\mathfrak{T}$ has a basis $\mathfrak{B}$, the open neighbourhoods may be restricted to be members of $\mathfrak{B}$. In Section 2.4 these observations are applied to prove deformation, blur and crack robustness for the Hausdorff metric.

If the base space X is the Euclidean space $\mathbb{R}^k$, some properties of pseudometric pattern spaces can be reformulated in more convenient ways. This holds especially for the axioms.

The following formulation is a specialisation of deformation robustness (Axiom 2.2.1) for each pseudometric pattern space that has a Euclidean base space.

Axiom 2.2.5. *For each P in $\mathcal{P}$ and each $\epsilon > 0$, a compact subset K of $\mathbb{R}^k$ and a $\delta > 0$ exist such that $\max_{x \in K} \|t(x) - x\| < \delta$ implies $t(P) \subseteq \mathrm{B}_d(P, \epsilon)$ for all t in $\mathrm{Clos}(\mathbb{R}^k, \mathcal{P})$.*

This expression of the axiom is equivalent to saying that for each P in $\mathcal{P}$ the function $t \mapsto t(P)$ from $\mathrm{Clos}(\mathbb{R}^k, \mathcal{P})$ to $\mathcal{P}$ is continuous, assuming $\mathrm{Clos}(\mathbb{R}^k, \mathcal{P})$ has the topology of compact convergence (defined in Appendix A). Since for $\mathbb{R}^k$ the latter topology is equivalent to the compact-open topology, Axiom 2.2.5 is equivalent with Axiom 2.2.1 for Euclidean base spaces.

Pseudometric pattern spaces under operations

Below, the relations between pseudometric pattern spaces are studied. The focus lies on pseudometric pattern spaces that are related by standard operations such as remapping the range of a pseudometric, taking the infimum over a transformation group, restriction of the pattern collection, extension of the pattern collection with the empty set, and complementation of the pattern collection.

Consider the operation of remapping the range of the pseudometric of a pseudometric pattern space in the sense of Theorem 2.1.3. It was already shown in Section 2.1 that the invariance under transformations does not changed. In addition, satisfaction of the axioms is unchanged under a remapping of the range.

The following theorem states that if a robustness axiom holds for a pseudometric pattern space, then the axiom also holds for each pseudometric pattern space whose topology is coarser.

Theorem 2.2.1. *Let X be a base space with a pattern collection $\mathcal{P}$. Suppose d and d' are pseudometrics on $\mathcal{P}$. If the topology on $\mathcal{P}$ induced by d is finer than that induced by d' and $(X, \mathcal{P}, d)$ satisfies one of Axioms 2.2.1, 2.2.2, 2.2.3 and 2.2.4, then the same axiom also holds for $(X, \mathcal{P}, d')$.*

Proof. Each of Axioms 2.2.1, 2.2.2, 2.2.3 and 2.2.4 has the following form: for each pattern A in $\mathcal{P}$, each object O and each $\delta > 0$, there is a some subcollection $\mathcal{B}$ of $\mathcal{P}$ such that $A \in \mathcal{B} \subseteq \mathrm{B}_d(A, \delta)$. The finer-than relation between the pseudometric topologies of d and d' implies that for each $A \in \mathcal{P}$ and each $\epsilon > 0$, there is a $\delta(\epsilon) > 0$ such that $\mathrm{B}_d(A, \delta(\epsilon)) \subseteq \mathrm{B}_{d'}(A, \epsilon)$. Combining this with the axioms for d gives the axioms for d'. □

The next theorem states that robustness of a pseudometric d implies robustness for the pseudometric d^G defined under the conditions of Theorem 2.1.1.

Theorem 2.2.2. *Let $(X, \mathcal{P}, d)$ be a pseudometric pattern space and let G be a subgroup of $\mathrm{Clos}(X, \mathcal{P})$. If $(X, \mathcal{P}, d)$ satisfies one of Axioms 2.2.1, 2.2.2, 2.2.3 and 2.2.4, then $(X, \mathcal{P}, d^G)$ satisfies same the axiom.*

Proof. Observe that $d^G(A, B) \leq d(A, B)$ for all $A, B \in \mathcal{P}$. The result now follows directly from Theorem 2.2.1. □

All important properties of a pseudometric pattern space are preserved if the collection $\mathcal{P}$ is restricted. In other words, these properties are hereditary. If $\mathcal{S}$ is a subcollection of $\mathcal{P}$, the restriction of $(X, \mathcal{P}, d)$ to $\mathcal{S}$ is the pseudometric pattern space $(X, \mathcal{S}, d|_{S\times S})$. The next theorem claims that properties that hold for the original pseudometric pattern space are implied for the restricted structure.

Theorem 2.2.3. *If a pseudometric pattern space $(X, \mathcal{P}, d)$ satisfies one of Axioms 2.2.1, 2.2.2, 2.2.3 and 2.2.4, then $(X, \mathcal{S}, d|_{S\times S})$ satisfies the same axiom.*

Proof. Each axiom can be written as a predicate starting with a list of quantifiers without negations before them or in between them. Furthermore, each quantification over $\mathcal{P}$ is universal ("for all"). Therefore each result holding for $(X, \mathcal{P}, d)$ continues to hold if $\mathcal{P}$ is replaced by $\mathcal{S}$. □

Throughout the remainder of the book, the notation for the restricted pseudometric is omitted: if $(X, \mathcal{P}, d)$ is pseudometric pattern space and $\mathcal{S}$ is a subcollection of $\mathcal{P}$, then $(X, \mathcal{S}, d)$ is used as a shorthand for $(X, \mathcal{S}, d|_{\mathcal{S}\times\mathcal{S}})$.

The collection of patterns $\mathcal{P}$ corresponding to a pseudometric pattern space $(X, \mathcal{P}, d)$ may not include the empty set. An example is the Hausdorff metric, which is defined on the nonempty bounded closed subsets of some metric base space X. Using Equation 2.7, it is possible to construct a new pseudometric pattern space $(X, \mathcal{P}^{\varnothing}, d^{\varnothing})$ that includes the empty set and is in many ways consistent with the old one. For example, by Theorem 2.1.5, the invariance group is the same for $(\mathcal{P}, d)$ and $(\mathcal{P}^{\varnothing}, d^{\varnothing})$. The following theorem states that the extended structure is consistent with the original structure, except for noise robustness (Axiom 2.2.4).

Theorem 2.2.4. *A pseudometric pattern space $(X, \mathcal{P}, d)$ satisfies one of Axioms 2.2.1 and 2.2.2 if and only if $(X, \mathcal{P}^{\varnothing}, d^{\varnothing})$ satisfies the same axiom. If the base space X is Hausdorff, then the same holds for Axiom 2.2.3.*

Proof. The proof can be split into two transitions

$$(X, \mathcal{P}, d) \leftrightarrow (X, \mathcal{P}, d^{\varnothing}|_{\mathcal{P}\times\mathcal{P}}) \leftrightarrow (X, \mathcal{P}^{\varnothing}, d^{\varnothing}).$$

The values of d and $d^{\varnothing}|_{\mathcal{P}\times\mathcal{P}}$ are related by the homeomorphism $\alpha \mapsto \alpha/(1+\alpha)$ from $[0, \infty)$ onto $[0, 1)$. Observe that this implies that satisfaction of the axioms is unchanged in the first transition.

This leaves proving equivalence in the second transition. This transition adds a number of constraints that involve the empty set. It is sufficient to prove that these constraints hold by definition. The proof proceeds by showing that satisfaction of Axioms 2.2.1, 2.2.2, and 2.2.3 separately is unchanged under the extension.

Axiom 2.2.1 quantifies over all A in $\mathcal{P}$. The extension of $\mathcal{P}$ to $\mathcal{P}^{\varnothing}$ results in the additional case $A = \varnothing$. This case is true because $t(\varnothing) = \varnothing$ for all $t \in \mathrm{Clos}(X, \mathcal{P})$.

Axiom 2.2.2 has two quantifications over $\mathcal{P}$. This means that the extension to $\mathcal{P}^{\varnothing}$ is slightly more complicated. The restriction of Axiom 2.2.2 to patterns A with empty boundary is always true for all pseudometric pattern spaces: choose $U = \varnothing$. This means that a possible disagreement must occur when the boundary of A is nonempty. Now the extra case introduced by setting $B \in \mathcal{P}$ is always true because $\mathrm{Bd}\,(B) \not\supseteq \mathrm{Bd}\,(A)$.

Axiom 2.2.3 resembles Axiom 2.2.2. It is safe to consider only patterns A with nonempty boundaries and cracks R of A that are nonempty. Under these circumstances Axiom 2.2.3 for $(X, \mathcal{P}^{\varnothing}, d^{\varnothing})$ implies Axiom 2.2.3 for $(X, \mathcal{P}, d^{\varnothing}|_{\mathcal{P}\times\mathcal{P}})$.

The implication in the other direction is found as follows. Choose A in $\mathcal{P}$, $\epsilon > 0$ and a crack R of A, where $\mathrm{Bd}\,(A)$ and R are nonempty. Choose an open neighbourhood U of R such that $d^{\varnothing}|_{\mathcal{P}\times\mathcal{P}}(A, B) < \epsilon$ for each $B \in \mathcal{P}$ satisfying $B - U = A - U$. By definition of a crack, R is in the closure of $\mathrm{Bd}\,(A) - R$. Because U is an open neighbourhood of R, it intersects $\mathrm{Bd}\,(A)$ in some point $x \notin R$. Using the assumption that X is Hausdorff (a property defined in Appendix A), choose disjoint open neighbourhoods of R and x; let V be the open neighbourhood corresponding to R. Now $W = U \cap V$ is open neighbourhood of R such that $B - W = A - W$ implies $B \neq \varnothing$. A a consequence, $d^{\varnothing}(A, B) < \epsilon$ for each $B \in \mathcal{P}^{\varnothing}$ satisfying $B - W = A - W$. □

Thus, proving one of Axioms 2.2.1 and 2.2.2 for the original patterns (without the empty set), means proving the same axiom for the extended set of patterns (including the empty set). For Axiom 2.2.3 this is also true, under the assumption that X is Hausdorff. In general, noise robustness (Axiom 2.2.4) does not carry over under the extension. In fact, each collection of patterns $\mathcal{P}$ that contains a singleton constitutes a counterexample. Observe that the above results also hold if the extension is with X instead of $\varnothing$.

Consider complementation, that is, replacing each pattern of $\mathcal{P}$ with its complement in X. The principle here is that "exchanging black and white must not matter for shape". This is made more precise as follows. Suppose that $(X, \mathcal{P}, d)$ is a pseudometric pattern space. A new collection of patterns, denoted by $\mathcal{P}^c$, is formed if each pattern A in $\mathcal{P}$ is replaced with $X - A$. On this collection, define a pseudometric d^c by setting

$$d^c(X - A, X - B) = d(A, B). \tag{2.8}$$

This results in the complement pseudometric pattern space $(X, \mathcal{P}^c, d^c)$. The following theorem states that a pseudometric pattern space is equivalent to its complement with respect to invariance and satisfaction of the axioms.

Theorem 2.2.5. *Let $(X, \mathcal{P}, d)$ be a pseudometric pattern space.*

1. *The pseudometric space $(\mathcal{P}, d)$ is invariant for g in $\mathrm{Clos}(X, \mathcal{P})$ if and only if $(\mathcal{P}^c, d^c)$ is invariant for g.*

2. *The pseudometric pattern space $(X, \mathcal{P}, d)$ satisfies one of Axioms 2.2.1, 2.2.2, 2.2.3 and 2.2.4 if and only if $(X, \mathcal{P}^c, d^c)$ satisfies the same axiom.*

Proof. The invariance claim follows directly from Equation 2.8.

The claim about the axioms is shown as follows. Consider the axioms as predicates for $(X, \mathcal{P}, d)$. It suffices to show that replacing A with $X - A$, B with $X - B$, and d with d^c results in an axiom equivalent to the original one. For Axiom 2.2.1 this is true because

$$d^c(X - A, t(X - A)) = d^c(X - A, X - t(A)) = d(A, t(A)).$$

Equivalence of Axioms 2.2.2, 2.2.3, and 2.2.4 follows from the following observations:

1. By the definition of the boundary of a set, $\mathrm{Bd}\,(X - A) = \mathrm{Bd}\,(A)$ and $\mathrm{Bd}\,(X - B) = \mathrm{Bd}\,(B)$.

2. A crack of $X - A$ is a crack of A and vice versa.

3. The equation $(X - B) - U = (X - A) - U$ is equivalent to $B - U = A - U$.

□

2.3 Embedding patterns in a function space

This section presents a method for constructing pseudometrics on a collection of patterns. The construction method is based on the assignment of real valued functions to patterns. Simple conditions on the assignment guarantee that the formed pseudometrics are invariant for a given transformation group.

A class of functions with a pseudonorm

The construction method assigns a function to each pattern. These functions belong to a class that is defined below. For each $p \geq 1$, define $\boldsymbol{I}_p^k$ to be the class of all functions $\boldsymbol{a} : \mathbb{R}^k \to \mathbb{R}$ for which $x \mapsto |\boldsymbol{a}(x)|^p$ is integrable. Throughout this thesis, Lebesgue integration is used. That is, integrals are defined in terms of the Lebesgue k-dimensional measure on $\mathbb{R}^k$ (which is discussed in the book by Munroe [102], for example). The integral of a function f over a set $S \subseteq \mathbb{R}^k$ is written as $\int_S f(x)\, dx$.

The class of functions has a number of important properties. The class $\boldsymbol{I}_p^k$ is a vector space under pointwise scalar multiplication and addition. Pointwise summation and subtraction of real valued functions $\boldsymbol{a}$ and $\boldsymbol{b}$ are denoted by $\boldsymbol{a}+\boldsymbol{b}$ and $\boldsymbol{a}-\boldsymbol{b}$, respectively. Write $\boldsymbol{a}\wedge\boldsymbol{b}$ and $\boldsymbol{a}\vee\boldsymbol{b}$ for the pointwise minimum and maximum of real valued functions, respectively. The class $\boldsymbol{I}_p^k$ is closed under pointwise addition, subtraction, minimum and maximum. That is, if $\boldsymbol{a}$ and $\boldsymbol{b}$ are members of $\boldsymbol{I}_p^k$, then $\boldsymbol{a}+\boldsymbol{b}$, $\boldsymbol{a}-\boldsymbol{b}$, $\boldsymbol{a}\wedge\boldsymbol{b}$, and $\boldsymbol{a}\vee\boldsymbol{b}$ are members of $\boldsymbol{I}_p^k$.

An important subclass of $\boldsymbol{I}_p^k$ is formed by the characteristic functions of Lebesgue measurable sets. For a subset P of $\mathbb{R}^k$, the characteristic function $\mathbf{1}_P$ is given by

$$\mathbf{1}_P(x) = \begin{cases} 1 & \text{if } x \text{ is in } P, \\ 0 & \text{if } x \text{ is in } \mathbb{R}^k - P. \end{cases} \tag{2.9}$$

For each Lebesgue measurable subset P of $\mathbb{R}^k$, the corresponding $\mathbf{1}_P$ is a member of $\boldsymbol{I}_p^k$, for each $p \geq 1$. The compact subsets of $\mathbb{R}^k$ are examples of Lebesgue measurable sets.

The class $\boldsymbol{I}_p^k$ is endowed with structure by the following definition.

Definition 2.3.1. For each $p \geq 1$, the function $\|\cdot\|_p : \boldsymbol{I}_p^k \to \mathbb{R}$ is defined by

$$\|\boldsymbol{a}\|_p = \sqrt[p]{\int_{\mathbb{R}^k} |\boldsymbol{a}(x)|^p \, dx}. \tag{2.10}$$

For each $p \geq 1$, the function $\|\cdot\|_p$ is a pseudonorm on $\boldsymbol{I}_p^k$. This means that it satisfies the following two axioms:

1. For each $\gamma \in \mathbb{R}$ and each $\boldsymbol{a}$ in $\boldsymbol{I}_p^k$, $\|\gamma\boldsymbol{a}\|_p = |\gamma| \, \|\boldsymbol{a}\|_p$.

2. For each $\boldsymbol{a}$ and $\boldsymbol{b}$ in $\boldsymbol{I}_p^k$, $\|\boldsymbol{a}+\boldsymbol{b}\|_p \leq \|\boldsymbol{a}\|_p + \|\boldsymbol{b}\|_p$.

The second axiom is known as the Minkowski integral inequality.

A pseudometric on the class of functions

The following definition derives a pseudometric $\boldsymbol{l}_p$ on $\boldsymbol{I}_p^k$ from the pseudonorm $\|\cdot\|_p$.

Definition 2.3.2. For $p \geq 1$, the pseudometric $\boldsymbol{l}_p$ on $\boldsymbol{I}_p^k$ is given by

$$\boldsymbol{l}_p(\boldsymbol{a}, \boldsymbol{b}) = \|\boldsymbol{a} - \boldsymbol{b}\|_p. \tag{2.11}$$

The construction of pseudometrics on a collection of patterns works as follows. Each P in a collection of patterns, say $\mathcal{P}$, is assigned a function $\boldsymbol{n}_P$ in $\boldsymbol{I}_p^k$. The construction of the method depends on the definition of this assignment. The assignment of functions to patterns determines a family $\{\, \boldsymbol{n}_P \,\}_{P\in\mathcal{P}}$ of functions which is indexed by the pattern collection $\mathcal{P}$. An example of an assignment is choosing $\boldsymbol{n}_P$ as the characteristic function $\mathbf{1}_P$ for each P in $\mathcal{P}$. Each assignment of functions to patterns results in a pseudometric on the patterns defined using the distance $\boldsymbol{l}_p$ on $\boldsymbol{I}_p^k$.

The following theorem places a condition on the assignment of functions to patterns under which the resulting pseudometric is invariant Recall from Appendix B that Dif^k denotes the class of all diffeomorphisms from $\mathbb{R}^k$ onto itself. Furthermore, $\delta_g(x)$ denotes the absolute value of the Jacobi determinant of g in x

Theorem 2.3.1. *Let $\{\, \boldsymbol{n}_P \,\}_{P\in\mathcal{P}}$ be a set of nonnegative valued members of $\boldsymbol{I}_p^k$ that is indexed by a collection $\mathcal{P}$. Construct a pseudometric d on $\mathcal{P}$ by*

$$d(A,B) = \boldsymbol{l}_p(\boldsymbol{n}_A, \boldsymbol{n}_B). \tag{2.12}$$

Let g be an element of Dif^k for which the action on $\mathcal{P}$ is defined. Suppose that

$$\boldsymbol{n}_{g(P)}(g(x)) = \delta_g(x)^{-1/r}\, \boldsymbol{n}_P(x) \tag{2.13}$$

for each P in $\mathcal{P}$ and each x in $\mathbb{R}^k$. Then, $(\mathcal{P}, d)$ is invariant for g.

Proof. It must be shown that for all A and B in $\mathcal{P}$,

$$\boldsymbol{l}_p(\boldsymbol{n}_{g(A)}, \boldsymbol{n}_{g(B)}) = \boldsymbol{l}_p(\boldsymbol{n}_A, \boldsymbol{n}_B). \tag{2.14}$$

Set $\boldsymbol{u} = \boldsymbol{n}_A - \boldsymbol{n}_B$, and $\boldsymbol{v} = \boldsymbol{n}_{g(A)} - \boldsymbol{n}_{g(B)}$. Now $\boldsymbol{v}(g(x)) = \delta_g(x)^{-1/r}\boldsymbol{u}(x)$. Substitution of variables in a k-dimensional integral gives

$$\begin{aligned}
(\|\boldsymbol{v}\|_p)^p &= \int_{\mathbb{R}^k} |\boldsymbol{v}(x)|^p \, dx \\
&= \int_{\mathbb{R}^k} \delta_g(x)\, |\boldsymbol{v}(g(x))|^p \, dx \\
&= \int_{\mathbb{R}^k} \delta_g(x) \left| \delta_g(x)^{-\frac{1}{p}} \boldsymbol{u}(x) \right|^p dx \\
&= \int_{\mathbb{R}^k} |\boldsymbol{u}(x)|^p \, dx \\
&= (\|\boldsymbol{u}\|_p)^p .
\end{aligned}$$

□

A normalised pseudometric on the class of functions

The next definition presents a normalised version of $\boldsymbol{l}_p$.

Definition 2.3.3. For each $p \geq 1$, the function $\boldsymbol{l}_p^*$ from $\boldsymbol{I}_p^k \times \boldsymbol{I}_p^k$ to $\mathbb{R}$ is given by

$$\boldsymbol{l}_p^*(\boldsymbol{a}, \boldsymbol{b}) = \|\boldsymbol{a} - \boldsymbol{b}\|_p / \|\boldsymbol{a} \vee \boldsymbol{b}\|_p. \tag{2.15}$$

Analogous to $\boldsymbol{l}_p$, the function $\boldsymbol{l}_p^*$ can be used to construct pseudometrics on a pattern collection. However, first it must be shown that $\boldsymbol{l}_p^*$ is a pseudometric. The following lemma is used to prove this.

Lemma 2.3.2. *Let $\boldsymbol{u}$, $\boldsymbol{a}$, and $\boldsymbol{b}$ be nonnegative valued members of $\boldsymbol{I}_p^k$. If $\boldsymbol{u} \leq \boldsymbol{a}$ and $\boldsymbol{u} \wedge \boldsymbol{b} = \boldsymbol{a} \wedge \boldsymbol{b}$, then $\boldsymbol{l}_p^*(\boldsymbol{u}, \boldsymbol{b}) \leq \boldsymbol{l}_p^*(\boldsymbol{a}, \boldsymbol{b})$.*

Proof. Observe that for functions $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{z}$ satisfying $\boldsymbol{0} \leq \boldsymbol{x} \leq \boldsymbol{y}$, and $\boldsymbol{0} \leq \boldsymbol{z}$, it follows that

$$\|\boldsymbol{x}\|_p / \|\boldsymbol{y}\|_p \leq \|\boldsymbol{x} + \boldsymbol{z}\|_p / \|\boldsymbol{y} + \boldsymbol{z}\|_p. \tag{2.16}$$

In particular, choose $\boldsymbol{x} = |\boldsymbol{u} - \boldsymbol{b}|$, $\boldsymbol{y} = |\boldsymbol{u} \vee \boldsymbol{b}|$, and $\boldsymbol{z} = \boldsymbol{a} - \boldsymbol{u}$. From $\boldsymbol{u} \leq \boldsymbol{a}$ and $\boldsymbol{u} \wedge \boldsymbol{b} = \boldsymbol{a} \wedge \boldsymbol{b}$, the equalities $|\boldsymbol{x} + \boldsymbol{z}| = |\boldsymbol{a} - \boldsymbol{b}|$ and $|\boldsymbol{y} + \boldsymbol{z}| = |\boldsymbol{a} \vee \boldsymbol{b}|$ follow. Substituting these equalities in Equation 2.16 gives

$$\begin{aligned} \boldsymbol{l}_p^*(\boldsymbol{u}, \boldsymbol{b}) &= \|\boldsymbol{u} - \boldsymbol{b}\|_p / \|\boldsymbol{u} \vee \boldsymbol{b}\|_p \\ &= \|\boldsymbol{x}\|_p / \|\boldsymbol{y}\|_p \\ &\leq \|\boldsymbol{x} + \boldsymbol{z}\|_p / \|\boldsymbol{y} + \boldsymbol{z}\|_p \\ &= \|\boldsymbol{a} - \boldsymbol{b}\|_p / \|\boldsymbol{a} \vee \boldsymbol{b}\|_p \\ &= \boldsymbol{l}_p^*(\boldsymbol{a}, \boldsymbol{b}). \end{aligned} \tag{2.17}$$

□

This lemma is applied in the next theorem, which states that $\boldsymbol{l}_p^*$ is a pseudometric.

Theorem 2.3.3. *For each $p \geq 1$, the function $\boldsymbol{l}_p^*$ is a pseudometric on the nonnegative valued members of $\boldsymbol{I}_p^k$.*

Proof. Axiom 2.1.1: Trivial.

Axiom 2.1.7: This means proving

$$\boldsymbol{l}_p^*(\boldsymbol{b}, \boldsymbol{c}) \leq \boldsymbol{l}_p^*(\boldsymbol{a}, \boldsymbol{b}) + \boldsymbol{l}_p^*(\boldsymbol{a}, \boldsymbol{c}). \tag{2.18}$$

Let $\boldsymbol{u} = \boldsymbol{a} \wedge (\boldsymbol{b} \vee \boldsymbol{c})$. Observe that $\boldsymbol{u} \vee \boldsymbol{b} \leq \boldsymbol{c} \vee \boldsymbol{b}$ and $\boldsymbol{u} \vee \boldsymbol{c} \leq \boldsymbol{b} \vee \boldsymbol{c}$. Since the functions are nonnegative it follows that $\|\boldsymbol{u} \vee \boldsymbol{b}\|_p \leq \|\boldsymbol{c} \vee \boldsymbol{b}\|_p$ and $\|\boldsymbol{u} \vee \boldsymbol{b}\|_p \leq$

$\|\boldsymbol{c} \vee \boldsymbol{b}\|_p$. This results in the inequality

$$\begin{aligned} \boldsymbol{l}_p^*(\boldsymbol{b}, \boldsymbol{c}) &= \|\boldsymbol{b} - \boldsymbol{c}\|_p / \|\boldsymbol{b} \vee \boldsymbol{c}\|_p \\ &\leq \|\boldsymbol{u} - \boldsymbol{b}\|_p / \|\boldsymbol{c} \vee \boldsymbol{b}\|_p + \|\boldsymbol{u} - \boldsymbol{c}\|_p / \|\boldsymbol{b} \vee \boldsymbol{c}\|_p \\ &\leq \|\boldsymbol{u} - \boldsymbol{b}\|_p / \|\boldsymbol{u} \vee \boldsymbol{b}\|_p + \|\boldsymbol{u} - \boldsymbol{c}\|_p / \|\boldsymbol{u} \vee \boldsymbol{c}\|_p \\ &= \boldsymbol{l}_p^*(\boldsymbol{u}, \boldsymbol{b}) + \boldsymbol{l}_p^*(\boldsymbol{u}, \boldsymbol{c}). \end{aligned} \tag{2.19}$$

Applying Lemma 2.3.2, it follows that $\boldsymbol{l}_p^*(\boldsymbol{u}, \boldsymbol{b}) \leq \boldsymbol{l}_p^*(\boldsymbol{a}, \boldsymbol{b})$ and $\boldsymbol{l}_p^*(\boldsymbol{u}, \boldsymbol{c}) \leq \boldsymbol{l}_p^*(\boldsymbol{a}, \boldsymbol{c})$. □

The next theorem, presented shortly, shows how the pseudometric $\boldsymbol{l}_p^*$ can be used to obtain pseudometrics that are invariant for a particular transformation group. In a sense, the normalised distance $\boldsymbol{l}_p^*$ is more powerful than the distance $\boldsymbol{l}_p$ because the condition required to obtain invariance is weaker.

Let CDif^k be the subgroup of Dif^k consisting of all transformations with constant Jacobi determinant. That is, for each g in CDif^k, the function $x \mapsto \delta_g(x)$ is constant. For such g, the value of $\delta_g(x)$ is denoted δ_g.

Theorem 2.3.4. *Let $\{ \boldsymbol{n}_P \}_{P \in \mathcal{P}}$ be a set of nonnegative valued members of $\boldsymbol{I}_p^k$ that is indexed by a collection $\mathcal{P}$. Construct a pseudometric d on $\mathcal{P}$ by*

$$d(A, B) = \boldsymbol{l}_p^*(\boldsymbol{n}_A, \boldsymbol{n}_B). \tag{2.20}$$

Let g be an element of CDif^k for which the action on $\mathcal{P}$ is defined. Suppose that there is a $\gamma_g > 0$ such that

$$\boldsymbol{n}_{g(P)}(g(x)) = \gamma_g \boldsymbol{n}_P(x) \tag{2.21}$$

for each P in $\mathcal{P}$ and each x in $\mathbb{R}^k$. Then, $(\mathcal{P}, d)$ is invariant for g.

Proof. It must be shown that for all $A, B \in \mathcal{P}$,

$$\boldsymbol{l}_p^*(\boldsymbol{n}_{g(A)}, \boldsymbol{n}_{g(B)}) = \boldsymbol{l}_p^*(\boldsymbol{n}_A, \boldsymbol{n}_B). \tag{2.22}$$

Applying substitution of variables using the constant δ_g gives

$$\begin{aligned} \boldsymbol{l}_p^*(\boldsymbol{n}_{g(A)}, \boldsymbol{n}_{g(B)}) &= \|\boldsymbol{n}_{g(A)} - \boldsymbol{n}_{g(B)}\|_p / \|\boldsymbol{n}_{g(A)} \vee \boldsymbol{n}_{g(B)}\|_p \\ &= j^{1/p} \|\boldsymbol{n}_{g(A)} \circ g - \boldsymbol{n}_{g(B)} \circ g\|_p / j^{1/p} \|\boldsymbol{n}_{g(A)} \circ g \vee \boldsymbol{n}_{g(B)} \circ g\|_p \\ &= \|\boldsymbol{n}_{g(A)} \circ g - \boldsymbol{n}_{g(B)} \circ g\|_p / \|\boldsymbol{n}_{g(A)} \circ g \vee \boldsymbol{n}_{g(B)} \circ g\|_p \\ &= \boldsymbol{l}_p^*(\boldsymbol{n}_{g(A)} \circ g, \boldsymbol{n}_{g(B)} \circ g) \\ &= \boldsymbol{l}_p^*(\gamma_g \boldsymbol{n}_A, \gamma_g \boldsymbol{n}_B) \\ &= \boldsymbol{l}_p^*(\boldsymbol{n}_A, \boldsymbol{n}_B). \end{aligned}$$

□

Applications

The section is concluded by a brief discussion of the applications of the previously described constructions.

First, it is shown that two pseudometrics induced by an assignment of functions using $\boldsymbol{l}_p$ and $\boldsymbol{l}_p^*$ are topologically equivalent. This implies that the two pseudometrics formed this manner satisfy the same robustness axioms (by Theorem 2.2.1).

Theorem 2.3.5. *Let $\{\, \boldsymbol{n}_A \,\}_{A\in\mathcal{P}}$ be a set of nonnegative valued members of $\boldsymbol{I}_p^k$ that is indexed by a collection $\mathcal{P}$. Then, the topologies induced on $\mathcal{P}$ by $\boldsymbol{l}_p$ and $\boldsymbol{l}_p^*$ are equivalent for each $p \geq 1$.*

Proof. First, it is shown that $(\mathcal{P}, \boldsymbol{l}_p)$ is finer than $(\mathcal{P}, \boldsymbol{l}_p^*)$ in terms of topology. Let A be in $\mathcal{P}$, and $\epsilon > 0$. Choose $\delta = \epsilon\|\boldsymbol{n}_A\|_p$. If $\boldsymbol{l}_p(\boldsymbol{n}_A, \boldsymbol{n}_B) < \delta$, then

$$\begin{aligned}\boldsymbol{l}_p^*(\boldsymbol{n}_A, \boldsymbol{n}_B) &= \boldsymbol{l}_p(\boldsymbol{n}_A, \boldsymbol{n}_B)/\|\boldsymbol{n}_A \vee \boldsymbol{n}_B\|_p \\ &< \delta/\|\boldsymbol{n}_A \vee \boldsymbol{n}_B\|_p \\ &\leq \delta/\|\boldsymbol{n}_A\|_p \\ &= \epsilon.\end{aligned}$$

Second, it is shown that $(\mathcal{P}, \boldsymbol{l}_p^*)$ is finer than $(\mathcal{P}, \boldsymbol{l}_p)$ in terms of topology. Let A be in $\mathcal{P}$, and $\epsilon > 0$. Observe that $f : \xi \mapsto \xi/(\|\boldsymbol{n}_A\|_p + \xi)$ is an order preserving bijection from the positive real numbers to the interval $(0, 1)$. It follows that

$$\begin{aligned}f(\boldsymbol{l}_p(\boldsymbol{n}_A, \boldsymbol{n}_B)) &= \boldsymbol{l}_p(\boldsymbol{n}_A, \boldsymbol{n}_B)/\left(\|\boldsymbol{n}_A\|_p + \boldsymbol{l}_p(\boldsymbol{n}_A, \boldsymbol{n}_B)\right) \\ &\leq \boldsymbol{l}_p(\boldsymbol{n}_A, \boldsymbol{n}_B)/\left(\|\boldsymbol{n}_A \wedge \boldsymbol{n}_B\|_p + \boldsymbol{l}_p(\boldsymbol{n}_A, \boldsymbol{n}_B)\right) \\ &= \boldsymbol{l}_p^*(\boldsymbol{n}_A, \boldsymbol{n}_B).\end{aligned}$$

Therefore, $\delta > 0$ may be chosen so that $\boldsymbol{l}_p^*(\boldsymbol{n}_A, \boldsymbol{n}_B) < \delta$ implies $\boldsymbol{l}_p(\boldsymbol{n}_A, \boldsymbol{n}_B) < \epsilon$. □

The construction methods and the corresponding results are applied in the following sections. The Hausdorff metric, discussed in Section 2.4, can be seen as an application of the construction method that uses the special pseudometric $\boldsymbol{l}_\infty$ which can be expressed as

$$\boldsymbol{l}_\infty(\boldsymbol{a}, \boldsymbol{b}) = \sup_{x\in\mathbb{R}^k} |\boldsymbol{a}(x) - \boldsymbol{b}(x)|. \tag{2.23}$$

The volume of symmetric difference, discussed in Section 2.5, can be seen as an application of the construction method where $\boldsymbol{l}_1$ is used. Replacing $\boldsymbol{l}_1$ with $\boldsymbol{l}_1^*$ gives rise to a new similarity measure which is called the normalised volume

of symmetric difference. In Section 2.6, visibility related structures are used to assign functions to patterns. Combining this assignment with the normalised pseudometric $\boldsymbol{l}_1^*$ results in an affine invariant metric.

2.4 The Hausdorff metric

Hausdorff was the first to topologise a collection of subsets of a topological space [76]. This was done using a metric which is known as the Hausdorff metric. The Hausdorff metric is defined on each collection of nonempty closed, bounded subsets of some metric base space. This section explores the Hausdorff metric and its properties. The discussion includes various equivalent definitions, the metric topology, and robustness issues.

In the most general case, the Hausdorff metric is defined on subsets of some metric base space. In this section, three distinct but equivalent expressions for this general definition of Hausdorff metric are discussed. The topology induced by the Hausdorff metric is investigated. An existing result is extended by proving that the extension of the Hausdorff metric with the empty set induces exactly the myope topology. Furthermore, it is observed that the Hausdorff metric is invariant for a transformation if and only if the base metric is invariant for it. The definition of the myope topology is applied to find simple proofs for deformation, blur, and crack robustness. It is also shown that for each base space with at least two elements, noise robustness does not hold. Finally, several variants on the Hausdorff metric are mentioned.

Definition

The definition of the Hausdorff metric, given shortly, is based on the notion of ϵ neighbourhoods. In a pseudometric space (X, ρ), the ϵ neighbourhood of a subset P of X, denoted with $\mathrm{N}_\rho(P, \epsilon)$, is the union of all open balls centred at points of P having radius ϵ. The Hausdorff metric is not defined on all subsets of a metric space. For this reason, restricted collections of subsets are given in the following definition.

Definition 2.4.1. Let X be a metric space. The collection $\mathcal{C}(X)$ consists of all closed and bounded subsets of X, and $\mathcal{C}'(X)$ consists of all nonempty elements of $\mathcal{C}(X)$.

The Hausdorff metric is defined on the collection $\mathcal{C}'(X)$ as follows.

Definition 2.4.2. Let (X, ρ) be a metric space. The *Hausdorff metric* h_ρ on $\mathcal{C}'(X)$ is given by

$$\mathsf{h}_\rho(A, B) = \inf\{\, \epsilon > 0 \mid \; A \subseteq \mathrm{N}_\rho(B, \epsilon) \text{ and } B \subseteq \mathrm{N}_\rho(A, \epsilon) \;\}. \tag{2.24}$$

The following theorem states that h_ρ is indeed a metric if ρ is a metric. In case of a pseudometric ρ, the resulting h_ρ is a pseudometric.

Theorem 2.4.1. *If (X, ρ) is a metric space, then h_ρ is a metric on $\mathcal{C}'(X)$. If ρ is only assumed to be pseudometric, then h_ρ is a pseudometric.*

Proof. Recall from Section 2.1 that h_ρ is a pseudometric if it satisfies self-identity (Axiom 2.1.1) and the alternative triangle inequality (Axiom 2.1.7). First, it is verified that the axioms for a pseudometric hold for h_ρ if the base metric ρ is a pseudometric.

Axiom 2.1.1: Observe that self-identity for ρ implies that $A \subseteq \mathrm{N}_\rho(A, \epsilon)$ for all $\epsilon > 0$. This means that $\mathsf{h}_\rho(A, A) = 0$.

Axiom 2.1.7: Let $R_\epsilon(A, B)$ denote the condition that $A \subseteq \mathrm{N}_\rho(B, \epsilon)$ and $B \subseteq \mathrm{N}_\rho(A, \epsilon)$. Observe that it is sufficient to show that $R_\beta(A, B)$ and $R_\gamma(A, C)$ implies $R_{\beta+\gamma}(B, C)$ for all $\beta, \gamma > 0$. This is implied by the inclusion

$$\mathrm{N}_\rho(\mathrm{N}_\rho(A, \beta), \gamma) \subseteq \mathrm{N}_\rho(A, \beta + \gamma) \tag{2.25}$$

that holds for all A, B and C in $\mathcal{C}'(X)$ and all $\beta > 0$ and $\gamma > 0$. This inclusion is a simple consequence of the fact that ρ is a pseudometric.

Finally, it is is verified that positivity for h_ρ is implied by positivity for ρ. Let A and B be in $\mathcal{C}'(X)$. Without loss a generality, assume there is a point a that is in A but not in B. There is an open ball $\mathrm{B}_\rho(a, \epsilon)$ with radius $\epsilon > 0$ that is disjoint with B. Otherwise the point a would have been a limit point of B, and therefore an element of B. Since A cannot be a subset of $\mathrm{N}_\rho(B, \epsilon)$, the distance $\mathsf{h}_\rho(A, B)$ is at least ϵ. □

There are other definitions of the Hausdorff metric. A familiar example is the expression of the Hausdorff metric as a supremum of infima. I did not find equivalence proofs for the other definitions of the Hausdorff metric in literature. The exact conditions under which the equivalences hold are not immediately clear. For this reason, the equivalence proofs are provided below.

The following theorem states that the "sup-inf" definition of the Hausdorff metric is equivalent with the previous definition for each metric base space X and all of $\mathcal{C}'(X)$.

Theorem 2.4.2. *The* directed *Hausdorff distance $\mathsf{h}^\triangleright_\rho$ on $\mathcal{C}'(X)$ is given by*

$$\mathsf{h}^\triangleright_\rho(A, B) = \sup_{a \in A} \inf_{b \in B} \rho(a, b). \tag{2.26}$$

The Hausdorff metric for A and B in $\mathcal{C}'(X)$ equals

$$\mathsf{h}_\rho(A, B) = \max\{\, \mathsf{h}^\triangleright_\rho(A, B), \mathsf{h}^\triangleright_\rho(B, A) \,\}. \tag{2.27}$$

Proof. Abbreviate $\boldsymbol{v}_P(x) = \inf_{p \in P} \rho(x,p)$. Observe that $a \in \mathrm{N}_\rho(B,\epsilon)$ if and only if $\boldsymbol{v}_B(a) < \epsilon$. Using this observation, an equivalent definition for the directed Hausdorff distance is derived by

$$\begin{aligned}
\mathsf{h}^{\triangleright}_\rho(A,B) &= \sup\{\, \boldsymbol{v}_B(a) \mid a \in A \,\} \\
&= \inf\{\, \epsilon > 0 \mid \ \boldsymbol{v}_B(a) \leq \epsilon \text{ for all } a \in A \ \} \\
&= \inf\{\, \epsilon > 0 \mid \ a \in \mathrm{N}_\rho(B,\epsilon) \text{ for all } a \in A \ \} \\
&= \inf\{\, \epsilon > 0 \mid A \subseteq \mathrm{N}_\rho(B,\epsilon) \,\}.
\end{aligned}$$

The Hausdorff metric can be written in terms of the alternative form of the directed Hausdorff distance as

$$\begin{aligned}
\mathsf{h}_\rho(A,B) &= \inf\{\, \epsilon > 0 \mid \ A \subseteq \mathrm{N}_\rho(B,\epsilon) \text{ and } B \subseteq \mathrm{N}_\rho(A,\epsilon) \ \} \\
&= \max\{\, \inf\{\, \epsilon > 0 \mid A \subseteq \mathrm{N}_\rho(B,\epsilon) \,\}, \inf\{\, \epsilon > 0 \mid B \subseteq \mathrm{N}_\rho(A,\epsilon) \,\} \,\}.
\end{aligned}$$

□

The domain of $\mathsf{h}^{\triangleright}_\rho(A,B)$ may be restricted so that the supremum is a maximum and the infimum is a minimum, resulting in the expression

$$\mathsf{h}^{\triangleright}_\rho(A,B) = \max_{a \in A} \min_{b \in B} \rho(a,b). \tag{2.28}$$

For this purpose, it is convenient to define the following collections of subsets of a space.

Definition 2.4.3. Let X be a topological space. The collection $\mathcal{K}(X)$ consists of all compact subsets of X. The collection $\mathcal{K}'(X)$ consists of all nonempty elements of $\mathcal{K}(X)$.

Each metric is a continuous function from $X \times X$ to $\mathbb{R}$. Using this fact, it follows that the maxima and minima in Equation 2.28 exist for each element of $\mathcal{C}'(X) \cap \mathcal{K}'(X)$. This means that the "max-min" expression of the Hausdorff metric determined by Equations 2.28 and 2.27 is defined on the collection $\mathcal{C}'(X) \cap \mathcal{K}'(X)$. If X is a Hausdorff space (Appendix A), then each compact subset of X is closed. Each metric space is a Hausdorff space. In this case, $\mathcal{K}'(X)$ is a subcollection of $\mathcal{C}'(X)$, implying that $\mathcal{C}'(X) \cap \mathcal{K}'(X) = \mathcal{K}'(X)$. In the special case of a finite-dimensional Euclidean space $\mathbb{R}^k$, the collections $\mathcal{C}'(\mathbb{R}^k)$ and $\mathcal{K}'(\mathbb{R}^k)$ coincide, implying that $\mathcal{C}'(X) \cap \mathcal{K}'(X) = \mathcal{C}'(X)$.

A less known equivalent expression for the Hausdorff metric is due to Baddeley [22]. This expression can be seen as a special application of the construction method of Section 2.3. In the proof of Theorem 2.4.2, the abbreviation $\boldsymbol{v}_P$ was used. This abbreviation corresponds to a well known structure in computational geometry, which is formally defined as follows. The Voronoi surface of

a nonempty subset P of X, denoted by $\boldsymbol{v}_P$, is a function from X to $\mathbb{R}$ given by $\boldsymbol{v}_P(x) = \inf_{p\in P} \rho(x,p)$. The following theorem states that the supremum of the pointwise differences of the Voronoi surfaces corresponding to two sets is an equivalent definition of the Hausdorff metric.

Theorem 2.4.3 (Baddeley, 1992). *Let (X,ρ) be a metric space. The Hausdorff metric on $\mathcal{C}'(X)$ is given by*

$$\mathsf{h}_\rho(A,B) = \sup_{x\in X} |\boldsymbol{v}_A(x) - \boldsymbol{v}_B(x)|. \tag{2.29}$$

Proof. The following is a more direct alternative for the proof given by Baddeley.

Observe that $|\boldsymbol{v}_A(x) - \boldsymbol{v}_B(x)| = \boldsymbol{v}_B(x)$ for $x \in A$ and $|\boldsymbol{v}_A(x) - \boldsymbol{v}_B(x)| = \boldsymbol{v}_A(x)$ for $x \in B$. Therefore, $\mathsf{h}_\rho(A,B) \le \sup_{x\in X} |\boldsymbol{v}_A(x) - \boldsymbol{v}_B(x)|$.

From the triangle inequality it follows that $\rho(a,b) \ge \rho(x,b) - \rho(x,a)$ for all a, b and x in X. By subsequently taking an infimum over A and a supremum over B over both sides of this inequality it follows that for all x in X,

$$\begin{aligned}
\mathsf{h}^{\triangleright}_\rho(A,B) &= \sup_{a\in A} \inf_{b\in B} \rho(a,b) \\
&\ge \sup_{a\in A} \inf_{b\in B} (\rho(x,b) - \rho(x,a)) \\
&= \left(\inf_{b\in B} \rho(x,b)\right) - \left(\inf_{a\in A} \rho(x,a)\right) \\
&= \boldsymbol{v}_A(x) - \boldsymbol{v}_B(x).
\end{aligned}$$

A similar inequality is obtained if A and B exchange roles. The two inequalities combined show that $\mathsf{h}_\rho(A,B) \ge |\boldsymbol{v}_A(x) - \boldsymbol{v}_B(x)|$ for each $x \in X$. This is equivalent to the desired inequality $\mathsf{h}_\rho(A,B) \ge \sup_{x\in X} |\boldsymbol{v}_A(x) - \boldsymbol{v}_B(x)|$. □

The metric topology

In general, the topology determined by the Hausdorff metric on a collection of subsets of X is not determined by the topology of the metric base space (X,ρ). Kelley [89] (page 131) gives an example in which two metrics on X with equivalent topologies induce Hausdorff metrics with different topologies on a class of subsets of X.

For the collection $\mathcal{K}'(X)$, the topology induced by the Hausdorff metric *is* completely determined by the metric topology of (X,ρ). Having the Hausdorff metric, $\mathcal{K}'(X)$ is a topological subspace of $\mathcal{K}(X)$, if the latter has the myope topology. The myope topology, defined below, is defined in terms of a topological space X.

The description of the myope topology given here is equivalent to that of Matheron [94]. The myope topology is defined in terms of a subbasis for $\mathcal{K}(X)$. The notion of a subbasis is explained in Appendix A. For each subset P of X, denote with $\mathcal{K}_P(X)$ and $\mathcal{K}^P(X)$ the subcollections of $\mathcal{K}(X)$ given by

$$\mathcal{K}_P(X) = \{\, K \in \mathcal{K}(X) \mid K \cap P \neq \varnothing \,\}, \tag{2.30}$$

$$\mathcal{K}^P(X) = \{\, K \in \mathcal{K}(X) \mid K \cap P = \varnothing \,\}. \tag{2.31}$$

The subbasis $\mathfrak{S}$ for the myope topology on $\mathcal{K}(X)$ is a family of subcollections of $\mathcal{K}(X)$. This family consists of all collections of the form $\mathcal{K}_V(X)$, where V is an open subset of X and all collections of the form $\mathcal{K}^C(X)$, where C is a closed subset of X. In other words, each element of $\mathfrak{S}$ is a collection of compact subsets of X that is either intersected by some open subset of X or contained in the complement of some closed subset of X. The basis $\mathfrak{B}$ for the myope topology for $\mathcal{K}(X)$ is the family of all finite intersections of elements in $\mathfrak{S}$. Each element of $\mathfrak{B}$ has the form

$$\mathcal{K}^C_{V_1,\ldots,V_n}(X) = \mathcal{K}_{V_1}(X) \cap \cdots \cap \mathcal{K}_{V_n}(X) \cap \mathcal{K}^C(X) \tag{2.32}$$

for open $V_i \subseteq X$ and closed $C \subseteq X$. The myope topology is the family of all unions of elements of $\mathfrak{B}$.

There is a connection between the myope topology and the Hausdorff metric. Assume that (X, ρ) is a metric space. Matheron [94] proves that the relation between the Hausdorff metric and the myope topology is as follows: if $\mathcal{K}(X)$ has the myope topology, then the metric space $(\mathcal{K}'(X), \mathsf{h}_\rho)$ is a topological subspace of $\mathcal{K}(X)$. The next theorem generalises this. It states that the Hausdorff metric extended with the empty set, defined in Equation 2.7 and denoted by $\mathsf{h}^{\varnothing}_\rho$, induces exactly the myope topology on $\mathcal{K}(X)$.

Theorem 2.4.4. *Let (X, ρ) be a metric space. The myope topology on $\mathcal{K}(X)$ coincides with the metric topology on $\mathcal{K}(X)$ induced by $\mathsf{h}^{\varnothing}_\rho$.*

Proof. The proof is achieved by adapting the proof of Proposition 1-4-4 by Matheron [94].

Let $\mathfrak{T}_H$ be the topology induces by the extended Hausdorff metric and let $\mathfrak{T}_M$ be the myope topology. The first part of the original proof shows that $\mathfrak{T}_M$ is finer than $\mathfrak{T}_H$. This part of the proof still works after extending $\mathcal{K}'(X)$ to $\mathcal{K}(X)$.

The second part of the original proof shows that $\mathfrak{T}_H$ is finer than $\mathfrak{T}_M$. This part of the proof needs some adjustments. It follows from that fact that the subbasis elements defining $\mathfrak{T}_M$ are open in $\mathfrak{T}_H$. This is implied by the fact that for each subbasis element containing K in $\mathcal{K}(X)$, there is a ball in the Hausdorff metric that contains K and is contained in the subbasis element.

For the case that $K \neq \varnothing$, the original proof is still sufficient. Below, only the adjustments in the proof are given.

Let $K \in \mathcal{K}(X)$ be contained in a subbasis element $\mathcal{K}_U(X)$, where U is open. Then, K cannot be empty. Case solved.

Let $K \in \mathcal{K}(X)$ be contained in a subbasis element $\mathcal{K}^C(X)$, where C is closed. The new case is that $K = \varnothing$. Observe that the open ball $\mathrm{B}_{\mathsf{h}_\rho}(\varnothing, \epsilon)$ for some $\epsilon < 1$ equals $\{\varnothing\}$ and is therefore contained in $\mathcal{K}^C(X)$. □

If $\mathcal{K}(X)$ is equipped with the myope topology, some facts for the space $\mathcal{K}(X)$ have interesting interpretations in terms of the Hausdorff metric [94]. Below, some are discussed.

The empty set has a special status: it is an isolated point in $\mathcal{K}(X)$. This corresponds to the fact that $\varnothing$ is the only element of the open ball $\mathrm{B}_{\mathsf{h}_\rho^\varnothing}(\varnothing, \epsilon)$, where $\epsilon < 1$. The latter follows directly from the definition of $\mathsf{h}_\rho^\varnothing$ in which the distance between the empty set and a nonempty set is always 1. A consequence is that the collection consisting of only the empty set is a connected component in $\mathcal{K}(X)$.

The collection of finite subsets of X is dense in $\mathcal{K}(X)$. This corresponds to the fact that for each K in $\mathcal{K}(X)$ and each $\epsilon > 0$, there is a finite subset K' with $\mathsf{h}_\rho^\varnothing(K, K') < \epsilon$. This means that under the ($\varnothing$-extended) Hausdorff metric, each compact set can be approximated with arbitrary accuracy using a finite set.

Some operations from mathematical morphology are continuous if $\mathcal{K}(X)$ has the myope topology. Examples are scaling $(\lambda, K) \mapsto \lambda K$ as a function from $\mathbb{R} \times \mathcal{K}(X)$ to $\mathcal{K}(X)$, and Minkowski addition $(K, L) \mapsto K \oplus L$ as a function from $\mathcal{K}(X) \times \mathcal{K}(X)$ to $\mathcal{K}(X)$. By Theorem 2.4.4, these operations are also continuous in terms of the $\varnothing$-extended Hausdorff metric.

Invariance

The Hausdorff metric is invariant for a transformation if and only if the base metric is invariant for that transformation. Clearly, each isometry for the base space is an isometry for the Hausdorff metric. The reverse implication follows from the fact that the function $x \mapsto \{x\}$ is an isometry from (X, ρ) to $(\mathcal{C}'(X), \mathsf{h}_\rho)$. Michael [96] calls a metric on $\mathcal{C}'(X)$ admissible if this property is satisfied.

Axiom satisfaction

Below, it is investigated which of the robustness axioms of Section 2.2 hold for the Hausdorff metric. The proofs apply the notion of a topological pattern space (which is defined on page 36). More precisely, the following observation

is used in the proofs: a robustness axiom holds for the metric pattern space $(X, \mathcal{K}'(X), \mathsf{h}_\rho)$ if and only if the same axiom holds for the topological pattern space $(X, \mathcal{K}'(X), \mathfrak{T})$, where $\mathfrak{T}$ is the relative myope topology for $\mathcal{K}'(X)$.

First it is shown that deformation robustness holds for the Hausdorff metric on the space of nonempty compact subsets in a metric base space.

Theorem 2.4.5. *Let (X, ρ) be a metric space. Then, the metric pattern space $(X, \mathcal{K}'(X), \mathsf{h}_\rho)$ is deformation robust in the sense of Axiom 2.2.1.*

Proof. Let A be in $\mathcal{K}'(X)$ and let $\mathcal{B} = \mathcal{K}^C_{V_1,\ldots,V_n}(X) \cap \mathcal{K}'(X)$ be a basis element of the relative myope topology containing A. Choose an element $a_i \in A \cap V_i$ for each $i = 1, \ldots, n$. The desired subset I of $\mathrm{Hom}(X)$ is given by

$$I = \{\, t \in \mathrm{Hom}(X) \mid \ t(A) \subseteq X - C \text{ and } t(a_i) \in V_i \text{ for } i = 1, \ldots, n \ \}. \quad (2.33)$$

It is not difficult to see that I is a finite intersection of subbasis elements, and therefore open in the compact-open topology. □

Blur robustness also holds on the space of all nonempty compact subsets.

Theorem 2.4.6. *Let (X, ρ) be a metric space. Then, the metric pattern space $(X, \mathcal{K}'(X), \mathsf{h}_\rho)$ is blur robust in the sense of Axiom 2.2.2.*

Proof. Let A be in $\mathcal{K}'(X)$ and let $\mathcal{B} = \mathcal{K}^C_{V_1,\ldots,V_n}(X) \cap \mathcal{K}'(X)$ be a basis element of the myope topology containing A. For each $i = 1, \ldots, n$, an open set W_i is chosen as follows. If V_i is disjoint with the boundary of A, choose x_i as a point in the intersection of V_i and the interior of A and In this case, let W_i be an open neighbourhood of $\mathrm{Bd}\,(A)$ disjoint with x_i. Otherwise, choose $W_i = X$. The desired open neighbourhood U of $\mathrm{Bd}\,(A)$ can be chosen as $U = (X - C) \cap \bigcap_{i=1}^n W_i$. □

Finally, crack robustness holds for all nonempty compact subsets.

Theorem 2.4.7. *Let (X, ρ) be a metric space. Then, the metric pattern space $(X, \mathcal{K}'(X), \mathsf{h}_\rho)$ is crack robust in the sense of Axiom 2.2.3.*

Proof. Let A be in $\mathcal{K}'(X)$ and let $\mathcal{B} = \mathcal{K}^C_{V_1,\ldots,V_n}(X) \cap \mathcal{K}'(X)$ be a basis element of the relative myope topology containing A. In addition, let R be a crack of A. For each $i = 1, \ldots, n$ choose an open set W_i as follows. Observe that it is possible to choose a point a_i in $A \cap V_i$ that is not in R. Choose W_i to be an open neighbourhood of R disjoint with a_i. The desired open neighbourhood U of R is given by $U = (X - C) \cap \bigcap_{i=1}^n W_i$. □

The following theorem states that the Hausdorff metric is not noise robust if it is defined on a collection $\mathcal{P}$ of subsets of X that satisfies a weak condition.

Theorem 2.4.8. *Let (X, ρ) be a metric space. Let $\mathcal{P}$ be subcollection of $\mathcal{C}'(X)$. Suppose that there exists an $A \in \mathcal{P}$ and an $x \notin A$ such that for each $\delta > 0$, there exists a $B \in \mathcal{P}$ for which $B - \mathrm{B}_\rho(x, \delta) = A - \mathrm{B}_\rho(x, \delta)$ and $x \in B$. Then, the metric pattern space $(X, \mathcal{P}, \mathsf{h}_\rho)$ is not noise robust in the sense of Axiom 2.2.4.*

Proof. Assume that the condition of the theorem holds. Choose A and x as in this condition. To prove that the Axiom 2.2.4 does not hold, it is sufficient to show that there is an $\epsilon > 0$ such that for each open neighbourhood U of x, there is a $B \in \mathcal{P}$ such that $B - U = A - U$ and $B \notin \mathrm{B}_{\mathsf{h}_\rho}(A, \epsilon)$.

Using that A is closed and x is not A, choose $\epsilon > 0$ so that $A \cap \mathrm{B}_\rho(x, \epsilon) = \varnothing$. Let U be an open neighbourhood of x. Using the definition of the metric topology, choose $\delta > 0$ such that $\mathrm{B}_\rho(x, \delta)$ is contained in U. By the assumption of the theorem, there exists a $B \in \mathcal{P}$ such that $B - \mathrm{B}_\rho(x, \delta) = A - \mathrm{B}_\rho(x, \delta)$ and $x \in B$. This implies two things. First, $B - U = A - U$. Second, $\mathsf{h}_\rho(A, B) \geq \inf_{a \in A} \rho(x, a) \geq \epsilon$. The latter is expressed equivalently as $B \notin \mathrm{B}_{\mathsf{h}_\rho}(A, \epsilon)$. □

Variants

An important property of the Hausdorff metric is that its value is dominated by the "worst points", that is, the points in a pattern whose distance to the other pattern is maximal. For this reason, the Hausdorff metric does not satisfy the noise robustness axiom. A number of less noise sensitive variants of the Hausdorff metric have been developed.

Huttenlocher and Rucklidge [83] introduce the *partial* Hausdorff distance on finite point sets. This distance is obtained by replacing the maximum with a fixed rank in Equation 2.28 (the max-min expression of the directed Hausdorff distance). These ranks may be different for the two directed partial Hausdorff distances. In general, the partial Hausdorff distance does not satisfy positivity, symmetry, or the triangle inequality.

Venkatasubramanian [122] replaces the maximum of Equation 2.28 with a summation (where the point sets are finite). The result is called the Σ Hausdorff distance. This similarity measure is less sensitive to noise points than the Hausdorff metric. The Σ Hausdorff distance satisfies positivity and symmetry, but does not satisfy the triangle inequality.

Baddeley [23] presents a metric variant of the Hausdorff metric. It can be seen as a generalisation of Equation 2.29. This variant is less sensitive to noise. The generalisation is similar to the function space approach of Section 2.3. Under special conditions on the base space, the supremum in Equation 2.29 can be generalised to the $\boldsymbol{l}_p$ distance on Voronoi surfaces, where $p = 1, \ldots, \infty$. Care must be taken so that the difference between the Voronoi surfaces is integrable. Baddeley shows that this can be achieved by restricting the underlying space or remapping the range of the underlying metric (using the techniques

in Section 2.1, for example).

2.5 The volume of symmetric difference

Similarity measures like the Bottleneck distance, the Fréchet distance, and the Hausdorff metric, are defined in terms of a metric defined on the base space. As a result, these similarity measures are only invariant for those transformations under which the base metric is invariant. This section analyses the volume of symmetric difference, which is a similarity measure defined in terms of a *measure* defined on the base space. As a result, the volume of symmetric difference is invariant for the group of volume preserving diffeomorphisms. A normalised version of the volume of symmetric difference, defined using the methods of Section 2.2, is invariant for the larger group of ratio-of-volume preserving diffeomorphisms. This class includes the affine transformations.

This section investigates the properties of the volume of symmetric difference. In addition, a normalised version of this similarity measure is introduced and analysed. The discussion includes a study of the metric topology, the invariance under transformations, and the satisfaction of the axioms.

Definition

The notions of length in one dimension, area in two dimensions and volume in three dimensions can be generalised to arbitrary dimensions: the k-dimensional volume of a subset P of $\mathbb{R}^k$, denoted by $\mathrm{vol}\,(P)$, is defined as the Lebesgue k-dimensional measure [102]. For k-simplices in $\mathbb{R}^k$, this definition of k-dimensional volume coincides with the familiar definition of volume as the determinant of k vectors. For a subset P of $\mathbb{R}^k$, the Lebesgue integral of the characteristic function over $\mathbb{R}^k$ equals the Lebesgue k-dimensional measure of P.

The symmetric (set) difference of two sets A and B, denoted $A \triangle B$, is the set of all points in A that are not in B and all points in B that are not in A. In formula, $A \triangle B = (A - B) \cup (B - A)$. Let $\mathcal{K}^+(\mathbb{R}^k)$ denote the collection of compact subsets of $\mathbb{R}^k$ with nonzero volume. The volume of symmetric difference and its normalised version are given in the following definition.

Definition 2.5.1. The *volume of symmetric difference* on $\mathcal{K}(\mathbb{R}^k)$ is given by

$$\mathsf{s}(A, B) = \mathrm{vol}\,(A \triangle B). \tag{2.34}$$

The *normalised volume of symmetric difference* on $\mathcal{K}^+(\mathbb{R}^k)$ is given by

$$\mathsf{s}^*(A, B) = \mathrm{vol}\,(A \triangle B) / \mathrm{vol}\,(A \cup B). \tag{2.35}$$

Defined on the collection of compact sets $\mathcal{K}(\mathbb{R}^k)$, the volume of symmetric difference does not have the positivity property. For example, the symmetric difference of two finite point sets always has zero volume. The positivity property holds for the subcollection of the compact subsets given in the following definition.

Definition 2.5.2. Let X be a topological space. The collection of solid subsets of X, denoted by $\mathcal{S}(X)$, consists of all elements of $\mathcal{K}(X)$ that equal the closure of some open subset of X. The notation $\mathcal{S}'(X)$ is used for the nonempty elements of $\mathcal{S}(X)$.

The collection $\mathcal{S}(X)$ can be characterised as all elements P of $\mathcal{K}(X)$ that are equal to the closure of their interior, that is, $\mathrm{Cl}\,(\mathrm{Int}\,(P)) = P$. Observe that each element of $\mathcal{S}'(X)$ has nonzero volume.

Theorem 2.5.1. *The function* s *is a pseudometric on* $\mathcal{K}(\mathbb{R}^k)$ *and a metric on* $\mathcal{S}(\mathbb{R}^k)$. *The function* s^* *is a pseudometric on* $\mathcal{K}^+(\mathbb{R}^k)$ *and a metric on* $\mathcal{S}'(\mathbb{R}^k)$.

Proof. First, it is shown that s and s^* are pseudometrics on $\mathcal{K}(\mathbb{R}^k)$ and $\mathcal{K}^+(\mathbb{R}^k)$, respectively. Recall that the notation $\mathbf{1}_P$ stands for the characteristic function for a subset P of $\mathbb{R}^k$. The similarity measures s and s^* may be expressed in terms of the pseudometrics $\boldsymbol{l}_1$ and $\boldsymbol{l}_1^*$ (defined in Section 2.3) through the equations

$$\mathsf{s}(A, B) = \boldsymbol{l}_1(\mathbf{1}_A, \mathbf{1}_B), \tag{2.36}$$

$$\mathsf{s}^*(A, B) = \boldsymbol{l}_1^*(\mathbf{1}_A, \mathbf{1}_B). \tag{2.37}$$

Second, positivity of s on $\mathcal{S}(\mathbb{R}^k)$ is proven. Without loss of generality, assume that A is not a subset of B. By definition of $\mathcal{S}(\mathbb{R}^k)$, the set difference $\mathrm{Int}\,(A) - B$ is nonempty and open, which implies that $\mathsf{s}(A, B) \geq \mathrm{vol}\,(A - B) > 0$. Similarly, the positivity of s^* on $\mathcal{S}'(\mathbb{R}^k)$ can be shown. □

The topology

Below, the topology induced by the volume of symmetric difference is studied. First, the relation between the myope topology (induced by the $\varnothing$-extended Hausdorff metric) and the topology induced by the volume of symmetric difference is investigated. Let h denote the Hausdorff metric based on $\mathbb{R}^k$ with the Euclidean metric.

Theorem 2.5.2. *For each* $k \geq 1$, *the metric topologies of* $(\mathcal{S}'(\mathbb{R}^k), \mathsf{h})$ *and* $(\mathcal{S}'(\mathbb{R}^k), \mathsf{s})$ *are incomparable.*

Proof. First, it is shown that the metric topology of $(\mathcal{S}(\mathbb{R}^k), \mathsf{h})$ is not finer than that of $(\mathcal{S}(\mathbb{R}^k), \mathsf{s})$. Let A be the Cartesian product of k closed intervals $[0,1]$. It is shown that for each $\delta > 0$, there is a B in $\mathcal{S}(\mathbb{R}^k)$ such that $\mathsf{h}(A,B) < \delta$ and $\mathsf{s}(A,B) \geq 1/2^k$. Let C_n be the union of closed intervals $[(i-\frac{1}{2})/n, i/n]$ for $i = 1, \ldots, n$. Let B_n be the Cartesian product of k copies of C_n. Clearly, n can be chosen so that $\mathsf{h}(A,B_n) = \sqrt{k}/(2n) < \delta$. At the same time $\mathsf{s}(A,B_n) = 1/2^k$ for all $n \geq 1$.

Now, it is shown that the metric topology of $(\mathcal{S}(\mathbb{R}^k), \mathsf{s})$ is not finer than that of $(\mathcal{S}(\mathbb{R}^k), \mathsf{h})$. Again, let A be the Cartesian product of k closed intervals $[0,1]$. For each $\gamma > 0$, define N_γ to be the Cartesian product of k copies of the closed interval $[2, 2+\gamma]$. For each γ, let B_γ be the union of A with N_γ. Clearly, $\gamma > 0$ can be chosen so that $\mathsf{s}(A,B) = \gamma^k < \delta$ given $\delta > 0$. On the other hand, $\mathsf{h}(A,B_\gamma) = 1$ for all $\gamma > 0$. □

The finer-than relation on topologies is preserved if the pattern collection is restricted. Therefore, by Theorem 2.5.2, the myope topology and the metric topology of s on $\mathcal{K}(\mathbb{R}^k)$ are incomparable.

One can also compare the topologies induced on $\mathcal{K}^+(\mathbb{R}^k)$ by s and its normalised variant s^*. The following corollary, a direct consequence of Theorem 2.3.5, shows that the two topologies are equal.

Corollary 2.5.3. *The pseudometric spaces* $(\mathcal{K}^+(\mathbb{R}^k), \mathsf{s})$ *and* $(\mathcal{K}^+(\mathbb{R}^k), \mathsf{s}^*)$ *have equivalent topologies.*

Invariance

The following theorem states that the group of volume preserving diffeomorphisms UDif^k, defined in Appendix B, the set of all diffeomorphisms for which the volume of symmetric difference is invariant.

Theorem 2.5.4. *The group* UDif^k *is the set of all elements of* Dif^k *for which* $(\mathcal{K}(\mathbb{R}^k), \mathsf{s})$ *is invariant.*

Proof. It follows from Equation 2.36 and Theorem 2.3.1 that $(\mathcal{K}(\mathbb{R}^k), \mathsf{s})$ is invariant for all elements of UDif^k. It must still be shown that UDif^k is the set of *all* diffeomorphisms under which $(\mathcal{K}(\mathbb{R}^k), \mathsf{s})$ is invariant. Suppose that g is a diffeomorphism that is not a member of UDif^k. Then there is a point x in $\mathbb{R}^k$ such that $\delta_g(x)$ is not equal to 1. Assume $\delta_g(x) < 1$, since the case $\delta_g(x) > 1$ is analogous. Because $x \mapsto \delta_g(x)$ is a continuous function from $\mathbb{R}^k$ to $\mathbb{R}$, an open neighbourhood of x is given by

$$U = \{\, y \in \mathbb{R}^k \mid \delta_g(y) < \frac{1}{2}(\delta_g(x) + 1) \,\}. \tag{2.38}$$

Choose A and B in $\mathcal{K}(\mathbb{R}^k)$ that are disjoint and that lie within U. This gives

$$\begin{aligned}\mathsf{s}(g(A), g(B)) &= \operatorname{vol}(g(A \cup B)) \\ &\leq \frac{1}{2}(\delta_g(x) + 1) \operatorname{vol}(A \cup B) \\ &< \operatorname{vol}(A \cup B) \\ &= \mathsf{s}(A, B).\end{aligned}$$

□

The group of ratio-of-volume preserving diffeomorphisms CDif^k, defined in Appendix B, is the set of all diffeomorphisms for which the Jacobi determinant is constant.

Theorem 2.5.5. *The group* CDif^k *is the set of all diffeomorphisms from* $\mathbb{R}^k$ *onto* $\mathbb{R}^k$ *under which* $(\mathcal{K}^+(\mathbb{R}^k), \mathsf{s}^*)$ *is invariant.*

Proof. First, it is shown that $(\mathcal{K}^+(\mathbb{R}^k), \mathsf{s}^*)$ is invariant for all elements of CDif^k. This follows from Equation 2.37 by Theorem 2.3.4.

Second, it is shown that CDif^k is the set of *all* diffeomorphisms from $\mathbb{R}^k$ onto $\mathbb{R}^k$ for which s^* is invariant. Let g be an element of Dif^k that is not in CDif^k. Choose points x and y in $\mathbb{R}^k$ such that $\delta_g(x) > \delta_g(y)$. Choose $\epsilon > 0$ strictly smaller than $\frac{1}{2}(\delta_g(x) - \delta_g(y))$. By continuity of $x \mapsto \delta_g(x)$ from $\mathbb{R}^k$ to $\mathbb{R}$, open neighbourhoods of x and y are given by

$$\begin{aligned}U &= \{\, z \in \mathbb{R}^k \mid \delta_g(z) \in (\delta_g(x) - \epsilon, \delta_g(x) + \epsilon) \,\}, \\ V &= \{\, z \in \mathbb{R}^k \mid \delta_g(z) \in (\delta_g(y) - \epsilon, \delta_g(y) + \epsilon) \,\}.\end{aligned}$$

Choose closed balls S and T (with nonzero radius) lying in U and V, respectively. Choosing $A = S_1$ and $B = S_1 \cup S_2$ gives

$$\mathsf{s}^*(g(A), g(B)) = \frac{\operatorname{vol}(g(S_2))}{\operatorname{vol}(g(S_1 \cup S_2))} \neq \frac{\operatorname{vol}(S_2)}{\operatorname{vol}(S_1 \cup S_2)} = \mathsf{s}^*(A, B).$$

□

Axiom satisfaction

Because the topology induced on $\mathcal{K}(\mathbb{R}^k)$ by s is not coarser than the topology induced on $\mathcal{K}(\mathbb{R}^k)$ by h, the robustness results known for the Hausdorff metric do not automatically hold for s. Below, it is shown that $(\mathbb{R}^k, \mathcal{K}(\mathbb{R}^k), \mathsf{s})$ satisfies Axioms 2.2.1, 2.2.2, 2.2.3, and 2.2.4. Since s and s^* induce equivalent topologies on $\mathcal{K}^+(\mathbb{R}^k)$, these axioms are also satisfied by $(\mathbb{R}^k, \mathcal{K}^+(\mathbb{R}^k), \mathsf{s}^*)$.

The following lemma is instrumental in proving the robustness axioms.

Lemma 2.5.6. *For each compact subset K of $\mathbb{R}^k$ and each $\epsilon > \mathrm{vol}\,(K)$, there is an open neighbourhood U of K such that $\mathrm{vol}\,(U - K) < \epsilon$.*

Proof. This follows directly from the fact that compact subsets of $\mathbb{R}^k$ are measurable [102, Theorem 13.7, page 111]. □

Theorem 2.5.7. *The pseudometric pattern space $(\mathbb{R}^k, \mathcal{K}(\mathbb{R}^k), \mathsf{s})$ is deformation robust in the sense of Axiom 2.2.1.*

Proof. Let A be in $\mathcal{K}(\mathbb{R}^k)$, and let $\epsilon > 0$. Using Lemma 2.5.6, choose an open neighbourhood U of $\mathrm{Bd}\,(A)$ satisfying $\mathrm{vol}\,(U) < \epsilon$. An open neighbourhood of the identity transformation is given by

$$I = \{\, t \in \mathrm{Dif}^k \mid\ t^{-1}(A - U) \subseteq \mathrm{Int}\,(A) \text{ and } t(A) \subseteq A \cup U\ \}. \tag{2.39}$$

The set I is open in the relative compact-open topology because $A - U$ is compact and $A \cup U$ is open. Assume that t is in I. From the definition of I, it follows that $A - U \subseteq t(A) \subseteq A \cup U$. This implies that $t(A) \triangle A \subseteq U$. Taking volumes over both sides of this set inclusion results in $\mathsf{s}(A, t(A)) \leq \mathrm{vol}\,(U) < \epsilon$. □

Theorem 2.5.8. *The pseudometric pattern space $(\mathbb{R}^k, \mathcal{K}(\mathbb{R}^k), \mathsf{s})$ is blur robust in the sense of Axiom 2.2.2.*

Proof. Let A be in $\mathcal{K}(\mathbb{R}^k)$, and let $\epsilon > 0$, Using Lemma 2.5.6, choose some open neighbourhood U of $\mathrm{Bd}\,(A)$ with volume smaller than ϵ. Clearly, $\mathsf{s}(A, B) < \epsilon$ for each $B \in \mathcal{S}^k$ satisfying $B - U = A - U$. □

Theorem 2.5.9. *The pseudometric pattern space $(\mathbb{R}^k, \mathcal{K}(\mathbb{R}^k), \mathsf{s})$ is crack robust in the sense of Axiom 2.2.3.*

Proof. Let A be in $\mathcal{K}(\mathbb{R}^k)$, let R be a crack of A, and let $\epsilon > 0$. Using Lemma 2.5.6, choose an open neighbourhood U of R with $\mathrm{vol}\,(U) < \epsilon$. For each B in $\mathcal{K}(\mathbb{R}^k)$, $B - U = A - U$ implies $\mathsf{s}(A, B) < \epsilon$. □

Theorem 2.5.10. *The pseudometric pattern space $(\mathbb{R}^k, \mathcal{K}(\mathbb{R}^k), \mathsf{s})$ is noise robust in the sense of Axiom 2.2.4.*

Proof. Let A be in $\mathcal{K}(\mathbb{R}^k)$, let x in $\mathbb{R}^k$, and let $\epsilon > 0$. Choose an open neighbourhood U of x with volume smaller than ϵ. Now $A - U = B - U$ implies $\mathsf{s}(A, B) < \epsilon$ for all B in $\mathcal{K}(\mathbb{R}^k)$. □

Corollary 2.5.11. *The pseudometric pattern space $(\mathbb{R}^k, \mathcal{K}^+(\mathbb{R}^k), \mathsf{s}^*)$ is deformation, blur, crack and noise robust in the sense of Axioms 2.2.1, 2.2.2, 2.2.3, and 2.2.4.*

Proof. The previous theorems show that each the mentioned axioms holds for $(\mathbb{R}^k, \mathcal{K}(\mathbb{R}^k), \mathsf{s})$. By Theorem 2.2.3, the axioms still hold after $\mathcal{K}(\mathbb{R}^k)$ is restricted to $\mathcal{K}^+(\mathbb{R}^k)$. It follows from Corollary 2.5.3 that the topology induced by s is finer than the topology induced by s^*. By Theorem 2.2.1, each of the axioms holds for the pseudometric space $(\mathbb{R}^k, \mathcal{K}^+(\mathbb{R}^k), \mathsf{s}^*)$. □

Variants

Below, various alternative normalisations of the volume of symmetric difference are discussed. First, a number of normalisations that are not pseudometrics are mentioned. Finally, an alternative normalisation that is based on the construction method of Section 2.3 is presented.

The normalisation of the volume of symmetric difference presented in this section divides by the volume of the union of the two patterns. However, other normalisations are possible. Alternatives are division by the sum of the volumes and the maximum of the volumes. These alternatives result in affine invariant similarity measures. Below it is demonstrated that the mentioned alternative normalisations do not satisfy the triangle inequality.

Define a normalisation version of s on $\mathcal{K}^+(\mathbb{R}^k)$ as

$$d(A, B) = \mathrm{vol}\,(A \triangle B) / \left(\mathrm{vol}\,(A) + \mathrm{vol}\,(B)\right). \tag{2.40}$$

A one-dimensional example shows that d does not satisfy the triangle inequality. Choose $A = [0, 1]$, $B = [0, 2]$ and $C = [1, 2]$. Then, $d(A, C) = 1 > 2/3 = d(A, B) + d(B, C)$.

The other normalisation is defined on $\mathcal{K}^+(\mathbb{R}^k)$ as

$$d(A, B) = \mathrm{vol}\,(A \triangle B) / \max\{\,\mathrm{vol}\,(A), \mathrm{vol}\,(B)\,\}. \tag{2.41}$$

Again, a one-dimensional example shows that the triangle inequality does not hold. Choose $A = [0, 1]$, $B = [0, 3]$ and $C = [2, 3]$. Then, $d(A, C) = 2 > 4/3 = d(A, B) + d(B, C)$.

The normalised volume of symmetric difference s^* can be expressed in terms of the construction method of Section 2.3. Here, the normalised function distance $\boldsymbol{l}_1^*$ is used to compare the characteristic functions $\mathbf{1}_A$ and $\mathbf{1}_B$ corresponding to patterns. By Theorem 2.3.4, s^* is invariant for CDif^k. However, it is also possible to form a pseudometric invariant for CDif^k using the function distance $\boldsymbol{l}_1$ by applying Theorem 2.3.1. The definition is as follows. Let $\boldsymbol{n}_P$ denote $\mathrm{vol}\,(P)^{-1}\mathbf{1}_P$. Then, the absolute difference (of normalised indicators) a is given by

$$\mathsf{a}(A, B) = \boldsymbol{l}_1(\boldsymbol{n}_A, \boldsymbol{n}_B). \tag{2.42}$$

Together with Veltkamp I proved that $(\mathbb{R}^k, \mathcal{S}'(\mathbb{R}^k), \mathsf{a})$ is deformation, blur, crack and noise robust [70].

2.6 Reflection visibility based distances

The visibility from a particular viewpoint gives a local description of a pattern. The visibilities from all possible viewpoints give a complete representation of the pattern. For this reason, the notion of visibility can be useful in pattern matching. Visibility is defined in terms of affine geometry, the concept does not depend on Euclidean distances. Therefore, visibility can be used as a tool for affine invariant shape recognition and affine pattern matching.

In this section, strong type of visibility is used to define a new similarity measure which is called the reflection visibility distance. In addition, a normalised version of this distance is given. The reflection visibility distance and its normalised version are constructed using the methods of Section 2.3.

In the previous section, it was shown that the volume of symmetric difference has better robustness and invariance properties than the Hausdorff metric. However, the volume of symmetric difference is only suitable for matching regions (solid sets). The reflection visibility distance, defined below, has similar advantages over the Hausdorff metric. This similarity measure is defined for "boundary patterns", that is, $(k-1)$-dimensional patterns in $\mathbb{R}^k$. Examples of such patterns are finite unions of line segments in the plane. It is shown that the reflection visibility distance satisfies the four robustness axioms and that its normalised version is invariant for affine transformations.

Definition

In Section 2.3, a construction method for (pseudometric) similarity measures was given. It was shown that pseudometrics on patterns can be formed by assigning a nonnegative real valued function to each pattern. These functions can then be compared using the pseudometrics $\boldsymbol{l}_p$ or $\boldsymbol{l}_p^*$. Below, the notion of visibility is used to assign a function for each pattern. When the normalised distance $\boldsymbol{l}_p^*$ is used to compare the functions, an affine invariant pseudometric is the result. Furthermore, the resulting pseudometric is deformation, blur, crack, and noise robust.

The definitions of the new similarity measures depend on the geometric constructions given below. For points x and y in $\mathbb{R}^k$, the notation $\mathrm{Seg}(x, y)$ stands for the open line segment connecting x and y. In set notation, the open line segment connecting x and y is given by

$$\mathrm{Seg}(x, y) = \{\, \alpha x + \beta y \in \mathbb{R}^k \mid \ \alpha + \beta = 1 \text{ and } \alpha, \beta > 0 \ \}. \tag{2.43}$$

Observe that an open line segment need not be an open subset of $\mathbb{R}^k$ in the topological sense. A point $y \in \mathbb{R}^k$ is visible from a point $x \in \mathbb{R}^k$ relative to a subset P of $\mathbb{R}^k$ if $\mathrm{Seg}(x, y)$ is disjoint with $\mathrm{Bd}\,(P)$. The set of all points of $\mathbb{R}^k$

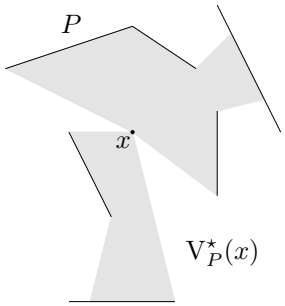


Figure 2.6: Visibility star

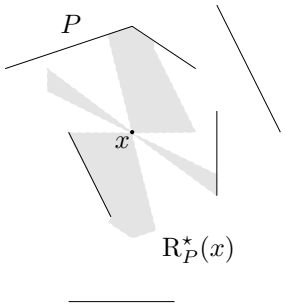


Figure 2.7: Reflection visibility star

that are visible from x relative to P is written as

$$\mathrm{V}_P(x) = \{\, y \in \mathbb{R}^k \mid \mathrm{Seg}(x, y) \cap \mathrm{Bd}\,(P) = \varnothing \,\}. \tag{2.44}$$

For each point x of $\mathbb{R}^k$, the visibility star of x (relative to P) is defined as the union of all $\mathrm{Seg}(x, y)$ for $y \in \mathrm{Bd}\,(P)$ that are visible from x. This is written formally as

$$\mathrm{V}^\star_P(x) = \bigcup_{y \in \mathrm{V}_P(x) \cap \mathrm{Bd}(P)} \mathrm{Seg}(x, y). \tag{2.45}$$

Figure 2.6 shows an example of a visibility star relative to a set A that is a finite union of line segments in the plane. The visibility star for the viewpoint x relative to A is the light grey region.

The reflection of a subset S of $\mathbb{R}^k$ about a point x of $\mathbb{R}^k$ is the set consisting of all points y in $\mathbb{R}^k$ such that $x - (y - x)$ is in S. Formulated otherwise, the reflection of S in x consists of all points $x + v$ such that $x - v$ is in S. The reflection visibility star, denoted $\mathrm{R}^\star_P(x)$, is the union of all open line segments with midpoint x that are contained in $\mathrm{V}^\star_P(x)$. In formula that is

$$\mathrm{R}^\star_P(x) = \{\, x + v \in \mathbb{R}^k \mid \mathrm{Seg}(x + v, x - v) \subseteq \mathrm{V}^\star_P(x) \,\}. \tag{2.46}$$

Observe that $\mathrm{R}^\star_P(x)$ equals $\varnothing$ for x in $\mathrm{Bd}\,(P)$. For x not in $\mathrm{Bd}\,(P)$, $\mathrm{R}^\star_P(x)$ equals the intersection of $\mathrm{V}^\star_P(x)$ with the reflection of $\mathrm{V}^\star_P(x)$ in x. In addition, observe that $\mathrm{R}^\star_P(x)$ is empty outside the convex hull of P. Figure 2.7 shows the reflection visibility star for a viewpoint x relative to the set P also used in Figure 2.6.

The reflection visibility surface of P, denoted $\boldsymbol{r}_P$, is a function from $\mathbb{R}^k$ to $\mathbb{R}$ that equals the volume of the reflection visibility star in each point. Formally, the reflection visibility surface is given by

$$\boldsymbol{r}_P(x) = \mathrm{vol}\,(\mathrm{R}^\star_P(x)). \tag{2.47}$$

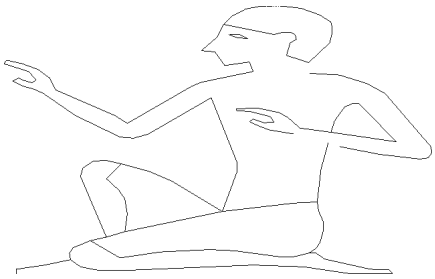


Figure 2.8: The pattern P

Figure 2.9: The function $\boldsymbol{r}_P$

Figure 2.8 shows a two-dimensional pattern P consisting of a finite number of line segments. Figure 2.9 shows the corresponding function $\boldsymbol{r}_P$ from $\mathbb{R}^2$ to $\mathbb{R}$ represented as a grey-scale image in which black corresponds with value 0 and white corresponds to the maximum value. The example pattern is a hieroglyphic with code "A1" from the hieroglyphica sign list [62].

Below, the reflection visibility surfaces are used to form metrics and pseudometrics by means of the construction methods of Section 2.3. First, the properties of the reflection visibility surfaces must be studied. The discussion is simplified by considering a subcollection of $\mathcal{K}'(\mathbb{R}^k)$. The definition of this subcollection is based on the following construct: an l-simplex in $\mathbb{R}^k$ is the (closed) convex hull of a given set of $l+1$ independent points in $\mathbb{R}^k$. The latter points are the vertices of the l-simplex. Using these constructs, a collection of subsets of $\mathbb{R}^k$ is defined as follows.

Definition 2.6.1. The collection $\mathcal{M}(\mathbb{R}^k)$ consists of all elements P of $\mathcal{K}(\mathbb{R}^k)$ such that $\mathrm{Bd}\,(P)$ can be expressed as the union of a finite collection of $(k-1)$-simplices having disjoint interiors such that the set of all vertices is independent and has size at least $k+1$.

It will be shown that distances based on the reflection visibility surfaces are pseudometrics on the collection $\mathcal{M}(\mathbb{R}^k)$. In addition, it will be shown that such distances form metrics on the subcollection $\mathcal{T}(\mathbb{R}^k)$ of $\mathcal{M}(\mathbb{R}^k)$ given next.

Definition 2.6.2. The collection $\mathcal{T}(\mathbb{R}^k)$ consists of all P in $\mathcal{M}(\mathbb{R}^k)$ for which $\mathrm{Int}\,(P) = \varnothing$.

At this point, the properties of the reflection visibility surfaces can be studied. The following lemma is helpful in proving that for each P in $\mathcal{M}(\mathbb{R}^k)$, the corresponding function $\boldsymbol{r}_P$ is continuous "almost everywhere".

Lemma 2.6.1. *Measurable subsets S and T of $\mathbb{R}^k$ with finite volume satisfy the inequality*

$$|\mathrm{vol}\,(S) - \mathrm{vol}\,(T)| \leq \mathrm{vol}\,(S \triangle T). \tag{2.48}$$

Proof. Observe that $\mathrm{vol}\,(S) - \mathrm{vol}\,(T) \leq \mathrm{vol}\,(S) - \mathrm{vol}\,(S \cap T) = \mathrm{vol}\,(S - T)$. Symmetrically, $\mathrm{vol}\,(T) - \mathrm{vol}\,(S) \leq \mathrm{vol}\,(T - S)$. Thus

$$\begin{aligned}
|\mathrm{vol}\,(S) - \mathrm{vol}\,(T)| &\leq \max\{\,\mathrm{vol}\,(S) - \mathrm{vol}\,(T), \mathrm{vol}\,(T) - \mathrm{vol}\,(S)\,\} \\
&= \max\{\,\mathrm{vol}\,(S - T), \mathrm{vol}\,(T - S)\,\} \\
&\leq \mathrm{vol}\,(S - T) + \mathrm{vol}\,(T - S) \\
&= \mathrm{vol}\,((S - T) \cup (T - S)) \\
&= \mathrm{vol}\,(S \triangle T).
\end{aligned}$$

□

The next lemma says that for elements P of $\mathcal{M}(\mathbb{R}^k)$, the function $\boldsymbol{r}_P$ is continuous at each point of the domain except, possibly, for a $(k-2)$-dimensional subset of the boundary. In the proof, the line through the points x and y is denoted $\mathrm{Line}(x, y)$. Furthermore, the Euclidean open ball centred at x having radius ϵ is denoted by $\mathrm{B}(x, \epsilon)$.

Lemma 2.6.2. *For each P in $\mathcal{M}(\mathbb{R}^k)$, there is a subset Q of $\mathrm{Bd}\,(P)$ that is a finite union of $(k-2)$-simplices such that the restriction $\boldsymbol{r}_P|_{\mathbb{R}^k - Q}$ is continuous.*

Proof. Let P be an element of $\mathcal{M}(\mathbb{R}^k)$. It is assumed that S is an open ball containing P. The boundary of P can be expressed as the union of a finite collection $\mathcal{D}$ of $(k-1)$-simplices. Let $\mathcal{Q}$ be the collection of $(k-2)$-simplices that are subsimplices of elements of $\mathcal{D}$; let Q be the union of $\mathcal{Q}$. Let V be set of all vertices of $\mathcal{D}$; let C be $\mathbb{R}^k - V$. It is sufficient to prove that the function $\boldsymbol{r}_P$ is continuous in each point of C.

First, continuity for x in $\mathbb{R}^k - \mathrm{Bd}\,(P)$ is shown. The number $\delta > 0$ is chosen so small that $\mathrm{B}(x, \delta)$ is disjoint with the boundary of P. Define $W_P(\delta)$ as the union of all lines that pass through $\mathrm{B}(x, \delta)$ and Q. In set notation, $W_P(\delta)$ is given by

$$W_P(\delta) = \bigcup_{y \in \mathrm{B}(x,\delta)} \bigcup_{z \in Q} \mathrm{Line}(y, z). \tag{2.49}$$

The set $W_P(\delta) \cap S$ contains the symmetric difference of $\mathrm{R}^\star_P(x)$ and $\mathrm{R}^\star_P(y)$ if $\|x - y\| < \delta$. The volume of $W_P(\delta) \cap S$ can be made arbitrarily small by choosing $\delta > 0$ sufficiently small. In particular, $\delta > 0$ may be chosen so that the volume of $W_P(\delta) \cap S$ is less than ϵ. Lemma 2.6.1 leads to the inequality

$$\begin{aligned}
|\boldsymbol{r}_P(x) - \boldsymbol{r}_P(y)| &\leq \mathrm{vol}\,(\mathrm{R}^\star_P(x) \triangle \mathrm{R}^\star_P(y)) \\
&\leq \mathrm{vol}\,(W_P(\delta) \cap S) \\
&< \epsilon.
\end{aligned} \tag{2.50}$$

This finishes the first part of the proof.

Second, continuity for x in $\mathrm{Bd}\,(P) - V$ is shown. Such an x is contained in a $(k-1)$-simplex D in $\mathcal{D}$. Furthermore, x is not a member of some $(k-2)$ subsimplex of D. Observe that $\boldsymbol{r}_P(x)$ is zero. Define

$$U(y) = \bigcup_{y+v \in D} \mathrm{Seg}(y-v, y+v) \cup \bigcup \{\, \mathrm{Line}(y,z) \mid \mathrm{Line}(y,z) \cap D = \varnothing \,\}. \quad (2.51)$$

The set $U(y) \cap S$ contains $\mathrm{R}^\star_P(y)$ for each y. The volume of $U(y) \cap S$ can be made arbitrarily small by choosing y sufficiently close to x. Therefore, $\delta > 0$ can be chosen so that

$$\begin{aligned} |\boldsymbol{r}_P(x) - \boldsymbol{r}_P(y))| &= \boldsymbol{r}_P(y) \\ &= \mathrm{vol}\,(\mathrm{R}^\star_P(y)) \\ &\leq \mathrm{vol}\,(U(y) \cap S) \\ &< \epsilon. \end{aligned} \quad (2.52)$$

□

The next lemma is instrumental in showing positivity of reflection visibility based distances on the collection $\mathcal{T}(\mathbb{R}^k)$.

Lemma 2.6.3. *If A and B are in $\mathcal{T}(\mathbb{R}^k)$ and $A \neq B$, then there is a point x in $\mathbb{R}^k - (A \cup B)$ such that $\boldsymbol{r}_A(x) \neq \boldsymbol{r}_B(x)$.*

Proof. Let A and B be distinct elements of $\mathcal{T}(\mathbb{R}^k)$. Using the definition of $\mathcal{T}(\mathbb{R}^k)$, A (B) is expressed as a union of a finite collection $\mathcal{A}$ ($\mathcal{B}$) of $(k-1)$-simplices with disjoint interiors. Let $V(\mathcal{A})$ and $V(\mathcal{B})$ denote the vertex sets corresponding to $\mathcal{A}$ and $\mathcal{B}$, respectively.

Without loss of generality, it is assumed that there is a point a in $A - B$ that is not in $V(\mathcal{A})$. Observe that one of the following three cases must hold:

1. There are no x and y in B such that $a \in \mathrm{Seg}(x, y)$.
2. There are x and y in the interiors of $(k-1)$-simplices of $\mathcal{B}$ such that $a \in \mathrm{Seg}(x, y)$.
3. Neither Case 1 nor Case 2 applies.

Case 1: Let C be the union of all open line segments connecting each two points of B. Because a is not in B, there is an open ball U centred at a that is disjoint with C. It follows from the definition of $\mathcal{T}(\mathbb{R}^k)$ that there is a point $a' \in A$ that is not a vertex of $\mathcal{A}$ such that $\mathrm{Seg}(a, a')$ is disjoint with $A \cup B$. If x is chosen in $\mathrm{Seg}(a, a') - V(\mathcal{A})$, then $\boldsymbol{r}_A(x) > 0$ while $\boldsymbol{r}_B(x) = 0$.

Case 2: This case is easy; $\boldsymbol{r}_A(a) = 0$ while $\boldsymbol{r}_B(a) > 0$.

Case 3: Let S be the $(k-1)$-simplex in $\mathcal{A}$ that contains the point a. Depending on the $(k-1)$-dimensional hyperplane H that contains S, one of two subcases applies.

Subcase 1: The hyperplane H does not contains all open line segments between elements of B that contain a. Here the point S must contain a point a' that is in $A - B$ such that Case 2 applies if a' is substituted for a.

Subcase 2: The hyperplane H contains all open line segments between elements of B that contain a. There is a vertex v in $V(\mathcal{A})$ that is visible from a relative to A. If $\mathrm{Seg}(a, v)$ contains a point b of B it is immediately clear that $\boldsymbol{r}_A(b) > 0$ while $\boldsymbol{r}_B(b) = 0$. Therefore assume that $\mathrm{Seg}(a, v)$ contains no point of B. Without loss of generality, it is assumed that the midpoint p of the segment $\mathrm{Seg}(a, v)$ is not the midpoint of two elements of B. The point p is a discontinuity in the total derivative of $\boldsymbol{r}_A$ since it designates the point that the vertex v appears or disappears from the reflection visibility star $\mathrm{R}^{\star}_A(x)$. On the other hand, for $\boldsymbol{r}_B$ the total derivative is continuous in an open neighbourhood U of p, where U is disjoint with $A \cup B$. This means that $\boldsymbol{r}_A$ and $\boldsymbol{r}_B$ must differ in some point x of U. □

At this point, the construction method of Section 2.3 can be applied. For elements of $\mathcal{M}(\mathbb{R}^k)$, the reflection visibility surfaces are continuous almost everywhere by Lemma 2.6.2. This implies that these surfaces are Lebesgue integrable, that is, $\boldsymbol{r}_A$ is in $\boldsymbol{I}^k_1$ for each A in $\mathcal{M}(\mathbb{R}^k)$. The reflection visibility distance and its normalised version are defined as follows.

Definition 2.6.3. The *reflection visibility distance* on $\mathcal{M}(\mathbb{R}^k)$ is given by

$$\mathsf{r}(A, B) = \boldsymbol{l}_1(\boldsymbol{r}_A, \boldsymbol{r}_B). \tag{2.53}$$

The normalised reflection visibility distance on $\mathcal{M}(\mathbb{R}^k)$ is given by

$$\mathsf{r}^*(A, B) = \boldsymbol{l}^*_1(\boldsymbol{r}_A, \boldsymbol{r}_B). \tag{2.54}$$

Theorem 2.6.4. *Both* r *and* r^* *are pseudometrics on* $\mathcal{M}(\mathbb{R}^k)$ *and metrics on* $\mathcal{T}(\mathbb{R}^k)$.

Proof. The fact that $\boldsymbol{l}_1$ is a pseudometric implies that r is a pseudometric. Similarly, r^* is a pseudometric because $\boldsymbol{l}^*_1$ is a pseudometric (Theorem 2.3.3).

Now it is shown that r and r^* are metrics on $\mathcal{T}(\mathbb{R}^k)$. First, positivity of r on $\mathcal{T}(\mathbb{R}^k)$ is shown. Let A and B be elements of $\mathcal{T}(\mathbb{R}^k)$ such that $A \neq B$. Consider the function f from $\mathbb{R}^k$ to $\mathbb{R}$ given by $f(x) = |\boldsymbol{r}_A(x) - \boldsymbol{r}_B(x)|$. It is sufficient to show that $\int_{\mathbb{R}^k} f(x)\, dx > 0$.

By Lemma 2.6.3, there is a point x in $\mathbb{R}^k - (A \cup B)$ such that $f(x) > 0$. It follows from Lemma 2.6.2 that the restriction $f|_{\mathbb{R}^k - (A \cup B)}$ is continuous. The

inverse image of the open ray $(f(x)/2, \infty)$ defines an open subset of $\mathbb{R}^k - (A \cup B)$ as in the equation

$$U = \{\, y \in \mathbb{R}^k - (A \cup B) \mid f(x)/2 < f(y) \,\}. \tag{2.55}$$

Because $A \cup B$ is closed, U is an open subset of $\mathbb{R}^k$. Because U is nonempty and open it has a positive volume $\gamma > 0$. This gives

$$\int_{\mathbb{R}^k} f(y)\, dy \geq \int_U f(y)\, dy \geq \int_U f(x)/2\, dy \geq \gamma f(x)/2 > 0. \tag{2.56}$$

This also implies positivity of r^* on $\mathcal{T}(\mathbb{R}^k)$. □

Invariance

It is a well known fact that affine transformations map line segments to line segments. In fact, for each affine transformation g, the visibility star of $g(x)$ relative to $g(A)$ is the image under g of the visibility star of x relative to A. This observation leads to the following theorem. Recall that UAf^k is the group of volume-preserving affine transformations.

Theorem 2.6.5. *The pseudometric space* $(\mathcal{M}(\mathbb{R}^k), \mathsf{r})$ *is invariant for* UAf^k. *The pseudometric space* $(\mathcal{M}(\mathbb{R}^k), \mathsf{r}^*)$ *is invariant for* Af^k.

Proof. First, an important property of the reflection visibility surface is shown. Let P be a member of $\mathcal{M}(\mathbb{R}^k)$. Let g be a member of Af^k. By definition of the reflection visibility star,

$$\mathrm{R}^\star_{g(P)}(g(x)) = g(\mathrm{R}^\star_P(x)). \tag{2.57}$$

Taking the volumes of both sides of the equation results in

$$\mathrm{vol}\left(\mathrm{R}^\star_{g(P)}(g(x))\right) = \delta_g(x)\, \mathrm{vol}\,(\mathrm{R}^\star_P(x)). \tag{2.58}$$

Substituting this in Equation 2.47 gives

$$\boldsymbol{r}_{g(P)}(g(x)) = \delta_g(x) \boldsymbol{r}_P(x). \tag{2.59}$$

By this equation, $\{\, \boldsymbol{r}_P \,\}_{P \in \mathcal{P}}$ satisfies the condition of Theorem 2.3.4. This gives

$$\mathsf{r}^*(g(A), g(B)) = \mathsf{r}^*(A, B). \tag{2.60}$$

Assume g is in UAf^k. Then, because $\delta_g(x) = 1$ for all x in $\mathbb{R}^k$, the condition of Theorem 2.3.1 is satisfied, giving that

$$\mathsf{r}(g(A), g(B)) = \mathsf{r}(A, B). \tag{2.61}$$

□

Axiom satisfaction

Below, it is shown that the reflection visibility distance and its normalised variant satisfy the robustness axioms. The proofs for the axioms are simplified by the following lemma which establishes a relation between the volume of symmetric difference of reflection visibility stars and visibility stars.

Lemma 2.6.6. *Let x be a point in $\mathbb{R}^k$ and let S and T be bounded subsets of $\mathbb{R}^k$. let S' and T' be the reflections of S and T about x, respectively. If $S'' = S \cap S'$ and $T'' = T \cap T'$, then*

$$\operatorname{vol}(S'' \triangle T'') \leq 2 \operatorname{vol}(S \triangle T). \tag{2.62}$$

Proof. It is sufficient to show that

$$\operatorname{vol}(S'' - T'') \leq 2 \operatorname{vol}(S - T).$$

The claimed inequality is obtained by "summing" this equation and the one in which S and T are exchanged. Using elementary set theory it is found that

$$\begin{aligned} S'' - T'' &= S'' - (T \cap T') \\ &= (S'' - T) \cup (S'' - T') \\ &\subseteq (S - T) \cup (S' - T'). \end{aligned} \tag{2.63}$$

Taking volumes over the previous set inclusion results in

$$\begin{aligned} \operatorname{vol}(S'' - T'') &\leq \operatorname{vol}((S - T) \cup (S' - T')) \\ &\leq \operatorname{vol}(S - T) + \operatorname{vol}(S' - T') \\ &= 2 \operatorname{vol}(S - T). \end{aligned} \tag{2.64}$$

□

Define the pointwise difference of the visibility stars relative to A and B by

$$\boldsymbol{u}_{A,B}(x) = \operatorname{vol}(\mathrm{V}^\star_A(x) \triangle \mathrm{V}^\star_B(x)). \tag{2.65}$$

By Lemmas 2.6.1 and 2.6.6, $2\boldsymbol{u}_{A,B}(x)$ is an upper bound for the expression $|\boldsymbol{r}_A(x) - \boldsymbol{r}_B(x)|$ for each x in $\mathbb{R}^k$. Observe that for each pattern A the reflection visibility function is zero in points outside each closed ball that contains A. Thus, if S is a closed ball containing both A and B, then

$$\mathsf{r}(A, B) \leq 2 \int_S \boldsymbol{u}_{A,B}(x)\, dx. \tag{2.66}$$

Below, four lemmas are given. Together with the latter inequality the lemmas imply the four distinct types of robustness for the reflection visibility distance and its normalised variant.

The following notation is used. For each subset P of $\mathbb{R}^k$, the Euclidean ϵ neighbourhood of P, denoted $\mathrm{N}(P,\epsilon)$ equals the union of all Euclidean open balls $\mathrm{B}(x,\epsilon)$ for $x \in P$. For each compact subset A of $\mathbb{R}^k$, the smallest closed ball containing A is denoted by $\mathrm{C}(A)$.

In the proof of the next lemma, the open ray emanating from x (not including x) and passing through y is denoted by $\mathrm{Ray}(x,y)$.

Lemma 2.6.7. *For each A in $\mathcal{M}(\mathbb{R}^k)$ and each $\epsilon > 0$, a $\delta > 0$ exists such that*

$$\int_{\mathrm{N}(\mathrm{C}(A),\delta)} \boldsymbol{u}_{A,t(A)}(x)\,dx < \epsilon \tag{2.67}$$

for all t in $\mathrm{Clos}(\mathbb{R}^k, \mathcal{M}(\mathbb{R}^k))$ satisfying $\max_{a\in A} \|t(a) - a\| < \delta$.

Proof. The proof uses the fact that A is in $\mathcal{M}(\mathbb{R}^k)$. Express the boundary of A as the union of a finite collection $\mathcal{D}$ of $(k-1)$-simplices. The collection $\mathcal{Q}$ consists of all $(k-2)$-simplices that are subsimplices of some element of $\mathcal{D}$. Let Q be the union of $\mathcal{Q}$.

Define $\mathcal{B}(A,\delta)$ as the collection of $t(A)$ for all t in $\mathrm{Clos}(\mathbb{R}^k, \mathcal{M}(\mathbb{R}^k))$ satisfying $\max_{a\in A} \|t(a) - a\| < \delta$.

Let $\epsilon > 0$. Express ϵ as a sum of positive numbers ϵ_1 and ϵ_2. Below, the integrals over two disjoint subsets of the domain are bounded by by ϵ_1 and ϵ_2.

First, the integral over a set of points in a neighbourhood of A is bounded so that it becomes less than ϵ_1. Choose $\delta_1 > 0$ so that for all B in $\mathcal{B}(A,\delta_1)$,

$$\int_{\mathrm{N}(\mathrm{Bd}(A),\delta_1)} \boldsymbol{u}_{A,B}(x)\,dx \le \mathrm{vol}\,(\mathrm{N}(\mathrm{Bd}\,(A),\delta_1))\,\mathrm{vol}\,(\mathrm{N}(\mathrm{C}(A),\delta_1)) < \epsilon_1.$$

Second, the integral over the set of points outside $\mathrm{N}(\mathrm{Bd}\,(A),\delta_1)$ is bounded by ϵ_2. Observe that the integral over the complement of $\mathrm{N}(\mathrm{C}(A),\delta_1)$ is zero. Therefore, it is sufficient to bound the integral over $\mathrm{N}(\mathrm{C}(A),\delta_1) - \mathrm{N}(\mathrm{Bd}\,(A),\delta_1)$ by ϵ_2. This holds if the value $\boldsymbol{u}_{A,B}(x)$ over all x in $\mathrm{N}(\mathrm{C}(A),\delta_1) - \mathrm{N}(A,\delta_1)$ is bounded from above by

$$\gamma = \epsilon_2 / \mathrm{vol}\,(\mathrm{N}(\mathrm{C}(A),\delta_1)). \tag{2.68}$$

Express γ as a sum of positive numbers γ_1 and γ_2. For each $\delta > 0$, define $W_x(\delta)$ as the union of all rays emanating from x that intersect the δ-neighbourhood of Q. In formula, that is

$$W_x(\delta) = \bigcup \{\, \mathrm{Ray}(x,y) \mid \mathrm{Ray}(x,y) \cap \mathrm{N}(Q,\delta) \neq \varnothing \,\}. \tag{2.69}$$

Since Q is a finite union of $(k-2)$-simplices, $\delta_2 > 0$ may be chosen so that for all x in $\mathrm{N}(\mathrm{C}(A),\delta_1) - \mathrm{N}(\mathrm{Bd}\,(A),\delta_1)$ and all B in $\mathcal{B}(A,\delta_2)$,

$$\mathrm{vol}\,((\mathrm{V}^\star_A(x) \mathbin{\triangle} \mathrm{V}^\star_B(x)) \cap W_x(\delta_2)) \le \mathrm{vol}\,(\mathrm{N}(\mathrm{C}(A),\delta_1) \cap W_x(\delta_2)) < \gamma_1. \tag{2.70}$$

Choose $\delta_3 > 0$ so that for each B in $\mathcal{B}(A, \delta_3)$, and each x in $\mathrm{N}(\mathrm{C}(A), \delta_1) - \mathrm{N}(\mathrm{Bd}\,(A), \delta_1)$,

$$\mathrm{vol}\,((\mathrm{V}^\star_A(x) \triangle \mathrm{V}^\star_B(x)) - W_x(\delta_2)) \leq \mathrm{vol}\,(\mathrm{N}(\mathrm{Bd}\,(A), \delta_3)) < \gamma_2. \tag{2.71}$$

The proof is finished by choosing δ as the minimum of δ_1, δ_2, and δ_3. □

Lemma 2.6.8. *For each A in $\mathcal{M}(\mathbb{R}^k)$ and each $\epsilon > 0$, a $\delta > 0$ exists such that*

$$\int_{\mathrm{N}(\mathrm{C}(A),\delta)} \boldsymbol{u}_{A,B}(x)\, dx < \epsilon \tag{2.72}$$

for all B in $\mathcal{M}(\mathbb{R}^k)$ satisfying $B - \mathrm{N}(\mathrm{Bd}\,(A), \delta) = A - \mathrm{N}(\mathrm{Bd}\,(A), \delta)$ and $\mathrm{Bd}\,(B) \supseteq \mathrm{Bd}\,(A)$.

Proof. Here it is convenient to define $\mathcal{B}(A, \delta)$ as the collection of B in $\mathcal{M}(\mathbb{R}^k)$ for which $B - \mathrm{N}(\mathrm{Bd}\,(A), \delta) = A - \mathrm{N}(\mathrm{Bd}\,(A), \delta)$ and $\mathrm{Bd}\,(B) \supseteq \mathrm{Bd}\,(A)$.

For the remainder, the proof is analogous to the proof of Lemma 2.6.7. □

Lemma 2.6.9. *For each A in $\mathcal{M}(\mathbb{R}^k)$, each crack R of A and each $\epsilon > 0$, a $\delta > 0$ exists such that*

$$\int_{\mathrm{N}(\mathrm{C}(A),\delta)} \boldsymbol{u}_{A,B}(x)\, dx < \epsilon \tag{2.73}$$

for all B in $\mathcal{M}(\mathbb{R}^k)$ satisfying $B - \mathrm{N}(R, \delta) = A - \mathrm{N}(R, \delta)$.

Proof. Let A be in $\mathcal{M}(\mathbb{R}^k)$, and let R be a crack of A. It is convenient to define $\mathcal{B}(A, \delta)$ as the collection of B in $\mathcal{M}(\mathbb{R}^k)$ for which $B - \mathrm{N}(R, \delta) = A - \mathrm{N}(R, \delta)$.

Write $\epsilon > 0$ as a sum of positive numbers ϵ_1 and ϵ_2.

First, the integral is bounded over a neighbourhood of R. Choose $\delta_1 > 0$ so that for all B in $\mathcal{B}(A, \delta_1)$,

$$\int_{\mathrm{N}(\mathrm{Bd}(R),\delta_1)} \boldsymbol{u}_{A,B}(x)\, dx \leq \mathrm{vol}\,(\mathrm{N}(\mathrm{C}(A), \delta_1))\, \mathrm{vol}\,(\mathrm{N}(\mathrm{Bd}\,(R), \delta_1)) < \epsilon_1.$$

Next, the integral outside the neighbourhood of R is bounded. For this purpose define $W_x(\delta)$ as in Equation 2.69, replacing Q with R. The number $\delta_2 > 0$ is chosen so that for all x in $\mathrm{N}(\mathrm{C}(A), \delta_1) - \mathrm{N}(\mathrm{Bd}\,(A), \delta_1)$ and all B in $\mathcal{B}(A, \delta_2)$,

$$\boldsymbol{u}_{A,B}(x) \leq \mathrm{vol}\,(\mathrm{N}(\mathrm{C}(A), \delta_1) \cap W_x(\delta_2)) < \gamma_2. \tag{2.74}$$

The required δ is the minimum of δ_1 and δ_2. □

Lemma 2.6.10. *Let k be at least 2. Then, for each A in $\mathcal{M}(\mathbb{R}^k)$, each x in $\mathbb{R}^k$ and each $\epsilon > 0$, a $\delta > 0$ exists such that*

$$\int_{\mathrm{N}(\mathrm{C}(A),\delta)} \boldsymbol{u}_{A,B}(x)\, dx < \epsilon \tag{2.75}$$

for all B in $\mathcal{K}(\mathbb{R}^k)$ satisfying $B - \mathrm{B}(x,\delta) = A - \mathrm{B}(x,\delta)$.

Proof. The proof is almost the same as that of Lemma 2.6.9. □

From the previous discussion it is evident that the four robustness axioms hold for the reflection visibility distance and its normalised variant.

Theorem 2.6.11. *Both $(\mathbb{R}^k, \mathcal{M}(\mathbb{R}^k), \mathsf{r})$ and $(\mathbb{R}^k, \mathcal{M}(\mathbb{R}^k), \mathsf{r}^*)$ are deformation, blur, and crack robust in the sense of Axioms 2.2.1, 2.2.2 and 2.2.3. If $k \geq 2$, then both spaces are noise robust in the sense of Axiom 2.2.4.*

Variants

The definitions of the reflection visibility distance r and its normalisation r^* given above might seem "cumbersome". Below, these definitions of are justified. It is shown that several simpler alternatives which achieve similar invariance properties do not have the robustness properties that r and r^* have.

The distances r and r^* are invariant for UAf^k and Af^k, respectively. This is stated in Theorem 2.6.5. The definitions of r and r^* are in terms of the reflection visibility surface $\boldsymbol{r}_P(x)$ corresponding to a pattern P. The value of $\boldsymbol{r}_P(x)$ is the volume of the reflection visibility star $\mathrm{R}^\star_P(x)$. The proof of Theorem 2.6.5 is entirely based on the fact that the reflection visibility star for each point satisfies Equation 2.57.

The construction method using the reflection visibility stars is generalised as follows. For each P and each x, a generic set $\mathrm{S}_P(x)$ is defined. Suppose that this definition satisfies the generalisation of Equation 2.57 given by

$$\mathrm{S}_{g(P)}(g(x)) = g(\mathrm{S}_P(x)) \tag{2.76}$$

for each $g \in \mathrm{Af}^k$. The generic sets for P and all x lead to a function $\boldsymbol{n}_P : \mathbb{R}^k \to \mathbb{R}$ given by $\boldsymbol{n}_P = \mathrm{vol}\,(\mathrm{S}_P(x))$. Now, Theorem 2.3.4 ensures that the pseudometric d given by $d(A,B) = \boldsymbol{l}^*_p(\boldsymbol{n}_P, \boldsymbol{n}_P)$ is invariant for Af^k.

Below, various examples of a definition for $\mathrm{S}_P(x)$ that satisfies Equation 2.76 are described. None of these examples results in a pseudometric pattern space that has all the properties of $(\mathbb{R}^k, \mathcal{T}(\mathbb{R}^k), \mathsf{r}^*)$ in terms of positivity, invariance, and robustness. For some of the examples, the corresponding distance is even ill-defined. The reflection visibility star is the only known choice for $\mathrm{S}_P(x)$ that achieves this. Each of the examples is two-dimensional.

Example 2.6.1. Define $\mathrm{S}_P(x)$ as $\mathrm{V}^\star_P(x)$, the visibility star of x relative to P. In this case, d is ill defined because the integral hidden in d does not exist.

Example 2.6.2. Let Δ be the collection of all triangles T such that $x \in T$, and the vertices of T are contained in P. Define $\mathrm{S}_P(x)$ as a minimum-volume element of Δ, or as $\varnothing$ if Δ is empty. The resulting space $(\mathbb{R}^k, \mathcal{T}(\mathbb{R}^k), d)$ is not noise robust.

Example 2.6.3. Let Δ be the collection of all triangles T such that $x \in T$, the vertices of T are contained in P, and $P \cap \mathrm{Int}\,(T) = \varnothing$. Define $\mathrm{S}_P(x)$ as a minimum-volume element of Δ, or as $\varnothing$ if Δ is empty. The resulting space $(\mathbb{R}^k, \mathcal{T}(\mathbb{R}^k), d)$ is not noise robust.

Example 2.6.4. Let Δ be the collection of all triangles T such that x is the centroid of T, the vertices of T are contained in P, and $P \cap \mathrm{Int}\,(T) = \varnothing$. Define $\mathrm{S}_P(x)$ as a minimum-volume element of Δ, or as $\varnothing$ if Δ is empty. The resulting space $(\mathbb{R}^k, \mathcal{T}(\mathbb{R}^k), d)$ is not noise robust.

2.7 Summary

In this section the theoretical results of the previous sections are summarised. Table 2.1 gathers the properties of the pseudometric pattern spaces. The first column indicates the pseudometric pattern space under consideration. The second column, indicates whether the positivity property (Axiom 2.1.2) holds or not, that is, whether the structure under consideration is a metric pattern space or not. The third column indicates a topological transformation group under which the pseudometric is invariant. The last four columns indicate if the pseudometric pattern space satisfies deformation, blur, crack, and noise robustness (Axioms 2.2.1, 2.2.2, 2.2.3, and 2.2.4, respectively).

The following conventions are used in the Table 2.1. The symbol X denotes a topological space, and $\wp(X)$ denotes the collection of all subsets of X. The symbol ρ always denotes a metric on X. The group of isometries in a metric space X (with metric ρ) is denoted by $\mathrm{Iso}(X)$.

The first rows in Table 2.1 describe "trivial" structures that are based on the indiscrete pseudometric z and the discrete metric c. The properties for the indiscrete pseudometric z on $\wp(X)$ follow directly from the fact that the value is always zero for the pseudometric. The properties for the discrete pseudometric follow from the fact that this metric has value one for two distinct patterns.

Table 2.1 contains results for various metric pattern spaces endowed with the Hausdorff metric h_ρ and its Euclidean version h are summarised. First, the results are given for the Hausdorff metric defined on the collection $\mathcal{K}'(X)$ of all nonempty compact subsets of a space X with metric ρ. In the next rows, the domain is restricted to the collection $\mathcal{S}'(X)$ of nonempty "solid patterns"

Space	Pos.	Inv.	Defo.	Blur	Crack	Noise
$(X, \wp(X), \mathsf{z})$	false	$\mathrm{Hom}(X)$	true	true	true	true
$(X, \wp(()X), \mathsf{c})$	true	$\mathrm{Hom}(X)$	false	false	false	false
$(X, \mathcal{K}'(X), \mathsf{h}_\rho)$	true	$\mathrm{Iso}(X)$	true	true	true	false
$(X, \mathcal{S}'(X), \mathsf{h}_\rho)$	true	$\mathrm{Iso}(X)$	true	true	true	false
$(X, \mathcal{T}(\mathbb{R}^k), \mathsf{h})$	true	$\mathrm{Iso}(X)$	true	true	true	false
$(\mathbb{R}^k, \mathcal{K}(\mathbb{R}^k), \mathsf{s})$	false	UDif^k	true	true	true	true
$(\mathbb{R}^k, \mathcal{S}(\mathbb{R}^k), \mathsf{s})$	true	UDif^k	true	true	true	true
$(\mathbb{R}^k, \mathcal{M}(\mathbb{R}^k), \mathsf{s})$	false	UDif^k	true	true	true	true
$(\mathbb{R}^k, \mathcal{K}'(\mathbb{R}^k), \mathsf{s})$	false	CDif^k	true	true	true	true
$(\mathbb{R}^k, \mathcal{S}'(\mathbb{R}^k), \mathsf{s})$	true	CDif^k	true	true	true	true
$(\mathbb{R}^k, \mathcal{M}(\mathbb{R}^k), \mathsf{s})$	false	CDif^k	true	true	true	true
$(\mathbb{R}^k, \mathcal{M}(\mathbb{R}^k), \mathsf{r})$	false	UAf^k	true	true	true	true
$(\mathbb{R}^k, \mathcal{T}(\mathbb{R}^k), \mathsf{r})$	true	UAf^k	true	true	true	true
$(\mathbb{R}^k, \mathcal{M}(\mathbb{R}^k), \mathsf{r}^*)$	false	Af^k	true	true	true	true
$(\mathbb{R}^k, \mathcal{T}(\mathbb{R}^k), \mathsf{r}^*)$	true	Af^k	true	true	true	true

Table 2.1: The properties of various pseudometric pattern spaces

(of Definition 2.5.2) and the collection $\mathcal{T}(\mathbb{R}^k)$ of nonempty "thin patterns" (of Definition 2.6.2). Deformation, blur and crack robustness hold on $\mathcal{K}'(X)$, by Theorems 2.4.5, 2.4.6 and 2.4.7, respectively. By Theorem 2.2.3, deformation, blur, and crack robustness are preserved under the restriction of $\mathcal{K}'(X)$ to $\mathcal{S}'(X)$ and $\mathcal{T}(\mathbb{R}^k)$. By Theorem 2.4.8, noise robustness does not hold for $\mathcal{K}'(X)$, $\mathcal{S}'(X)$ and $\mathcal{T}(\mathbb{R}^k)$.

Table 2.1 contains information for the volume of symmetric difference s defined on the collections $\mathcal{K}(\mathbb{R}^k)$, $\mathcal{S}(\mathbb{R}^k)$ and $\mathcal{M}(\mathbb{R}^k)$. By Theorem 2.5.4, s defined on $\mathcal{K}(\mathbb{R}^k)$ is invariant for UDif^k. Deformation, blur, crack and noise robustness hold on $\mathcal{K}(\mathbb{R}^k)$, by Theorems 2.5.7, 2.5.8, 2.5.9 and 2.5.10, respectively. By Theorem 2.2.3, each of the four types of robustness also hold on $\mathcal{S}(\mathbb{R}^k)$ and $\mathcal{M}(\mathbb{R}^k)$.

For the normalised volume of symmetric difference s^* the results are similar to those of s. The main difference is that by Theorem 2.5.5, s^* defined on $\mathcal{K}'(\mathbb{R}^k)$ is invariant for CDif^k. Satisfaction of the robustness axioms for s^* follow from the results for s by Theorem 2.2.3 and Corollary 2.5.11.

The reflection visibility distance r is considered on the collections $\mathcal{M}(\mathbb{R}^k)$ and $\mathcal{T}(\mathbb{R}^k)$. By Theorem 2.6.5, r is invariant for UAf^k. By Theorem 2.6.11, deformation, blur, crack and noise robustness hold. For the normalised reflection visibility distance r^*, the results are the same, except that this distance is

invariant for Af^k.

2.8 Experimental results

In this section, practical experiments are performed using computer programs that compute the Euclidean Hausdorff metric h (which is shorthand for h_{l_2}) and the normalised reflection visibility distance r^*. Patterns from a database are compared with one another using both h and r^*. The pattern database consists of hieroglyphics which are modelled as finite unions of line segments in the plane. These hieroglyphics were obtained from the hieroglyphica sign list [62]. The experiments compare the robustness of h and r^*.

The experiments of this section compare the behaviour of h and r^* under increasing amounts of distortion. Each robustness axiom of Section 2.2 corresponds to a particular type of distortion. These types of distortion are deformation, blur, crack, and noise. In Section 2.4, it was proven that h is deformation, blur, and crack robust, but not noise robust. In Section 2.6 it was proven that r^* is robust for each of the four types of distortion. It is expected that the outcome of the experiments confirms the theoretical results concerning the robustness of h and r^*. This is made more precise by two hypotheses. The first hypothesis is that, in practical experiments, both similarity measures behave in a robust manner if distortion of the types deformation, blur, and crack occurs in patterns. That is, the distance between the altered pattern and the original pattern is small, if the alterations are small. The second hypothesis is that r^* behaves robust if noise is added to patterns, while h does not. That is, if small outliers are added, h measures a large distance between the original and the altered pattern, while r^* only measures a small distance.

Figure 2.10 shows six patterns $A_1, \ldots, A_5, B$. Pattern A_1 is a hieroglyphic. Patterns $A_2, \ldots, A_5$ are increasingly distorted versions of A_1: Pattern A_2 is a slightly deformed image of A_1. Pattern A_3 is a version of A_2 in which parts have been connected and separated at small cracks. Pattern A_4 is a blurred version of A_3. Pattern A_5 is A_4 with noise. Pattern B is a different hieroglyphic, not similar to A_1.

For each pattern, a square containing it is computed. After that, the patterns were scaled such that the square containing them had diameter 1. This causes the values of h to lie between 0 and 1, just like for r^* (which is invariant under this scaling). Then, the distances between each two patterns in the sequence $A_1, \ldots, A_5, B$ were computed. This was done for both the Euclidean Hausdorff metric h and the normalised reflection visibility distance r^*. Tables 2.2 and 2.3 are the distance matrices for h and r^*, respectively.

Under both h and r^*, the distance from the original increases as the pattern becomes more distorted. That is, $\mathsf{h}(A_1, A_i)$ and $\mathsf{r}^*(A_1, A_i)$ increase as i

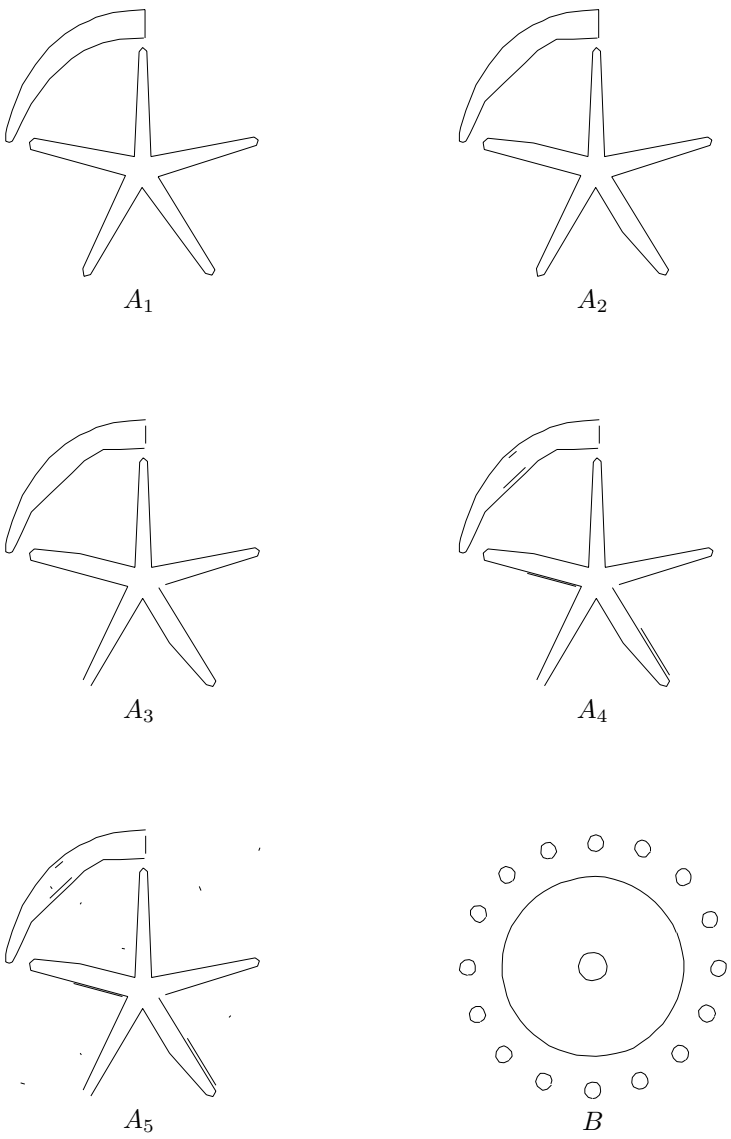


Figure 2.10: Increasingly distorted patterns

h	A_1	A_2	A_3	A_4	A_5	B
A_1	0.00000	0.01581	0.01581	0.01581	0.29000	0.24083
A_2	0.01581	0.00000	0.01414	0.01581	0.29000	0.24083
A_3	0.01581	0.01414	0.00000	0.01581	0.29000	0.24083
A_4	0.01581	0.01581	0.01581	0.00000	0.29000	0.24083
A_5	0.29000	0.29000	0.29000	0.29000	0.00000	0.13509
B	0.24083	0.24083	0.24083	0.24083	0.13509	0.00000

Table 2.2: Distance matrix for h and patterns $A_1, \ldots, A_5, B$

r^*	A_1	A_2	A_3	A_4	A_5	B
A_1	0.00000	0.06249	0.08986	0.09635	0.13371	0.79073
A_2	0.06249	0.00000	0.03700	0.05021	0.08841	0.79242
A_3	0.08986	0.03700	0.00000	0.01415	0.06306	0.79581
A_4	0.09635	0.05021	0.01415	0.00000	0.05214	0.79660
A_5	0.13371	0.08841	0.06306	0.05214	0.00000	0.79200
B	0.79073	0.79242	0.79581	0.79660	0.79200	0.00000

Table 2.3: Distance matrix for r^* and patterns $A_1, \ldots, A_5, B$

increases. The first three increases in distance are relatively small for both h and r^*. The fourth step, adding noise, has far more impact on h as it has on r^*. Table 2.2 indicates that the Hausdorff distance between A_1 and B is smaller than the Hausdorff distance h between A_1 and A_5. This example illustrates that the Hausdorff distance is not noise robust. For the human eye, A_5 is quite similar to A_1, while A_1 and B seem to have no visual similarities at all. The distance r^* agrees with this, since $\mathsf{r}^*(A_1, A_5)$ is much smaller than $\mathsf{r}^*(A_1, B)$ in Table 2.3.

In the test results both h and r^* are robust for deformation, crack, and blur. This confirms the first hypothesis. However, when noise is added, h, not being noise robust gives counter-intuitive distance values. The noise robustness of the reflection metric is confirmed by the test results. This confirms the second hypothesis. In addition, it can be seen that h totally ignore small alterations as long as these do not affect the "worst matching point". On the other hand, the distance r^* changes proportionally under the alterations, reflecting the fact that the patterns become less similar after more changes.

2.9 Discussion

This chapter presented a new theory for similarity measures and pseudometrics in particular. The four robustness axioms are the most important new contributions. These axioms describe a set of properties that can be desirable if a pseudometric is to be used as a basis for pattern matching or shape recognition.

The set of all pseudometrics that are defined on a fixed collection forms a complete lattice of pseudometric topologies. On one extremity, the discrete metric induces the coarsest of all topologies. On the other extremity, the indiscrete pseudometric induces the finest of all topologies. Each of the robustness axioms "pushes" the pseudometric in one direction, namely, that of the indiscrete pseudometric. The indiscrete pseudometric trivially satisfies all of the robustness axioms. This fact is a clear indication that the set of properties described in this chapter is not complete: a good pseudometric should have the properties, but not all pseudometric having the properties are good. To complete the set of axioms, additional "discernment axioms" that push the pseudometric back into the direction of the discrete metric are needed. In fact, the positivity property (of metrics) is an example of such an axiom. The formulation of more powerful discernment axioms for pseudometric pattern spaces is a topic for further research.

Section 2.6 introduces the reflection visibility distance and its normalised version. These similarity measures are metrics on a collection of $(k-1)$-dimensional subsets of $\mathbb{R}^k$. In contrast with parametrisation based metrics (the Fréchet metric, for example), the reflection visibility distance and its nor-

malisation are defined on patterns that have multiple connected components. In addition, the new metrics satisfy the axioms of Section 2.2. Furthermore, the normalised reflection visiblity distance are invariant for affine transformations. Finding metrics, that retain the good properties mentioned above but which have simpler definitions and can be computed more efficiently is a topic for future research.

Chapter 3

Computation of the minimum distance

This chapter deals with the computational side of pattern matching using similarity measures. More precisely, the problems of computing a similarity measure and minimising a similarity measure under a transformation group are dealt with. Each transformation that achieves a minimum can be considered as a geometric match between the patterns. That means that the minimisation of a similarity measure under a transformation group, is a form of pattern matching. Assuming the conditions of Theorem 2.1.1, the minimisation of a similarity measure under a transformation group results in a new similarity measure that is independent of the action of a transformation group. That is, the latter similarity measure compares patterns independent of any transformations that are applied separately to both patterns. This means that the minimisation of similarity measures under transformations can be applied in shape recognition.

The following examples illustrate the use of minimisation under transformations. Minimisation of the Euclidean Hausdorff metric (defined in Section 2.4) under Euclidean isometries can be used to match and recognise patterns if each of the two patterns is rotated and translated independently. Minimisation of the normalised volume of symmetric difference (defined in Section 2.5) under affine transformations can be used to match and recognise solid sets (regions) independent of rotation, shear, scaling and translation applied to the patterns. The same can be done for finite unions of line segments in the plane if the normalised reflection visibility distance (defined in Section 2.6) is used.

The problems studied in this chapter are defined in terms of an objective function which is defined as follows. Consider a pattern collection $\mathcal{P}$, a sim-

ilarity measure d on $\mathcal{P}$, a transformation group G for $\mathcal{P}$ and two patterns A and B in $\mathcal{P}$. The choices of d, G, A and B determine the objective function $f : G \to \mathbb{R}$ which is given by

$$f(g) = d(g(A), B). \tag{3.1}$$

In the sections below, the function f is studied for various choices of d and G, while A and B are fixed.

The following conventions apply throughout this chapter. The notation $f(G)$ is used for the image set of f, that is, $f(G) = \{\, f(g) \mid g \in G \,\}$. The *infimum of* f is the value $\inf f(G)$. A element g in G *minimises* f if $f(g) = \inf f(G)$. It is assumed that for each of the examples of f discussed in this chapter, there is always some g in G that minimises f. This is convenient since it means that the word minimum can safely be used instead of the word infimum.

The objective function f leads to the following computational problems:

1. Constructing an explicit representation for the function f.
2. Reporting all local minima of f.
3. Reporting all global minima of f.
4. Finding an element g of G that minimises f.
5. Compute a transformation g such that $f(g) - \inf f(G) < \epsilon$.
6. Return a transformation g satisfying $f(g) \leq (1 + \beta)\epsilon$ if $\inf f(G) \leq \epsilon$. Return NONE if $\inf f(G) > (1+\beta)\epsilon$. Otherwise, return some transformation $g \in G$.
7. Computing the value $\inf f(G)$.
8. Approximating the value $\inf f(G)$ within an accuracy of $\epsilon > 0$.
9. Approximating the value $\inf f(G)$ with a fixed ratio.
10. Decide whether $\inf f(G) \leq \epsilon$ given some $\epsilon > 0$.

Problem 1, computing an explicit representation, typically involves finding the maximally connected subsets H of G such that each restriction $f|_H$ is an algebraic function. Problem 2, finding all local minima, involves computing the set of all g in G for which there exists an open neighbourhood H such that $f(g)$ equals $\inf f(H)$. Problem 3, finding all global minima, involves computing the set of all g in G such that $f(g)$ equals $\inf f(G)$. Problem 4 involves computing just one minimising transformation. Problem 5 is a weaker, approximate

version of problem 4. Problem 6, investigated by Heffernan and Shirra [78], is also a weaker version of Problem 4. Most existing algorithms for Problem 7, determining the infimum value, actually solve Problem 4. Problems 8, 9 and 10 are weaker versions of Problem 7. Problem 8, approximating the infimum value within a given accuracy $\epsilon > 0$, is the subject of Chapter 4.

The focus of this chapter lies on Problems 4 and 7, which are related with pattern matching and shape recognition, respectively. Algorithms for Problems 7 and 4 compute the orbit pseudometric d^G (of Theorem 2.1.1) and a transformation that determines its value, respectively. If G is the trivial group, consisting of the identity only, then Problem 7 is equivalent to computing the distance $d(A, B)$.

This chapter describes combinatorial and computational results related with the objective function f. Asymptotic bounds are given for the worst case complexity of the representation of f as an arrangement. In addition, algorithms for Problems 7 and 4 are discussed and analysed.

Section 3.1 discusses the general strategies that are used in existing algorithms for Problems 7 and 4. This includes a characterisation of the objective function f in a general setting. In the subsequent sections, specific choices of d and G are considered. Section 3.2 analyses the problem of exact congruence matching Section 3.2. Here, d is the discrete metric (discussed in Section 2.1, and G is the group of Euclidean isometries. In particular, new asymptotic bounds are given for the worst case number of symmetries and congruences in higher dimensions. Section 3.3 discusses the best known results for the minimisation of the Hausdorff metric. The minimisation of the volume of symmetric difference is the subject of Section 3.4. Here, new bounds are given for the complexity of the objective function f as an arrangement. In addition, the local description of f as a rational function is studied. Section 3.5 presents new algorithms that compute the reflection visibility distance.

3.1 General minimisation

This section discusses algorithmic strategies that apply to the minimisation of various similarity measures under various transformation groups.

Assumptions

Below, a number of conventions and assumptions are described. These apply throughout this chapter.

It is assumed that the input patterns A and B are represented by collections of geometric primitives. More precisely, it is assumed that the pattern A (B), a subset of $\mathbb{R}^k$, has an associated representation $\mathcal{A}$ ($\mathcal{B}$). Here, $\mathcal{A}$ ($\mathcal{B}$) is a finite

collection of geometric primitives. The set A (B) is the union of $\mathcal{A}$ ($\mathcal{B}$). The cardinality of $\mathcal{A}$ ($\mathcal{B}$) is denoted by n (m). If n is mentioned without m, then it denotes the maximum of the cardinalities of $\mathcal{A}$ and $\mathcal{B}$.

The transformation group G is represented in a space $\mathbb{R}^r$, where the dimension r depends on the group under consideration. Transformations in G are points in $\mathbb{R}^r$. For example, planar isometries are considered points in $\mathbb{R}^3$. Here, two coordinates represent translation and one coordinate represents rotation. In the field of motion planning, the space $\mathbb{R}^r$ is often called the configuration space.

The representation of transformations is assumed to be such that the function $(g, x) \mapsto g(x)$ from $G \times \mathbb{R}^k$ to $\mathbb{R}^k$ (known as the evaluation mapping) is an algebraic function[1], where $G \times \mathbb{R}^k$ is considered a subset of $\mathbb{R}^{r+k}$. In addition, the function $g \mapsto g^{-1}$ from G onto itself (known as the inversion mapping) is assumed to be an algebraic function, where G is considered a subset of $\mathbb{R}^r$.

Events

Suppose the pattern A is transformed under transformations $g \in G$ resulting in a patterns $g(A)$, while B stays fixed. In the interaction between the transformed geometric primitives of $\mathcal{A}$ and the static geometric primitives of $\mathcal{B}$, *events* occur. Each event designates a change in the expression of the objective function f in terms of the geometric primitives in $\mathcal{A}$ and $\mathcal{B}$. The set of all events determines the combinatorial structure of the objective function f.

The notion of an event is made more clear by the following examples. Consider the volume of symmetric difference in two dimensions, where the collections of primitives $\mathcal{A}$ and $\mathcal{B}$ consist of triangles. Here, an event occurs when a vertex of a transformed triangle of $\mathcal{A}$ lies on an edge of some triangle of $\mathcal{B}$ or a vertex of a triangle of $\mathcal{B}$ lies on an edge of some transformed triangle of $\mathcal{A}$. As another example, consider the Hausdorff metric, where A and B are finite point sets in $\mathbb{R}^k$. For "almost all" transformations g, the value of the Hausdorff metric on $g(A)$ and B equals the distance $\|g(a) - b\|$ for a unique point pair (a, b) in $A \times B$. Events occur at transformations g for which the point pair is not uniquely determined.

For the examples of this chapter, events are given by a collection of equalities and inequalities in terms of the coordinates of transformations as points in $\mathbb{R}^r$. In fact, each event corresponds to a semialgebraic[2] subset of $\mathbb{R}^r$ which has

[1] A real valued algebraic function is a function that can be expressed as a root of an equation in which a polynomial, in the independent and dependent variables, is set equal to zero. A function with range $\mathbb{R}^k$ is algebraic if each of the coordinate functions is algebraic.

[2] A subset of $\mathbb{R}^k$ is called semialgebraic if it can formed by finite union and finite intersection of sets of the form $\{ x \in \mathbb{R}^k \mid p(x) < 0 \}$ and $\{ x \in \mathbb{R}^k \mid p(x) = 0 \}$, where p is a polynomial in k variables.

constant degree[3]. Thus, the event set generates a collection $\mathcal{H}$ of semialgebraic subsets of $\mathbb{R}^r$.

The collection of semialgebraic subsets of $\mathbb{R}^r$ generated by an event set induces an arrangement $\mathcal{R}(\mathcal{H})$, which is a decomposition of $\mathbb{R}^r$ into cells. A cell in $\mathcal{R}(\mathcal{H})$ is a maximally connected subset of $\mathbb{R}^r$ that lies in the intersection of a subcollection $\mathcal{S}$ of $\mathcal{H}$ and that does not intersect some element of $\mathcal{S}-\mathcal{K}$. Consult the survey by Agarwal and Sharir [2] for more information about arrangements.

General minimisation techniques

Consider the collection $\mathcal{H}$ of semialgebraic sets that is generated by the event set. In each of examples that are studied here (except the discrete metric), the objective function f is continuous and piecewise algebraic. More precisely, for each cell C of $\mathcal{R}(\mathcal{H})$ the restriction $f|_C$ is an algebraic function. A brute force method for the minimisation of f is based on the assumption that the infimum of $f|_C$ can be computed for each cell C. This method simply computes the minimum of the infima of f over all cells C in $\mathcal{R}(\mathcal{H})$, resulting in the value $\inf f(G)$.

Algorithm 1 summarises the brute force method that was informally described above. The algorithm is a called brute force method because it traverses the whole arrangement that represents f to compute the infimum of f. Computationally, this approach is at least as expensive as computing a complete representation of the objective function f (Problem 1). The total computation time is likely to be dominated by the computation of the arrangement generated by the event set.

Algorithm 1 Brute force global minimisation

Input: Finite collections $\mathcal{A}$ and $\mathcal{B}$ representing A and B, respectively.

Output: The minimum value of $d(g(A), B)$ over g in G.

1: Using $\mathcal{A}$ and $\mathcal{B}$, determine the event set E.
2: Using E, determine the collection $\mathcal{H}$ of semialgebraic subsets of $\mathbb{R}^r$.
3: Compute the arrangement $\mathcal{R}(\mathcal{H})$ induced by $\mathcal{H}$.
4: **for** each cell C in $\mathcal{R}(\mathcal{H})$ **do**
5: Compute $l(C) = \inf f(C)$.
6: **end for**
7: Return $\min\{\, l(C) \mid C \text{ in } \mathcal{R}(\mathcal{H}) \,\}$.

The concepts used in the brute force algorithm reappear in other min-

[3]The degree of a semialgebraic set is constant if it can be expressed in polynomials of constant degree.

imisation methods. In computational geometry, three (possibly overlapping) strategies are used to obtain fast algorithms:

1. Apply the brute force method, using an efficient method to compute the arrangement induced by the event set.

2. While finding the global minimum, avoiding computation the complete arrangement.

3. Devise an efficient algorithm for testing $\inf f(G) \leq \epsilon$ (Problem 10), and apply parametric search.

Some of the fastest known algorithms are based on the first strategy [39, 81, 82]. Obviously, such algorithms can never beat the fastest algorithm that computes the arrangement representing f.

Only few examples of the second strategy are found in literature [40]. This is because it is difficult to develop a method that avoids the computation of local minima that are not global minima. Section 3.3 mentions an example of this strategy for the l_∞ based Hausdorff metric in the plane under translation. Here, the worst case execution time is asymptotically lower than the worst case computation time for the whole arrangement.

The third strategy applies when an algorithm for Problem 10 is available [38]. Such a decision algorithm can sometimes be converted into a minimisation algorithm using Megiddo's parametric search technique [95]. This typically increases the execution time with a polylogarithmic factor. Agarwal, Sharir and Toledo [3] apply parametric search to a variety of problems, including minimisation of the Hausdorff metric on finite unions of line segments under translation.

3.2 Exact congruence matching

Consider finite subsets A and B of $\mathbb{R}^k$. Recall that the discrete metric has value 0 if $A = B$ and 1 otherwise. Under the discrete metric, computing $\inf f(G)$ is equivalent to deciding whether there is some transformation g in G such that $g(A)$ equals B. The problem is also known as exact congruence matching.

Exact congruence is a fundamental problem. This is illustrated by that fact that exact congruence matching can be reduced to the infimum computation problem for any metric d. In addition, the set of global minima includes all exact congruences. This implies that the worst case number of g in G for which $g(A)$ equals B is a lower bound for the worst case number of minima of f under an arbitrary metric d.

Events

For exact congruence matching, the events are very simple: they occur at transformations $g \in G$ for which $g(A)$ equals B. Under the assumptions stated below, the events form a finite subset of G. First, it is assumed that the transformation group G on $\mathbb{R}^k$ consists of affine transformations. Second, it is assumed that A and B are finite subsets of $\mathbb{R}^k$ that are in general position[4]. Third, it is assumed that A and B have sizes at least $k+1$.

Combinatorial bounds

There is a close relation between congruence and symmetry, in particular between the number of symmetries and the number of congruences. Symmetry of geometric figures is a well studied subject; Schulte [113] gives an overview. A few facts are mentioned here. Most results on the symmetry of geometric figures concern the regular solids. Plato (427-347 B.C.) classified the regular solids in three dimensions. Schäfli (1814-1895) determined the regular polytopes in higher dimensions and their symmetry groups. The theory of regular polytopes and symmetry groups was further extended by Coxeter [44] and Grünbaum [64]. There is a complete classification of the regular polytopes in $\mathbb{R}^k$. For each type of regular polytopes, the number of facets, number of vertices, and the size of the symmetry group are known.

Here, symmetry is studied without the regularity restriction. Below, new asymptotic bounds are derived for the worst case (that is, maximal) size of the symmetry groups as a function of point set cardinality.

A few definitions are necessary. A bijection g from $\mathbb{R}^k$ onto itself is said to be a congruence of A and B if $g(A) = B$. Given a transformation group G on $\mathbb{R}^k$, A and B are said to be congruent if G contains some congruence of A and B. The notation $G_{A:B}$ stands for the set of all congruences of A and B under G. If g is a congruence of P and itself, then it is said that g is a symmetry of P. The set of symmetries of P under G is abbreviated as $G_P = G_{P:P}$.

The following theorem states that under each transformation group on $\mathbb{R}^k$, there is a bijection between the congruences of A and B and the symmetries of the point sets A and B.

Theorem 3.2.1. *Let G be a transformation group on $\mathbb{R}^k$. Let A and B be finite subsets of $\mathbb{R}^k$. If $g \in G$ is a congruence of A and B, then*

$$gG_A = G_{A:B} = G_B g. \tag{3.2}$$

[4]A finite subset S of $\mathbb{R}^k$ is said to be in general position if for each $l \leq k$, no $l+1$ points of A lie in an $(l-1)$-dimensional hyperplane.

Proof. Assume that $g(A)$ equals B for a particular $g \in G$. It is only shown that $gG_A = G_{A:B}$. The proof for the other equality is analogous.

First, the containment $gG_A \subseteq G_{A:B}$ is shown: Suppose h is in gG_A. This means $h = gk$, where $k(A) = A$. Then,

$$h(A) = g(k(A)) = g(A) = B. \tag{3.3}$$

This means $h \in G_{A:B}$.

Second, the containment $G_{A:B} \subseteq gG_A$ is shown: Suppose h is in $G_{A:B}$. This means $h(A) = B$. Observe that

$$g^{-1}h(A) = g^{-1}(B) = A. \tag{3.4}$$

This means that $g^{-1}h$ is a member of G_A: Therefore $h = g(g^{-1}h)$ is a member of gG_A. □

A consequence of this lemma is that for each point set cardinality $n = |A| = |B|$, the worst case number of symmetries under G equals the worst case number of congruences under G.

A useful technique in exact congruence matching is the normalisation of transformation groups to "less complex" transformation groups. Below, the notion of normalisation made precise. Let $\mathcal{F}(\mathbb{R}^k)$ be the collection of finite subsets of $\mathbb{R}^k$. Let H be a subgroup of G. Formally, a normalisation from G to H is a function $\langle\cdot\rangle : A \mapsto \langle A\rangle$ from $\mathcal{F}(\mathbb{R}^k)$ to G satisfying $\langle A\rangle G_A \langle A\rangle^{-1} \subseteq H$. In other words, a normalisation assigns a transformation $\langle A\rangle$ to each pattern A such that the symmetry group of A in G becomes a subgroup of H after conjugation with $\langle A\rangle$.

The following theorem states that if there is a normalisation from G to H, then the number of symmetries of a pattern A under G equals the number of symmetries of the normalised pattern $\langle A\rangle(A)$ under H.

Theorem 3.2.2. *If $\langle A\rangle$ is a normalisation from G to H, then $|G_A| = |H_{\langle A\rangle(A)}|$.*

Proof. Something stronger is shown, namely that the function $m_A : g \mapsto \langle A\rangle g \langle A\rangle^{-1}$ is a bijection from G_A onto $H_{\langle A\rangle(A)}$.

First, it is shown that the range of the function m_A is $H_{\langle A\rangle(A)}$. That is, $\langle A\rangle g \langle A\rangle^{-1}$ is a member of $H_{\langle A\rangle(A)}$ for each $g \in G_A$. By definition of a normalisation, $\langle A\rangle g \langle A\rangle^{-1}$ is an element of H. This leaves showing that

$$\langle A\rangle g \langle A\rangle^{-1}(\langle A\rangle(A)) = \langle A\rangle g(A) = \langle A\rangle(A). \tag{3.5}$$

It follows from elementary group theory that m_A is a bijection [15]. More precisely, m_A is a group isomorphism that is called the conjugation by $\langle A\rangle$. Thus G_A and $H_{\langle A\rangle(A)}$ have the same group structure. This finishes the proof. □

A rotation is an isometry from $\mathbb{R}^k$ onto itself that fixes the origin. That is, an isometry $g \in \mathrm{Iso}^k$ is a rotation if $g(o) = o$, where o stands for the origin of $\mathbb{R}^k$. There exist simple normalisations of Iso^k and Sim^k to the group of rotations around the origin. Such normalisations are given below.

The centroid of a finite subset P of $\mathbb{R}^k$ is the element of $\mathbb{R}^k$ that is defined by $\mathrm{c}(P) = |P|^{-1} \sum_{p \in P} p$. A normalisation of Iso^k to the group of rotations is given by $x \mapsto x - \mathrm{c}(A)$. The same expression can be used as a normalisation from Sim^k to the group of rotations around the origin composed with scalar multiplication. The latter group can be normalised to the group of rotations using the maximum distance of any point of A to the origin, $\nu(A) = \max_{a \in A} \|a\|$. The normalisation is given by $x \mapsto \nu(A)^{-1} x$.

Using similar techniques, the transformation groups of translations, homotheties and stretch transformations can be normalised to the trivial group which consists only of the identity. The number of symmetries in these cases is therefore exactly one.

The relation between symmetries under various affine subgroups is clear: if a normalisation exists between two groups, then the worst case number of symmetries under both groups is the same. For this reason, the remainder of the discussion focuses on the worst case number of symmetries under rotation around the origin.

A lemma, given below, states what bounds on the number of correspondences between points are needed to determine a finite number of rotations in $\mathbb{R}^k$. Correspondences between points are defined more precisely as follows. Suppose that $x_1, \ldots, x_t$ are independent points of $\mathbb{R}^k$ and $y_1, \ldots, y_t$ are independent points of $\mathbb{R}^k$. In this case, define a point-point correspondence of size t as a set of point pairs (x_i, y_i) for $i = 1, \ldots, t$. A transformation g is said to satisfy the point-point correspondence if $g(x_i) = y_i$ for each index $i = 1, \ldots, t$. An orthogonal transformation is a linear transformation from $\mathbb{R}^k$ onto itself whose matrix representation M has its own transpose as its inverse, that is, $MM^T = I$. Each rotation around the origin is an orthogonal transformation.

In the following lemma and its proof, the notations $\lfloor \alpha \rfloor$ and $\lceil \alpha \rceil$ are used to denote the floor and ceiling of a real number α, respectively.

Lemma 3.2.3. *For each dimension k there is a positive integer c such that each point-point correspondence of size at least $\lceil k/2 \rceil$ is satisfied by at most c orthogonal transformations.*

Proof. The orthogonality condition $MM^T = I$ is expressed in matrix form by

$$\begin{pmatrix} m_{11} & \ldots & m_{1k} \\ \vdots & & \vdots \\ m_{k1} & \ldots & m_{kk} \end{pmatrix} \begin{pmatrix} m_{11} & \ldots & m_{k1} \\ \vdots & & \vdots \\ m_{1k} & \ldots & m_{kk} \end{pmatrix} = \begin{pmatrix} \delta_{11} & \ldots & \delta_{1k} \\ \vdots & & \vdots \\ \delta_{k1} & \ldots & \delta_{kk} \end{pmatrix}. \tag{3.6}$$

Here, δ_{ij} (the Kronecker delta) is defined as 1 if $i = j$ and 0 if $i \neq j$. Let m_i denote the i-th row of M. The matrix expression corresponds to the following equations in which the matrix elements m_{ij} are variables. There are $k(k-1)/2$ distinct equations of the form

$$m_i {m_j}^T = 0, \tag{3.7}$$

where $1 \leq i < j \leq k$. There are k additional equations of the form

$$m_i {m_i}^T = 1, \tag{3.8}$$

where $1 \leq i \leq k$. Thus, the orthogonality condition consists of $k(k+1)/2$ distinct quadratic equations, each having either k or $2k$ variables.

It is shown that a point-point correspondence of size $\lceil k/2 \rceil$ adds so many new equations that the total system of equations has a finite set of solutions whose size depends on k only. First, the orthogonality equations are conveniently distributed. The orthogonality equations of the first form (Equation 3.7) are partitioned into k groups of size $\lfloor k/2 \rfloor$ as follows. Each orthogonality equation is represented by a pair of row indices. Row $i = 1, \ldots, k$ is assigned the pairs $(i, i+j)$ where $0 \leq j \leq \lfloor k/2 \rfloor - 1$. If $i + j > k$, then $i + j$ is replaced with $i + j - k$. Now, each row has been assigned $\lfloor k/2 \rfloor$ of the orthogonality equations (of which 1 is like Equation 3.7, and $\lfloor k/2 \rfloor - 1$ are like Equation 3.8). The assignment is such that for each two pairs, the combination of indices is distinct, implying that the corresponding equations are distinct. There remain at most $k/2$ orthogonality equations which can not be distributed evenly over the rows.

At this point, each row has been associated with $\lfloor k/2 \rfloor$ orthogonality equations. The $\lceil k \rceil$ point-point correspondences are used to assign $\lfloor k/2 \rfloor$ additional equations to each row. Linking a point x to a point y results in an equation

$$m_i x^T = y \tag{3.9}$$

for each row $i = 1, \ldots, k$. By definition of a point-point correspondence each related point pair (x_i, y_i) in a correspondence results in a distinct set of equations (the points in both tuples are independent). Therefore, a size $k/2$ point-point correspondence suffices to assign k equations to each row.

Now that each row has k equations, the elements of M can be eliminated one by one. Without loss of generality, assume that each coordinate of each point in the correspondence is nonzero. First, the equations generated by the point-point correspondences are used. These are used to express the variables m_{ic} in columns $c = 1, \ldots, \lfloor k/2 \rfloor$ in terms of the variables m_{ic} in the columns $c = \lfloor k/2 \rfloor + 1, \ldots, k$. The variables in the next $\lceil k/2 \rceil - 1$ columns are eliminated using applications of Equation 3.7. The final column is solved by applying

Equation 3.8 for each row. In each of the $k^2/2$ variable reductions using Equations 3.7 and 3.8, there is a choice between two possibilities. Therefore the number of distinct solutions for M is $2^{k^2/2}$ which is a candidate for the integer c stated in the theorem. □

Theorem 3.2.4. *Let k be a fixed dimension. Let A be a finite subset of $\mathbb{R}^k$ that is in general position and has at least k elements. Then, the worst case number of symmetries of A under rotations around the origin is $\Omega(n^{\lfloor k/2 \rfloor})$ and $O(n^{\lceil k/2 \rceil})$.*

Proof. The lower bound claim is shown by defining a construction for each $n \geq 1$. Let $r = \lfloor k/2 \rfloor$. Consider $\mathbb{R}^k$ as a Cartesian product of r planes and an additional real line if k is odd. For each $i = 1, \ldots, r$, the corresponding plane is spanned by coordinates $2i$ and $2i+1$ in $\mathbb{R}^k$. Now create a set of $n \geq 1$ points as follows. For each $s = 1, \ldots, n$, a point a^s is constructed. For each plane $i = 1, \ldots, r$, coordinates $2i$ and $2i+1$ are given by the equations

$$a^s_{2i} = \cos(2\pi s/(n+1)), \tag{3.10}$$

$$a^s_{2i+1} = \sin(2\pi s/(n+1)). \tag{3.11}$$

If k is odd, the last coordinate is set to zero. This results in our construction $A = \{ a_1, \ldots, a_n \}$.

It is clear that in each of the r planes i, each rotation between coordinates $2i$ and $2i+1$ does not affect the other coordinates. For each plane separately, there are $n-1$ distinct rotations that leaves the projection of A in that plane unchanged. Since for each of the r planes rotations may be applied independently, the number of symmetries is $\Omega(n^{\lfloor k/2 \rfloor})$.

This leaves the upper bound. Let $l = \lceil k/2 \rceil$. Choose a subset $\{ a_1, \ldots, a_l \}$ of A. If g is a rotation symmetry, then

$$g(\{ a_1, \ldots a_l \}) \subseteq A. \tag{3.12}$$

By Lemma 3.2.3, the set of all $g \in G$ satisfying this equation has size at most cn^l, where c depends on k only. Because the rotational symmetries of A form a subset of the set of solutions to Equation 3.12, the total number of rotational symmetries of A is $cn^l = O(n^{\lceil k/2 \rceil})$. □

Theorem 3.2.2 shows that the same bounds hold for the worst case number of symmetries under the Euclidean isometries Iso^k and the similarity transformations Sim^k. Because of the connection between symmetry and congruence shown in Theorem 3.2.1, the worst case bounds also apply to the number of congruences under these groups.

Computation

Consider the exact congruence problem for finite subsets A and B of $\mathbb{R}^k$. Using the normalisation steps discussed in the previous section (which take linear time), exact congruence under Sim^k and Iso^k can be reduced to congruence under rotations around the origin. From now on it is assumed that this normalisation is applied to A and B. Throughout, with the exact congruence problem, determining congruence under rotations around the origin is meant.

No algorithm for exact congruence matching can beat the time bound $\Omega(n \log n)$ because the set equality problem is reducible to it [21]. Since exact congruence matching itself is reducible to the minimisation problem under an arbitrary metric, this entire class of problems also has this worst case lower bound.

The upper bound of Theorem 3.2.4 leads directly to an algorithm for exact congruence matching: generate all correspondences for a tuple of points with size $\lceil k/2 \rceil$ and verify whether some of them leads to a congruence. Since each verification takes $O(n \log n)$ time, the overall computation time of this algorithm is $O(n^{\lceil k/2 \rceil + 1} \log n)$.

In algorithms for exact congruence matching, the following preprocessing step is standard. Both A and B are projected onto the surface of the unit sphere. Each point on the sphere is labelled with the distance of the original point to the origin.

For dimensions at most 3, there are optimal $O(n \log n)$ time algorithms to determine exact congruence. For dimension 2, Atallah [20] describes such an algorithm. This algorithm uses a string matching algorithm (described by Hall and Dowling [74]). Optimal algorithms for dimension 3 are given by Atkinson and Alt et al. [21, 12]. The algorithm by Alt et al. uses an algorithm for planar graph isomorphism (described by Aho, Hopcroft and Ullman [4]).

Alt et al. [12] describe a dimension reduction method that can be used for the exact congruence problem in each dimension. A (slightly different) description of the dimension reduction step is given below. Let A and B be point sets of size n in $\mathbb{R}^k$ One point a in A is correlated with each point of B. For each of these n correlations (a, b), a rotation is applied to A such that the last coordinate of a becomes 1. Use A to compute a labelled subset A' of dimension $k-1$ by keeping the first $k-1$ coordinates of each original point and extending the label with the k-th coordinate. The same procedure is applied to B, resulting in a labelled point set B' of dimension $k-1$. This dimension reduction step results in n exact congruence problems for labelled point sets of dimension $k-1$. It is known that for $k = 3$ such a label preserving isometry can be found in $O(n \log n)$. time. Thus, the total time of the resulting exact congruence matching algorithm is $O(n^{k-2} \log n)$.

The principle of dimension reduction has been used to obtain algorithms

with lower time bounds. A randomised algorithm by Akutsu [5] works in $O(n^{\lfloor k/2 \rfloor} \log n)$ time. This algorithm is of the Monte Carlo type, that is, it gives false results with a controllable probability. Braß and Knauer [31] describe a deterministic algorithm that achieves a dimension step of three. This results in the current best time bound of $O(n^{\lceil k/3 \rceil} \log n)$. The same authors argue that an $O(n \log n)$ time algorithm might exist.

Summary

The combinatorial results for exact congruence matching are summarised below. The results all follow from Theorems 3.2.1 and 3.2.4. The bounds on the worst case number of exact congruences for finite point sets in $\mathbb{R}^k$ are given in the next table (where the notation Rot^k stands for the group of rotations around the origin in $\mathbb{R}^k$).

Group	Lower bound	Upper bound
Id^k Lat^k Thet^k Stret^k	1	1
Rot^k Iso^k Sim^k	$\Omega(n^{\lfloor k/2 \rfloor})$	$O(n^{\lceil k/2 \rceil})$

The fastest known algorithms that solve the exact congruence matching problem are summarised in the below table.

Group	Time bound	Source
Id^k Lat^k Thet^k Stret^k	$O(n \log n)$	
Rot^k Iso^k Sim^k	$O(n^{\lceil k/3 \rceil} \log n)$	[31]

3.3 The Hausdorff metric

This section deals with the combinatorial and computational aspects of minimising the Hausdorff metric. The discussion is limited to finite point sets in $\mathbb{R}^k$ Only the l_p based versions of the Hausdorff metric (defined in Equation 1.8)

are considered. Recall that the l_p based Hausdorff metric for $p = 1, \ldots, \infty$ is given by

$$\mathsf{h}_{l_p}(A, B) = \max\{ \max_{a \in A} \min_{b \in B} l_p(a, b), \max_{b \in B} \min_{a \in A} l_p(b, a) \}. \tag{3.13}$$

The events

Consider finite subsets A and B of $\mathbb{R}^k$. The events for the Hausdorff metric as a function of a subgroup G of the Euclidean isometries are characterised below. The Hausdorff metric $\mathsf{h}_{l_p}(g(A), B)$ is a piecewise algebraic function in $g \in G$. In the interior of each piece, the Hausdorff metric equals $l_p(g(a), b)$, for a unique point pair (a, b) in $A \times B$. The boundaries of the pieces, the events, consist of the transformations g for which there is no unique defining pair of points defining $\mathsf{h}_{l_p}(g(A), B)$. When $\mathsf{h}^{\triangleright}_{l_p}(A, B) \geq \mathsf{h}^{\triangleright}_{l_p}(B, A)$, there are two possible types of events:

1. (min-change in B) There exist an $a^* \in A$ and distinct $b_1, b_2 \in B$ such that
$$l_p(g(a^*), b_1) = l_p(g(a^*), b_2) = \mathsf{h}^{\triangleright}_{l_p}(A, B).$$
2. (max-change in A) There exist distinct $a_1, a_2 \in A$ and a $b^* \in B$ such that
$$l_p(g(a_1), b^*) = l_p(g(a_2), b^*) = \mathsf{h}^{\triangleright}_{l_p}(A, B).$$

The first event indicates that the point in B whose distance is minimal to the point of A that has the maximum distance to B changes. The second event indicates that the point in A whose (minimum) distance to B is maximal changes. Two additional events occur when $\mathsf{h}^{\triangleright}_{l_p}(A, B) \leq \mathsf{h}^{\triangleright}_{l_p}(B, A)$. These are analogous to the first two events. The fifth and last event occurs when the maximum of the two directed Hausdorff distances changes.

The events and the corresponding arrangement can be computing efficiently using Voronoi surfaces. Recall that the l_p Voronoi surface for a finite subset P of $\mathbb{R}^k$ is the real valued function $\boldsymbol{v}_P$ on $\mathbb{R}^k$ given by $\boldsymbol{v}_P(x) = \min_{p \in P} l_p(x, p)$. The Hausdorff metric is expressed in terms of Voronoi surfaces as

$$\mathsf{h}(A, B) = \max\{ \max_{a \in A} \boldsymbol{v}_B(a), \max_{b \in B} \boldsymbol{v}_A(b) \}. \tag{3.14}$$

For a subgroup G of Iso^k, the expression $f(g) = d(g(A), B)$ equals

$$f(g) = \max\{ \max_{a \in A} \boldsymbol{v}_B(g(a)), \max_{b \in B} \boldsymbol{v}_A(g^{-1}(b)) \}. \tag{3.15}$$

An arrangement that represents f can be computed efficiently using algorithms that are mentioned below.

Combinatorial bounds

First, upper bounds on the the complexity of f are considered. If G equals Lat^k, the group of translations in $\mathbb{R}^k$, then it follows from Equation 3.15 that f equals the upper envelope (that is, the pointwise maximum) of a finite number of translated Voronoi surfaces. Using this observation, Huttenlocher, Kedem and Sharir [82] show that the complexity of f for $G = \mathrm{Lat}^2$ is $O(nm(n+m)\alpha(nm))$ for general l_p and $O(nm(n+m))$ for l_1 and l_∞. Here, α stands for the inverse Ackermann function. For translations in three dimensions this gives a bound of $O(n^2m^2(n+m)\alpha(nm))$ for l_2. For other transformation groups G (represented in $\mathbb{R}^r$) and general dimension k, a rough upper bound for the complexity of f for l_2 and l_∞ is given by

$$O(\Big(nm^{\lceil k/2 \rceil} + mn^{\lceil k/2 \rceil}\Big)^r). \tag{3.16}$$

This is a simple application of two facts. First, the complexity of l_2 and l_∞ based Voronoi surfaces of n points is $O(n^{\lceil k/2 \rceil})$. Second, the overlay of n semi-algebraic subsets of in $\mathbb{R}^r$ has complexity $O(n^r)$.

In addition to upper bounds, nontrivial lower bounds have been found for the worst case complexity of f. Rucklidge [111] investigates the complexity of the objective function f for various choices of the transformation group G. This complexity is equivalent to that of arrangement in $\mathbb{R}^r$ induced by the events. Rucklidge gives lower bounds for the worst case number of connected regions in transformation space for which f is smaller than some given value. This number is a lower bound for the worst case complexity of f. For the l_1, l_2 and l_∞ based Hausdorff metrics under translation the constructions result in a worst case lower bound of $\Omega(n^3)$. For the l_2 based Hausdorff metric under Euclidean isometries the worst case lower bound is $\Omega(n^5)$. Chew et al. [38] use constructions in higher dimensions to show that for the l_∞ based Hausdorff metric under translation there is an $\Omega(n^{\lfloor 3k/2 \rfloor})$ lower bound for the worst case complexity of f.

Computation

The Hausdorff metric between fixed polygons A and B in the plane can be computed in $O((n+m)\log(n+m))$ time (Alt, Behrends and Blömer [7]). For the minimisation of the Hausdorff metric under translations in two and three dimensions the following results have been found. Chew and Kedem [40] present an algorithm that works for the l_1 and l_∞ based Hausdorff metric in two dimensions in $O(nm\log^2(nm))$ time. Note that this algorithm beats the $\Omega(n^3)$ worst case lower bound for the number of minima. Huttenlocher, Kedem and Sharir [81, 82] use the upper envelopes of Voronoi diagrams to find the

minimum. This results in an algorithm that works in $O(nm(n+m)\log(nm))$ time for the l_2 based Hausdorff metric in two dimensions. In addition, the techniques result in an $O(n^2m^2(n+m)^{1+\epsilon})$ time algorithm (for each $\epsilon > 0$) for the l_2 based Hausdorff metric in $\mathbb{R}^3$.

Chew et al. [39] solve the problem of deciding whether the l_2 based Hausdorff metric under Euclidean isometries in the plane is less than some given ϵ. This decision algorithm performs an exhaustive search of a space of possible combinatorial changes, resulting in a time of $O((n+m)^5\log(nm))$. Application of parametric search converts this into an $O((n+m)^5\log^2(nm))$ minimisation algorithm.

For translation in general dimension k, the following is known. Chew et al. [38] apply parametric search to an algorithm for the decision problem. The decision problem can be solved by testing if the intersection of a collection of unions of cubes is empty or not. This, in turn, is reduced to determining the maximum coverage. This is the problem of determining the maximum number of different box unions that contain some point. The minimisation algorithms that result the application of parametric search work in time $O(n^3\log^2 n)$ for $k=3$, $O(n^{(4k-2)/3}\log^2 n)$ for $3 < k \leq 8$, and $O(n^{5k/4}\log^2 n)$ for $k > 8$. For the minimisation of the l_2 based Hausdorff metric under translation in $\mathbb{R}^k$, there is an algorithm that works in time $O(n^{\lceil 3k/2\rceil+1+\epsilon})$ for each $\epsilon > 0$.

Summary

The following table summarises the upper bounds on the combinatorial complexity of the Hausdorff metric on finite point sets for various dimensions, transformation groups and base metrics.

Group	Base metric	Upper bound	Source
Lat^2	l_1, l_∞	$O(nm(n+m))$	[82]
Lat^2	l_2	$O(nm(n+m)\alpha(nm))$	
Iso^2	l_2	$O((nm)^3)$	Eq. 3.16
Lat^k	l_∞	$O((nm^{\lceil k/2\rceil}+mn^{\lceil k/2\rceil})^k)$	

The following able summarises the known lower bounds on the worst case combinatorial complexity of the Hausdorff metric on finite point sets for various dimensions, transformation groups, and base metrics.

Group	Base metric	Lower bound	Source
Lat^2	l_1, l_∞	$\Omega(n^3)$	[111]
Lat^2	l_2	$\Omega(n^3)$	
Iso^2	l_2	$\Omega(n^5)$	
Lat^k	l_∞	$\Omega(n^{\lfloor 3k/2\rfloor})$	[38]

The following table summarises the fastest known algorithms for the minimisation of the Hausdorff metric on finite point sets under various transformation groups.

Group	Base metric	Time bound	Source
Id^k	l_2	$O((n+m)\log(n+m))$	[7]
Lat^2	l_1, l_∞	$O(nm\log^2(nm))$	[40]
Lat^2	l_2	$O(nm(n+m)\log(nm))$	[82]
Lat^3	l_2	$O(n^2m^2(n+m)^{1+\epsilon})$ for all $\epsilon > 0$	
Iso^2	l_2	$O((n+m)^5\log^2(nm))$	[39]
Lat^k	l_∞	$O(n^3\log^2 n)$ if $k=3$ $O(n^{(4k-2)/3}\log^2 n)$ if $3<k\leq 8$ $O(n^{5k/4}\log^2 n)$ if $k>8$	[38]
	l_2	$O(n^{\lfloor 3k/2\rfloor+1+\epsilon})$ for all $\epsilon > 0$	

3.4 The volume of symmetric difference

In this section, new results are presented for the volume of symmetric difference as a function of transformation. Here, the patterns A and B are unions of finite collections of geometric primitives $\mathcal{A}$ and $\mathcal{B}$, respectively. It is assumed that $\mathcal{A}$ ($\mathcal{B}$) is a collection of k-simplices such that if two k-simplices intersect, their intersection is an l-face of both simplices, where $l < k$. Here, an l-face of a k-simplex S is simply an l-simplex formed by $l+1$ vertices of S. For example, in two dimensions $\mathcal{A}$ ($\mathcal{B}$) consists of triangles which may only intersect in common edges or vertices. Observe that the unions A and B formed in this manner may have multiple connected components and holes.

Recall that the volume of symmetric difference is defined as follows. The symmetric difference of the sets A and B is defined as $A \triangle B = (A-B)\cup(B-A)$. In each dimension $k \geq 1$, the k-dimensional volume of a subset P of $\mathbb{R}^k$ is denoted by $\mathrm{vol}\,(P)$ (implying that in two dimensions, $\mathrm{vol}\,(P)$ denotes the area of P). The volume of symmetric difference is given by $\mathtt{s}(A,B) = \mathrm{vol}\,(A \triangle B)$.

The events

Consider the maximal subsets of the transformation group G for which the intersection of each transformed simplex of $\mathcal{A}$ with each simplex of $\mathcal{B}$ is combinatorially constant. The boundaries of these regions are described by the events. More precisely, the events are sets of transformations $g \in G$ for which $g(S)$ intersects T where S is some l-face of some simplex in $\mathcal{A}$, T is some l'-face of some simplex in $\mathcal{B}$, and $l + l' = k - 1$. The collection of events determines an

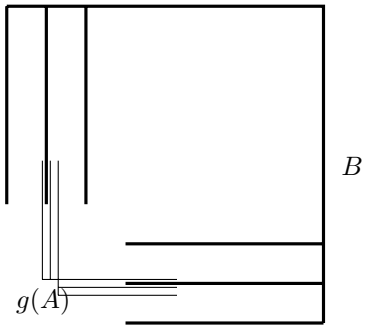


Figure 3.1: Existing construction for translation

arrangement $\mathcal{R}(\mathcal{H})$ in the space $\mathbb{R}^r$ (that represents the transformation space G). The complexity of this arrangement is studied below.

Combinatorial bounds

The complexity of the objective function f is linear in the complexity of the arrangement induced by the events. Recall that r is the lowest dimension needed to represent the transformation group G via an algebraic function. The worst case combinatorial complexity of the arrangement induced by n semialgebraic subsets of $\mathbb{R}^r$ is $\Theta(n^r)$. Each of the $\Theta(nm)$ events gives rise to a semialgebraic subset of $\mathbb{R}^r$. The arrangement induced by these semialgebraic sets has combinatorial complexity $O((nm)^r)$. Below, tight lower bounds are given for the worst case complexity of this arrangement for various transformation groups.

Mount, Silverman and Wu [99] analyse the volume of intersection of simple polygons as a function of translations in two dimensions. This function is a piecewise polynomial of degree two in the two variables that represent translation. The lower bound is achieved using the well known construction by Pollack et al. [107]. The skeleton of this construction is depicted in Figure 3.1. The line segments in the figure actually represent sufficiently thin strips with non-zero area, so $t(A)$ and B can be considered genuine polygons. The vertical teeth can be horizontally translated in $\Omega(nm)$ distinct ways such that $g(A)$ and B do not intersect. Independently, the horizontal teeth can be vertically translated in $\Omega(nm)$ disjoint ways. The "horizontal" and "vertical" combinations can all be formed independent of each other. There is no path in translation space that brings one combination to another without causing an intersection. Therefore they represent $\Omega((nm)^2)$ distinct path-connected components in the space of translations.

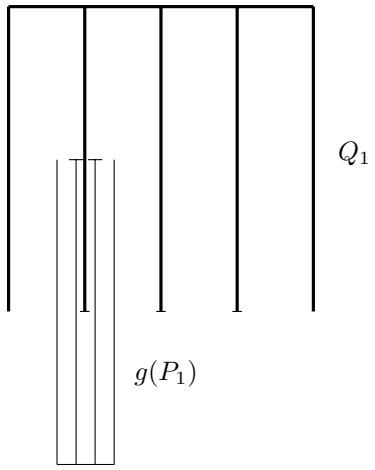


Figure 3.2: Building block for new constructions

Here worst case constructions for various affine subgroups are presented. Each of them is based on the building blocks P_i and Q_i, described below. There are special building blocks for the horizontal direction ($i = 1$) and the vertical direction ($i = 2$). Combining multiple instances of these components result in worst case bounds for translations, isometries, homotheties, similarity transformations, stretch transformations, and affine transformations in two dimensions. Each of the corresponding constructions is connected. Later, it is shown that constructions in higher dimensions k can be made using building blocks for each direction $i = 1, \ldots, k$.

For each choice of parameters p and q, the building blocks P_i and Q_i have sizes $\Theta(p)$ and $\Theta(q)$, respectively. The i-translation of magnitude $\alpha \in \mathbb{R}$ is the mapping $t^i_\alpha : x \mapsto x + \alpha e_i$, where e_i is the i-th basis vector. The components P_i and Q_i are such that there are $\Omega(pq)$ "separate" i-translations t^i_α such that $t^i_\alpha(P_i)$ and Q_i are disjoint. Figure 3.2 shows the skeletons of the "horizontal" building blocks P_1 and Q_1. The horizontal versions of the building blocks are such that there are $\Omega(pq)$ distinct translations in horizontal direction. In addition, each path in Af^2 between two such translations contains a g such that the interiors of $g(P_1)$ and Q_1 intersect. This means that the set of transformations $g \in \mathrm{Af}^2$ satisfying $g(P_1) \cap Q_1 = \varnothing$ consists of $\Omega(pq)$ path-connected components.

The "horizontal" building blocks P_1 and Q_1, described below, lead directly to the "vertical" building blocks P_2 and Q_2 by exchanging first and second coordinates. The building blocks P_1 and Q_1 have p and q teeth, respectively.

The distances within the building blocks depend on the heights of P_1 and Q_1, that are both equal to a given scaling parameter $\gamma > 0$, which can be chosen as desired. Below, it is shown how the dimensions of P_1 and Q_1 depend on γ. It must be ensured that the configurations of the horizontal and vertical building blocks can be manipulated independently. Therefore the width of Q_1, denoted σ_Q, is be chosen strictly smaller than γ. Because Q_1 has q teeth, the distance between the teeth of Q_1, denoted τ_Q, equals $\sigma_Q/(q-1)$. The total width of P_1 is chosen $\sigma_P = \tau_Q(p-1)/p$. At the ends of the middle $q-2$ teeth of Q_1, there are extensions in the form of horizontal segments. The length of these segments must be chosen so small that, in each the configurations of P_1 and Q_1, there is still some "elbow room". This is achieved by choosing the length of the segments, denoted ξ_Q, strictly smaller than $\tau_P/2$. Now is must be ensured that no path in affine transformation space can separate P_1 and Q_1 by placing similar extensions on the middle $p-2$ teeth of P_1. This can be done by choosing ξ_P, the length of the extensions, so long that the space between the extensions $\tau_P - \xi_P$ is sufficiently small. More precisely, ξ_P is must be chosen so that $(\tau_P - \xi_P)/\xi_Q$ is strictly smaller than $\sigma_P/(2\tau_Q)$, that is, the length ξ_P should be chosen strictly larger than $\tau_P - \xi_Q\sigma_P/(2\tau_Q)$. Each path in affine transformation space that would make the gap between the extensions of P_1 wider than ξ_Q would cause the P_1 to be wider than τ_Q, causing P_1 and Q_1 to intersect. This finishes the description of the skeletons of P_i and Q_i.

The next step is to "inflate" P_1 and Q_1 into simple polygons by computing the Minkowski sum with a small cube C with some width $\delta > 0$, $C = [-\delta, \delta] \times [-\delta, \delta]$. By our previous choices of the dimensions of P_1 and Q_1, the value $\delta > 0$ can be chosen so small that the combinatorial properties of P_1 and Q_1 remain preserved:

- There are still $\Omega(pq)$ translations that are not path-connected in the "free" affine transformation space.
- In each combination of teeth for P_1 and Q_1, there is still a vertical freedom of translation of length at least σ_Q.
- There is still some "elbow room" $\epsilon > 0$ such that each point of P_1 has an open cubic neighbourhood of width ϵ disjoint with Q_1, and vice versa.

From now on, it is assumed that P_1 and Q_1 have been inflated in this manner. The same is assumed for the analogous building blocks P_2 and Q_2.

Now that our building blocks are finished, it is easy to re-do the worst case lower bound for translation. Choose A as the union of disjoint translates $t(P_1)$ and $t'(P_2)$. Form B as the union of $t(Q_1)$ and $t'(Q_2)$. In this example, both A and B consists of two components. These can be turned into simple polygons by connecting them with strips. Figure 3.3 shows the skeleton of the construction.

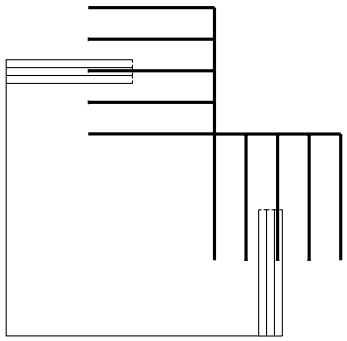


Figure 3.3: New construction for translation

Now, tight a worst case construction is presented for two simple polygons under affine transformation. The affine construction consists of three translation examples (Figure 3.3). Three instances of this translation example will be placed somewhere in the plane. Observe that the result consists of a total of six applications of the primitive constructions (P_i, Q_i), each corresponding to one parameter in affine transformation space. It is assumed that each g in Af^2 is represented in $\mathbb{R}^6$ as

$$g : x \mapsto \begin{pmatrix} g_1 & g_3 \\ g_2 & g_4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} g_5 \\ g_6 \end{pmatrix}. \tag{3.17}$$

The placement of each translation example is such that the configuration of each of the constructions can be modified using a pair of parameters of the transformation without affecting the other two. The construction is such that each of the six parameters can be used to reconfigure a building block (Figure 3.2) independently.

The three translation constructions can be combined to form an affine example. For this purpose, three scaled and translated versions of the translation construction of Figure 3.3, having superscripts 1, 2 and 3, will be used. The first version, (A_1, B_1), has scaling parameter 1 and intersects the horizontal axis in $\mathbb{R}^2$ at a distance δ from the origin. The second version, (A_2, B_2), also has scaling parameter 1 and intersects the vertical axis in $\mathbb{R}^2$ at the distance δ from the origin. The third version, (A_3, B_3), intersects the origin and has scaling parameter $\epsilon > 0$. The lower bound for affine transformation, (A, B) is formed by $A = A_1 \cup A_2 \cup A_3$ and $B = B_1 \cup B_2 \cup B_3$. By choosing δ large enough and $\epsilon > 0$ small enough, the three instances can be manipulated independently, resulting in $\Omega((pq)^6)$ path-connected components in free transformation space. First $\epsilon > 0$ is chosen so that the instance 3 can be manipulated using the trans-

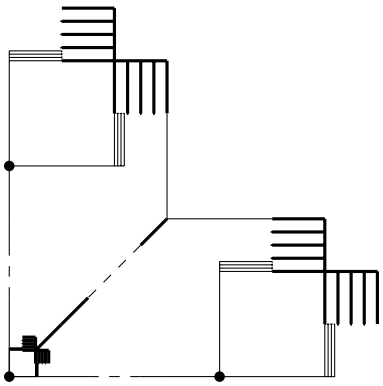


Figure 3.4: Construction for affine transformations

lation parameters g_5 and g_6 without affecting instances 1 and 2. By choosing the distance $\delta > 0$ large enough, instance 1 can be manipulated using parameters g_1 and g_2 without affecting 2 and 3. By symmetry the same holds for instance 2 and parameters g_2 and g_4. The whole construction is turned into a pair of simple polygons by adding connecting line segments. This results in the construction schematically drawn in Figure 3.4. The dotted lines represent long distances.

Using similar techniques as given above connected polygons can be constructed for the transformation groups of isometries, stretch transformations, and homotheties. In fact, these examples can be formed by leaving out primitive instances from the above affine example. The principle is that each dimension of transformation space corresponds exactly one building block. These results are summarised as follows.

Theorem 3.4.1. *For the transformation groups* Lat^2, Iso^2, Thet^2, Sim^2 Stret^2 *and* Af^2, *the worst case complexity of the volume of symmetric difference of* connected *polygons in* $\mathbb{R}^2$ *as a function of transformation is* $\Omega(n^r)$, *where* r *is the dimension of the transformation space.*

If the connectedness requirement is dropped, similar lower bounds can be given in dimension of 1 and higher, as stated in the next result.

Theorem 3.4.2. *For the transformation groups* Lat^k, Iso^k, Thet^k, Sim^k Stret^k *and* Af^k, *the worst case complexity of the volume of symmetric difference of unions of finite collections of* k*-simplices in* $\mathbb{R}^k$ *(where* $k > 0$*) is* $\Omega(n^r)$, *where* r *is the dimension of the transformation space.*

The volume of intersection as a function of transformation

At this point, the combinatorial bounds for the arrangement in transformation space in which for each cell, the intersection is combinatorially constant are known. Recall that this arrangement is generated by all events in which a vertex of some simplex of $\mathcal{A}$ ($\mathcal{B}$) intersects a $(k-1)$-face of some simplex of $\mathcal{B}$ ($\mathcal{A}$). Clearly, for each cell in the arrangement, the volume of intersection (or symmetric difference) is an algebraic function.

Consider the volume of intersection function from transformations $G \subseteq \mathrm{Af}^2$ to $\mathbb{R}$ given by $i(g) = \mathrm{vol}\,(g(A) \cap B)$. There is a simple relation with the volume of symmetric difference:

$$\mathsf{s}(g(A), B) = \delta_g \,\mathrm{vol}\,(A) + \mathrm{vol}\,(B) - 2i(g). \tag{3.18}$$

The relation with the normalised volume of symmetric difference is the following:

$$\mathsf{s}^*(g(A), B) = \frac{\delta_g \,\mathrm{vol}\,(A) + \mathrm{vol}\,(B) - 2i(g)}{\delta_g \,\mathrm{vol}\,(A) + \mathrm{vol}\,(B) - i(g)}. \tag{3.19}$$

Below, the volume of intersection function i is analysed.

If G equals Lat^2 (represented as $\mathbb{R}^2$), the restriction of f to a cell in the arrangement is a polynomial on $\mathbb{R}^2$ with degree two [99]. Below, the local properties of i if G equals Af^2 (represented as $\mathbb{R}^6$) are studied. The value $i(g)$ for $g \in \mathrm{Af}^2$ is a sum over all triangle pairs (S, T) in $\mathcal{A} \times \mathcal{B}$ of the volume of intersection $\mathrm{vol}\,(g(S) \triangle T)$. The intersection $g(S) \cap T$ is the convex hull of a finite set of points. Each such point can be: a vertex of $g(S)$, a vertex of T, an intersection of an edge of $g(S)$ and an edge of T.

The coordinates of the three types of vertices can be expressed in terms of $g \in \mathrm{Af}^2$, where Af^2 is represented in $\mathbb{R}^6$ by Equation 3.17. Three types of vertices in the intersection as a function c from $\mathbb{R}^6$ to $\mathbb{R}^2$ are described below. If the intersection vertex is $g(a)$ for some vertex a of S, then the coordinates of $c(g)$ are written as degree-1 polynomials in g by

$$c(g) = \begin{pmatrix} a_1 g_1 + a_2 g_3 + g_5 \\ a_1 g_2 + a_2 g_4 + g_6 \end{pmatrix}. \tag{3.20}$$

If intersection-vertex is a "static" vertex b of T, then the coordinates are a degree-0 polynomials in g:

$$c(g) = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}. \tag{3.21}$$

Finally, consider the case when the intersection-vertex is formed by edges E and F belonging to S and T, respectively. Let v and w be the direction vectors

of E and F, respectively. Furthermore, let a and b be vertices incident to E and F, respectively. let L_g denote the "linear part" of g, the matrix

$$L_g = \begin{pmatrix} g_1 & g_3 \\ g_2 & g_4 \end{pmatrix}. \tag{3.22}$$

The intersection vertex $c(g)$ is found by equating $g(a) + \alpha L_g(v)$ with $b + \beta w$. For each g, β, denoted $\beta(g)$, is a quotient of determinants

$$\beta(g) = \det(L_g(v), g(a) - b)/\det(L_g(v), w). \tag{3.23}$$

The denominator in the expression $\beta(g)$ is a monomial in the first four coordinates of $\mathbb{R}^6$ of degree 1:

$$\det(L_g(v), w) = v_1 w_2 g_1 + v_2 w_2 g_3 + -v_1 w_1 g_2 + -v_2 w_1 g_4. \tag{3.24}$$

Here, $\det(x, y)$ stands for the determinant of the matrix with first column x and second column y. This leads to the expression of the edge-edge intersection as

$$c(g) = \begin{pmatrix} b_1 + \beta(g) w_2 \\ b_2 + \beta(g) w_2 \end{pmatrix}. \tag{3.25}$$

Each of the two coordinates of $c(g)$ is a rational function in $\mathbb{R}^6$, that is, a quotient of polynomials in $\mathbb{R}^6$. In fact, the polynomial in the denominator is the monomial of Equation 3.24.

At this point, expressions for the vertices that make up the intersection of $g(S)$ and T is known. The volume of this intersection is a sum of determinants, where each determinant is an expression

$$c_1(g) c_2'(g) - c_2(g) c_1', \tag{3.26}$$

where c and c' are distinct intersection-vertices expressed in g. Therefore,

$$\mathrm{vol}(g(S) \cap T) = p(g)/q(g), \tag{3.27}$$

where p is a polynomial in $\mathbb{R}^6$ and q is a monomial in $\mathbb{R}^6$. The monomial p is a finite product of expressions $\det(L_g(v), w)$ for various distinct direction vectors v and w (having the form given in Equation 3.24). This is so because the terms of 3.26 must be normalised in order to be added.

The volume of $g(A) \cap B$ equals the sum of volumes of $g(S) \cap T$ over all S in $\mathcal{A}$ and T in $\mathcal{B}$. In this sum there appear $O(nm)$ terms as in Equation 3.27, with distinct monomial $q(g)$ (which depends on the direction vectors of intersecting edges). These observations lead to the following result.

Theorem 3.4.3. *Let $\mathcal{A}$ and $\mathcal{B}$ be finite collections of triangles in $\mathbb{R}^2$, having unions A and B, respectively. The volume of intersection* $\mathrm{vol}\,(g(A) \cap B)$ *as a function of $g \in \mathrm{Af}^2$ is piecewise rational from $\mathbb{R}^6$ to $\mathbb{R}$. This rational is the quotient of a polynomial and a monomial both having degree $O(nm)$.*

Now, consider the special case where A and B are finite unions of axis-parallel rectangles (closed intervals) in $\mathbb{R}^2$. In this case, the monomial of Equation 3.24 can have only one of the four forms g_1, g_2, g_3, and g_4. Thus each term $c_1(g)c_2'(g)$ and $c_1'(g)c_2(g)$ of Equation 3.26 can be put in the following canonical form:

$$p(g)/(g_1^2 g_2^2 g_3^2 g_4^2), \tag{3.28}$$

where $p(g)$ is a polynomial in $\mathbb{R}^6$. Once converted to this canonical form, the terms can be added freely without need of normalisation of numerators and denominators. In order to get a term in canonical form, its degree of at most 4, needs to be raised with at most 8. This gives the next result.

Theorem 3.4.4. *Let $\mathcal{A}$ and $\mathcal{B}$ be finite collections of axis-parallel rectangles in $\mathbb{R}^2$, having unions A and B, respectively. The volume of intersection* $\mathrm{vol}\,(g(A) \cap B)$ *as a function of $g \in \mathrm{Af}^2$ is piecewise rational from $\mathbb{R}^6$ to $\mathbb{R}$. This rational is the quotient of a polynomial of degree* 12 *and a monomial of degree* 8*.*

Computation

Below, the computation of the minimum of the volume of symmetric difference under various transformation groups is discussed. The few existing results on this topic are dealt with.

First, minimisation under Id^k is discussed. This is simply computing the volume of symmetric difference for two finite unions of k-simplices in $\mathbb{R}^k$ This computation is proceeds as follows: For each $(k-1)$-face of a k-simplex in in $\mathcal{A}$ and $\mathcal{B}$ store a pointer to the simplex it belongs to. Compute the incidence graph of the arrangement $\mathcal{R}(\mathcal{A} \cup \mathcal{B})$ induced by the collection of all simplices of both $\mathcal{A}$ and $\mathcal{B}$. During a depth-first search of the incidence graph the volumes of all cells that are in the symmetric difference can be summed using the pointers. If the technique by Edelsbrunner, O'Rourke and Seidel [52] is used to compute the incidence graph of the arrangement, the total time is bounded by $O((n+m)^k)$.

Existing results for the area of overlap of two polygons [99] directly lead to a minimisation algorithm for the two-dimensional volume of symmetric difference under translation. Consider the area of overlap as a function i from Lat^2 (represented in $\mathbb{R}^2$) to $\mathbb{R}$ given by $i(g) = \mathrm{vol}\,(g(A) \cap B)$. The value $i(g)$ is a polynomial of degree two in the coordinates of g. The area (two-dimensional volume) of symmetric difference as a function of translation, denoted f, is

simply the sum of the areas of the polygons minus two times the area of overlap as a function of translation:

$$f(g) = \mathrm{vol}\,(A) + \mathrm{vol}\,(B) - 2i(g). \tag{3.29}$$

This means that the two functions are equivalent from a combinatorial viewpoint. A representation of area of overlap function for translation can be computed in time $O(nm\log(nm) + |i|)$. Here, $|i|$ denotes the complexity of a representation of the function i as an arrangement of line segments in $\mathbb{R}^2$. The algorithm that computes the representation can be extended so that it maintains the maximum of local maxima. In fact, this allows the maximum of i, the minimum of f, to be found without increasing the time bound.

Summary

Bounds on the worst case complexity of the volume of symmetric difference as a function of transformation are given in the next table. The result for translation in two dimensions is known [99]. The other results follow from Theorem 3.4.2.

Group	Dimension group (r)	Worst case
Id^k	0	$\Theta((nm)^r)$
Lat^k	k	
Iso^k	$k(k+1)/2$	
Thet^k	$k+1$	
Sim^k	$k(k+1)/2+1$	
Stret^k	$2k$	
Af^k	$k(k+1)$	

3.5 Reflection visibility based similarity measures

Computing the visibility based similarity measures r and r^* of Section 2.6 is a nontrivial task. This section contains the first algorithms for this problem. The running times of the algorithms are expressed in terms of the complexity of intermediate structures whose complexity can be expected to significantly lower than the worst case in practical situations. Vital in the computation strategy are partitions of the plane based on reflection visibility. A randomised algorithm constructs these reflection visibility partitions using existing algorithms for visibility graphs and trapezoidal decompositions. The construction algorithm is optimal if its worst case performance is measured in terms of input

size only. The distances r and r^* can be computed by traversing the overlay of two reflection visibility partitions,

Visibility graphs and partitions

The visibility graph is a well studied structure. For a collection of n planar line segments, the visibility graph is the graph having the endpoints of the line segments as vertices, and having edges between vertices for which the corresponding endpoints can be connected by an open line segment disjoint with all segments in the collection. If the number of visibility edges is e, an output-sensitive algorithm by Pocchiola and Vegter [106] computes the visibility graph in $O(n \log(n) + e)$ time and $O(n)$ space.

The reflection visibility distance and its normalised version can be computed using alternative forms of visibility partitions. Visibility partitions consist of equivalence classes with constant combinatorial visibility. Plantinga and Dyer [105] call this structure the *viewpoint space partition*. The dual of the visibility partition is called the *aspect graph*. Results for aspect graphs have been found by Kriegman and Ponce [91], Bowyer and Dyer [30], and Gigus et al. [59]. The number of possible views, the size of the visibility partition, was investigated, under varying assumptions, by Agarwal and Sharir [1], and de Berg et al. [47]. For polygons, results about visibility partitions were found by Guibas et al. [65], Aronov et al. [17], and Bose et al. [29].

Here, the focus lies on the structure of visibility partitions as arrangements. In addition to partitions based on "standard" visibility, partitions based on two alternative forms of visibility are considered. These are called trans visibility and reflection visibility. As a start, standard visibility is considered. The corresponding partition is described through an alternative approach. The same approach is used in the treatment of the other two partitions. It is shown that for each of the partitions, the worst case complexity is $\Omega(n^4)$. Let e be the number of visibility edges (at most quadratic in n) and w be the number of vertices in the partition (at most quadratic in $n + e$). Randomised algorithms are presented that compute each of the three partitions in $O((n+e)\log(n)+w)$ time. Finally, the techniques are applied in computing the reflection visibility distance and its normalised version. It is shown that the reflection visibility based distances can be computed by constructing and traversing an arrangement that is an overlay of two reflection visibility partitions. Let A and B be unions of n and m segments, respectively. If the overlay of the two corresponding reflection visibility partitions has complexity c, the reflection distance between A and B can be computed in $O(c(n + m))$ time.

The following sections characterise the visibility, trans visibility, and reflection visibility partitions. Each of them is provided with a worst case construction for the combinatorial complexity, and give randomised algorithms that are

worst case optimal in the number of segments n.

Visibility partitions

Throughout, $\mathcal{S} = \{ S_1, \ldots, S_n \}$ is a finite collection of closed line segments and $P = \{ p_1, \ldots, p_v \}$ is the corresponding set of endpoints. In addition, it is assumed that the endpoints are in general position. For convenience, set $A = \bigcup \mathcal{S}$. The presentation is simplified by the introduction of an additional "line segment". Let D be an open rectangle containing the union of segments A, and let S_0 be its boundary. Throughout, the set S_0 is treated as an ordinary line segment. This gives an extended collection of segments $\mathcal{S}' = \{ S_0 \} \cup \mathcal{S}$. Set $A' = S_0 \cup A$. Each ray emanating from some point in D intersects A'.

Recall from Section 2.6 that the visibility star $\mathrm{V}^\star_A(x)$ is the union of all open line segments connecting the point x with visible points of A. Consider the endpoints and segments bounding the visibility star $\mathrm{V}^\star_A(x)$ ordered by slope with respect to x. This describes the structure of the visibility star. The visibility star is a finite union of triangles. Each triangle is an intersection of three half-planes. Two of the half-planes are bounded by lines through x and a point in P. The third half-plane is bounded by the line through a segment of $\mathcal{S}$. If a segment $S_i \in \mathcal{S}$ has a visible endpoint p_j, and x is collinear with S_i, then the triangle "degenerates" to the open line segment $\mathrm{Seg}(x, p_j)$.

A compact description for the structure of the visibility star $\mathrm{V}^\star_A(x)$ is needed for each viewpoint $x \in D$. For this purpose, a collection of identifiers is defined. These refer either to segments or endpoints. An identifier is an (integer) index subscripted with a p or an s, indicating an endpoint or a segment, respectively. The identifiers have the following linear order: Each point $a \in A'$ is assigned an identifier $\mathrm{id}(a)$ as follows. If $a = p_i$, then set $\mathrm{id}(a) = i_p$. If $a \in S_i - P$, then set $\mathrm{id}(a) = i_s$.

The structure of the visibility star is represented using a tuple of identifiers. Choose some closed disc centred at x disjoint with A. The boundary of such a disc is called the *view circle*, denoted by C_x. Each point $c \in C_x$ is given a label $l(c)$, as follows: for each point $a \in A'$ visible from x, compute the intersection $c \in C_x \cap \mathrm{Seg}(x, a)$ and set $l(c) = l(a)$.

The view circle C_x is a disjoint union of inverse images $l^{-1}(d)$, for each identifier d. Each non-empty inverse image $l^{-1}(d)$, the subset of points in C_x with label d, can be decomposed into its connected components. These components are either single points or open arcs. The *view map* of x, denoted with $\mathrm{Vmp}_{\mathcal{S}}(x)$, is a labelled circuit graph whose vertices are the components with their constant labels. Vertices labelled with endpoint identifiers (i_p) are called *p-vertices*. Vertices labelled with segment identifiers (i_s) are called *s-vertices*. Edges of $\mathrm{Vmp}_{\mathcal{S}}(x)$ are defined by pairs of vertices, constant-label components, with intersecting closures. Figure 3.5 shows the labelled view

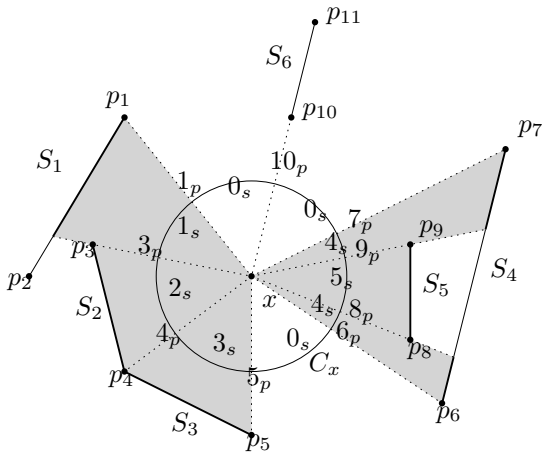


Figure 3.5: A view map

circle, inducing the view map, for a collection of six closed line segments having eleven distinct endpoints. Labels of p-vertices are indicated on the dotted lines on the outside of C_x. Labels of s-vertices are indicated inside the view circle between successive dotted lines.

The view map $\mathrm{Vmp}_{\mathcal{S}}(x)$ can be represented using a tuple of labels encountered when traversing all edges, starting with some initial vertex and some incident edge. Of all possible tuples, the lexicographically smallest one represents the view map. This representation does not depend on the direction (clockwise or counter-clockwise) in which the labels occur on the view circle. The view map is identified with this unique tuple. In the situation of Fig. 3.5 this gives:

$$\mathrm{Vmp}_{\mathcal{S}}(x) \quad = \quad (0_s, 1_p, 1_s, 3_p, 2_s, 4_p, 3_s, 5_p, 0_s, 6_p, 4_s, 8_p, 5_s, 9_p, 4_s, 7_p, 0_s, 10_p).$$

The view map $\mathrm{Vmp}_{\mathcal{S}}(x)$ is a function of points $x \in D$. Define points $x, y \in D$ to be equivalent if their view maps (labelled graphs) are isomorphic, that is, $\mathrm{Vmp}_{\mathcal{S}}(x) = \mathrm{Vmp}_{\mathcal{S}}(y)$. This equivalence relation results in a partition of D into equivalence classes. This *visibility partition* is denoted by $\mathcal{Q}_v(\mathcal{S})$. If x and y lie in the same class $Q \in \mathcal{Q}_v(\mathcal{S})$ of the partition, the visibility stars $\mathrm{V}^\star_A(x)$ and $\mathrm{V}^\star_A(y)$ have the same structure.

The visibility partition is affine invariant: The partition for the affine transformed set $\mathcal{S}$ equals the affine transformed partition (including the labels). The other two partitions of this section also have this property.

A single class in the visibility partition can have more than one connected

component. For example, in the case of a single segment $\mathcal{S} = \{ S_1 \}$ with endpoints $P = \{ p_1, p_2 \}$, the open half-planes left and right of S_1 are the connected components of the equivalence class in $\mathcal{Q}_v(\mathcal{S})$ having view map $(0_s, 1_p, 1_s, 2_p)$.

The visibility partition has the geometric structure of an arrangement induced by a finite union of closed line segments. Each cell in this arrangement is a connected component of an equivalence class in the partition. As the viewpoint x moves continuously within D, changes occur in $\mathrm{Vmp}_{\mathcal{S}}(x)$. Each time such a change occurs, the set of vertices visible from x changes. The sets of viewpoints x on which changes in the viewmap occur form one-dimensional boundaries in the arrangement describing the visibility partition.

A collection of "event segments" can be constructed for the view map. Let $E_{\mathcal{S}}$ be the collection of (directed) edges in the visibility graph. That is, $E_{\mathcal{S}}$ consist of all pairs of endpoint-indices (i, j) such that p_j is visible from p_i. Extend $E_{\mathcal{S}}$ to a collection $E'_{\mathcal{S}}$ by also including the endpoint-index pairs of each segment in $\mathcal{S}$. Given an endpoint p_i, sort all endpoints p_{j_k}, with $(i, j_k) \in E'_{\mathcal{S}}$, by clockwise angle. This results in a list of endpoint identifiers $j_1, \ldots, j_c$. Let s_k be the segment-identifier of the segment visible from p_i inbetween the angles of p_{j_k} and $p_{j_{k+1}}$ relative to p_i (where $k+1$ is modulo c). Below, event segments are constructed that bound the set of points in D from which p_i is visible.

First, the collection of event segments $\mathcal{P}_i$ is defined. For each $k = 1, \ldots, c$, include in $\mathcal{P}_i$ the closure of the visible part of segment S_{s_k} (visible from p_i). This includes parts of the special segment with index 0_s.

Second, the segment collection $\mathcal{B}_i$ connecting pairs of segments in $\mathcal{P}_i$ is constructed. For each $k = 1, \ldots, c$, construct a closed segment between the two intersections of $\mathrm{Ray}(p_i, p_{j_k})$ with segments in $\mathcal{S}'$. If these two intersection coincide, include no segment in $\mathcal{B}_i$, for that particular k.

The third and last types of segments $\mathcal{X}_i$ are extensions of segments in $\mathcal{S}$. Consider each segment S in $\mathcal{S}$ having endpoint p_i, and having another endpoint p_j. Include in $\mathcal{X}_i$, the closed segment having endpoints p_i, and the intersection of $\mathrm{Ray}(p_i, p_j)$ with $\bigcup \mathcal{P}_i$ that is closest to p_i.

The three types of segments result in the arrangement describing the visibility partition. Let $\mathcal{P}$, $\mathcal{B}$, and $\mathcal{X}$, denote the unions of $\mathcal{P}_i$, $\mathcal{B}_i$, and $\mathcal{X}_i$, over all $i = 1, \ldots, m$, respectively.

Theorem 3.5.1. *The boundaries in the visibility partition are formed by the event segments:*

$$\bigcup_{Q \in \mathcal{Q}_v(\mathcal{S})} \mathrm{Bd}\,(Q) = \bigcup \mathcal{P} \cup \bigcup \mathcal{B} \cup \bigcup \mathcal{X}.$$

Figure 3.6 shows the visibility partition for four line segments having seven distinct endpoints. The elements of $\mathcal{P}$ coincide with the segments of $\mathcal{S}$. The segments in $\mathcal{B}$ and $\mathcal{X}$ are shown dotted and dashed, respectively. The points

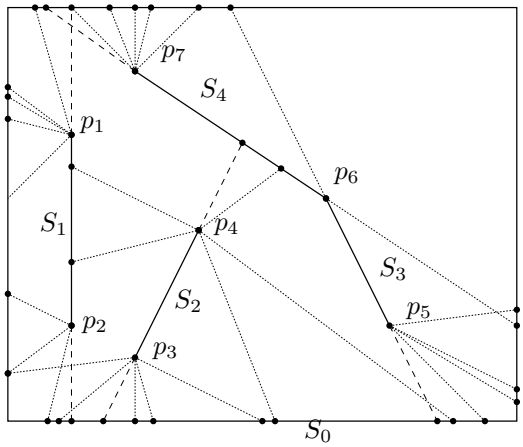


Figure 3.6: A visibility partition

where event segments meet are indicated as dots. The rectangle containing the segments is the "segment" S_0.

Next, the complexity of the visibility partition is investigated. Let e be the number of visibility edges ($e = |E_{\mathcal{S}}|$). It is not difficult to see that total number of event segments in $\mathcal{P}$, $\mathcal{B}$, and $\mathcal{X}$ is $O(n+e)$. Assume this number is h. Since the event segments may intersect, there is an $O(h^2)$ upper bound on the complexity of the visibility partition $\mathcal{Q}_v(\mathcal{S})$ (represented as an arrangement). Since h is $O(n^2)$, it follows that the complexity of $\mathcal{Q}_v(\mathcal{S})$ in terms of n only is $O(n^4)$. Next, it is shown show that this bound is tight.

An $\Omega(n^4)$ lower-bound construction for the complexity of the visibility partition is sketched. Let $n = 2l$. For each $i = 1, \ldots l$, define points $q_i = (i-1, -l)$ and $q_{l+i} = (il, 0)$. Of all lines through these point pairs, at least l^2 have distinct slopes. These l^2 lines form l^4 distinct intersection points. It is possible to perturb the set of q_i $(i = 1, \ldots, n)$ slightly such that the resulting set of p_i $(i = 1, \ldots, n)$ is in general position. To each p_i, a small line segment S_i is attached. By making these attached line segments sufficiently small, the arrangement of event segments gets $\Omega(n^4)$ vertices. Figure 3.7 shows the lower bound construction.

The visibility partition $\mathcal{Q}_v(\mathcal{S})$ corresponding to $\mathcal{S}$ can be computed by generating the collection of event segments and constructing the arrangement induced by these segments. Below it is described how the event segments can be generated efficiently. First, the visibility graph for $\mathcal{S}$, is computed in $O(n \log(n) + e)$ time using algorithms by Ghosh and Mount [58] or Pocchiola

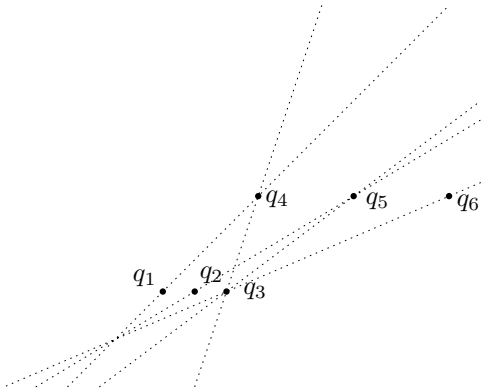


Figure 3.7: A lower bound construction

and Vegter [106]. For each vertex, the visible vertices and line segments are needed. The visible line segments between successive visible vertices (which are ordered counterclockwise) can be found from a visibility graph using the simple algorithm sketched below.

Given is $E'_{\mathcal{S}}$, the set of index pairs belonging to mutually visible vertices or vertices that are joined by a segment of $\mathcal{S}$. The following preprocessing step is performed: for each $(i,j) \in E'_{\mathcal{S}}$, store the edges that "turn left" and "turn right". More precisely, for (i,j), store the index k with $(j,k) \in E'_{\mathcal{S}}$ such that $\mathrm{Ray}(p_j, p_k)$ has minimal angle with $\mathrm{Ray}(p_i, p_j)$ in counter-clockwise (clockwise) direction. This preprocessing step can be performed in $O((n+e)\log(n))$ time. The algorithm proceeds with a number of "iterations" over the total set of indices. For each directed edge $(i,j) \in E'_{\mathcal{S}}$, maintain pointers $\mathrm{l}(i,j)$ and $\mathrm{r}(i,j)$ to visibility edges. After the last iteration, the pointer $\mathrm{l}(i,j)$ ($\mathrm{r}(i,j)$) indicates the line segment visible from vertex p_i directly in counter-clockwise (clockwise) direction of p_j. The pointers $\mathrm{l}(i,j)$ are initialised as follows. If there is a segment of $\mathcal{S}$ adjacent to p_i that is visible in counter-clockwise direction relative to p_j, then let $\mathrm{l}(i,j)$ point to this segment. Otherwise, let $\mathrm{l}(i,j)$ point to the element $(j,k) \in E'_{\mathcal{S}}$ that turns right relative to (i,j). In each iteration, the following is done for each $(i,j) \in E'_{\mathcal{S}}$. If $\mathrm{l}(i,j)$ is not a segment of $\mathcal{S}$ blocking the view from p_i counter-clockwise from p_j, then replace $\mathrm{l}(i,j)$ by $\mathrm{l}(\mathrm{l}(i,j))$. An analogous procedure is applied to the pointers $\mathrm{r}(i,j)$. After $O(\log(e)) = O(\log(n))$ iterations, the segments that block the view on each side of a directed visibility edge are found. This takes a total of $O((n+e)\log(n))$ time.

Now that the view map is found for for each vertex, the total collection of event segments can be generated in $O(n+e)$ time. Let w be the number of inter-

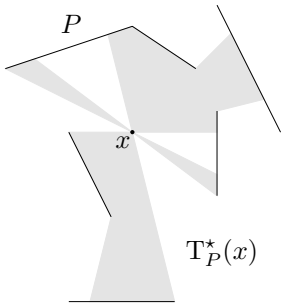


Figure 3.8: A trans visibility star

sections in the collection of event segments thus generated. Using randomised incremental construction, described by Mulmuley [100, pages 84–94], a trapezoidal decomposition of this collection can be built in $O((n+e)\log(n)+w)$ expected time. The arrangement defined by the event segments can be obtained by merging together trapezoids into polygonal cells. Thus, the visibility partition, represented as an arrangement, can be computed using randomised techniques in $O((n+e)\log(n)+w)$.

Theorem 3.5.2. *The visibility partition of n segments has worst case complexity $\Theta(n^4)$. Using randomisation, it can be computed in $O((n+e)\log(n)+w)$ time, where e is the number of visibility edges, and $w = O((n+e)^2)$ is the number of vertices in the arrangement.*

Trans visibility partitions

Here, a different type of visibility star called the trans visibility star is considered. It is an intermediate step between normal visibility, and reflection visibility. The combinatorial structure of the trans visibility star is described by a trans view map. The regions in which the trans view map is constant determine a trans visibility partition. Adaptation of the constructions for the trans visibility partition leads to the reflection visibility partition.

The trans visibility star $\mathrm{T}^\star_A(x)$ is the union of all open line segments between points of A that contain x and are disjoint with A:

$$\mathrm{T}^\star_A(x) = \bigcup\{\,\mathrm{Seg}(y,z) \mid \; y,z \in A,\, x \in \mathrm{Seg}(y,z), \text{ and } A \cap \mathrm{Seg}(y,z) = \varnothing \;\}. \tag{3.30}$$

Figure 3.8 shows a trans visibility star for the set A also used in Figures 2.6 and 2.7.

Consider a view circle C_x with radius $\rho > 0$ centred at x. The view circle is labelled to find the structure of the trans visibility star at x. The labelling l of C_x is defined as follows. Introduce polar coordinates, where the point x acts as the origin, and some fixed direction has the angle 0. Choose an angle $\epsilon > 0$ smaller than the angle between each two endpoints (modulo 2π). For each angle $\alpha \in [0, 2\pi)$, assign the point $c = (\alpha, \rho)$ on C_x a label. Let $\mathrm{Ray}(x, \alpha)$ be the open ray emanating from x in the direction α. There are three cases:

1. If $\mathrm{Ray}(x, \alpha)$ or $\mathrm{Ray}(x, \pi + \alpha)$ intersects no point of A, then $l(c) = 0_s$.
2. Otherwise, if $\mathrm{Ray}(x, \pi + \alpha)$ contains a visible point p of P, and one of $\mathrm{Ray}(x, \pi + \alpha - \epsilon)$ and $\mathrm{Ray}(x, \pi + \alpha + \epsilon)$ does not intersect A, then $l(c) = \mathrm{id}(p)$.
3. In all other cases, $l(c) = \mathrm{id}(a)$, where a is the visible point of A intersected by $\mathrm{Ray}(x, \alpha)$.

It is possible that the previous rules overlap for some $c \in C_x$. Those cases are resolved by only choosing the identifier with minimum index assigned to c.

The resulting labelling l of the view circle C_x indicates the structure of the trans visibility star at the point x. The corresponding labelled circuit graph is called the *trans view map*, denoted by $\mathrm{Tmp}_{\mathcal{S}}(x)$. The trans visibility star is a finite union of triangles, where each triangle, the intersection of three half-planes, is given by an adjacent *psp* vertex triple in the trans view map (where the segment-label is nonzero). The lines through x and the two p-vertices define two half-planes, and the s-vertex defines the third half-plane. Sometimes, just the endpoint of a segment is visible from x. In that case, the triangle "degenerates" to the open line segment connecting x with this endpoint. Figure 3.9 shows a trans view map for five segments with eleven distinct endpoints. The labels in the trans view map are indicated along the view circle. Points that belong to the visible part of A are drawn thick.

Order tuples of labels lexicographically using the underlying order on the identifiers. This way, a unique identifier-tuple describing the trans view map is obtained. For the situation of Fig. 3.9, this results in the following trans view map:

$$\mathrm{Tmp}_{\mathcal{S}}(x) \quad = \quad (0_s, 2_p, 1_s, 1_p, 5_s, 3_p 0_s, 2_p, 3_s, 4_p, 2_s, 3_p).$$

Identifying points $x, y \in D$ if $\mathrm{Tmp}_{\mathcal{S}}(x)$ equals $\mathrm{Tmp}_{\mathcal{S}}(y)$, gives a *trans visibility partition* $\mathcal{Q}_t(\mathcal{S})$. The trans visibility partition has the structure of an arrangement induced by a finite number of line segments.

Next, the collection of event segments for the trans view map is described. Given an endpoint p_i, sort all endpoints p_{j_k}, with $(i, j_k) \in E'_{\mathcal{S}}$, by clockwise angle, resulting in a list of endpoint identifiers $j_1, \ldots, j_c$. For each $k = 1, \ldots, c$,

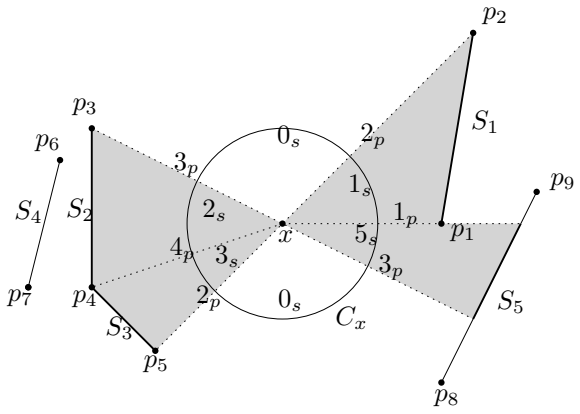


Figure 3.9: A trans view map

s_k is the segment-identifier of the segment visible between p_{j_k} and $p_{j_{k+1}}$. The event segments bound the set of points in D from which p_i is trans visible. The construction is very similar to that of the visibility partition in the previous section.

First, the collection of event segments $\mathcal{P}_i$ is defined. For each $k = 1, \ldots, c$, include in $\mathcal{P}_i$, the closure of the part of segment S_{s_k} that is visible from p_i. Except for $s_k = 0$, in which case no segment is included.

Second, a collection of segments $\mathcal{B}_i$ is defined. These segments connect segments of $\mathcal{P}_i$ at the endpoints. For each $k = 1, \ldots, c$, consider the intersections of $\mathrm{Ray}(p_i, p_{j_k})$ with segments in $\mathcal{B}_i$. If two segments of $\mathcal{B}_i$ intersect the ray in distinct points, the segment connecting these points is included. If there are two coinciding intersections, do not include a segment. If there is only one intersection, the closed segment connecting p_i with this intersection is included.

Third and finally, a segment collection $\mathcal{X}_i$ is constructed. Consider each segment $S \in \mathcal{S}$ having p_i as an endpoint, where p_j is the other endpoint. If the intersection of $\mathrm{Ray}(p_i, p_j)$ with $\mathcal{B}_i$ closest to p_i, is distinct from p_j, then include a closed segment connecting p_i with this intersection.

Again, let $\mathcal{P}$, $\mathcal{B}$, and $\mathcal{X}$ be the unions of $\mathcal{P}_i$, $\mathcal{B}_i$, and $\mathcal{X}_i$, respectively.

Theorem 3.5.3. *The boundaries in the trans visibility partition are formed by the event segments:*

$$\bigcup_{Q \in \mathcal{Q}_t(\mathcal{S})} \mathrm{Bd}\,(Q) = \bigcup \mathcal{P} \cup \bigcup \mathcal{B} \cup \bigcup \mathcal{X}.$$

Figure 3.10 shows a trans visibility partition. The elements of $\mathcal{P}$ coincide

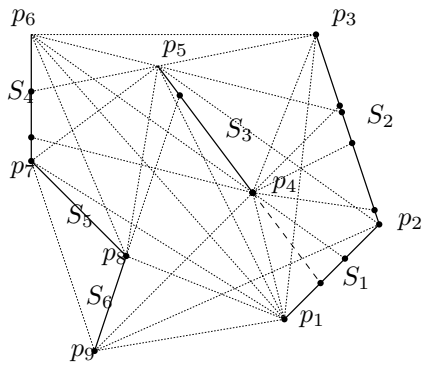


Figure 3.10: A trans visibility partition

with the segments of $\mathcal{S}$. The segments in $\mathcal{B}$ and $\mathcal{X}$ are shown dotted and dashed, respectively.

Using an analysis similar to that for the visibility partition, it follows that the complexity of the trans visibility partition is $O(n^4)$. The worst case lower bound of this complexity is also $\Omega(n^4)$. This bound can be achieved with a construction similar to that presented in the previous section for the visibility partition. That construction ensured $\Omega(n^4)$ intersections between event segments. The same result is obtained here by copying the construction, and adding four extra segments that make up a rectangle containing all these intersections.

Computation of the trans visibility partition is analogous to the computation of the visibility partition. Again, there are $O(n + e)$ event segments. Therefore, the trans visibility partition can be constructed using the visibility partition algorithm from the previous section, resulting in the same time.

Theorem 3.5.4. *The trans visibility partition of n segments has worst case complexity $\Theta(n^4)$. Using randomisation, it can be computed in $O((n+e)\log(n)+w)$ time, where e is the number of visibility edges, and $w = O((n+e)^2)$ is the number of vertices in the arrangement.*

Reflection visibility partitions

Recall that the reflection visibility star, denoted $\mathrm{R}^\star_A(x)$, is defined as the union of all open line segments with midpoint x that are contained in $\mathrm{V}^\star_A(x)$. An equivalent definition is obtained when $\mathrm{V}^\star_A(x)$ is replaced with the trans visibility star $\mathrm{T}^\star_A(x)$. The reflection visibility star leads to the notion of a reflection view

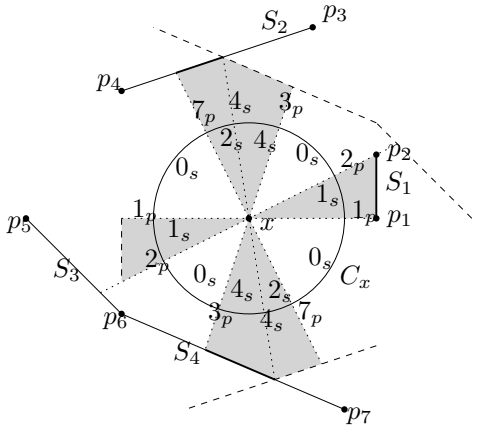


Figure 3.11: A reflection view map

map. In turn, this defines the reflection visibility partition. It is shown that the reflection visibility partition is closely connected to the trans visibility partition.

Consider a view circle C_x centred at x. The view circle is labelled to describe the structure of the reflection visibility star at x. The labelling l of a view circle C_x is as follows. Polar coordinates are used so that each point $c \in C_x$, is represented as $c = (\alpha, \rho)$, where ρ is the radius of C_x. Let $L(x, \alpha)$ be the line through x and (α, ρ). Let $\epsilon > 0$ be smaller than the angle between each two endpoints. There are three cases:

1. If $L(x, \alpha)$ intersects a visible point of S_0, then $l(c) = 0_s$.
2. If $L(x, \alpha)$ intersects a visible endpoint $p \in P$ and $L(x, \alpha - \epsilon)$ or $L(x, \alpha + \epsilon)$ intersects a visible point of S_0, then $l(c) = \mathrm{id}(p)$.
3. In all other cases, set $l(c) = \mathrm{id}(a)$, where a is the visible point in $A \cap L(x, \alpha)$ closest to x.

Overlaps in the rules are resolved by choosing the minimum-index identifier.

The labelling l defines a labelled circuit graph called the *reflection view map*, denoted by $\mathrm{Rmp}_{\mathcal{S}}(x)$. Figure 3.11 shows the view circle along with the labels of the reflection view map. The reflection visible part is shown thick. The dashed lines are reflections of segments in the view point.

The reflection view map $\mathrm{Rmp}_{\mathcal{S}}(x)$ at a point x contains the structure of the reflection visibility star $\mathrm{R}^{\star}_{A}(x)$ at a given point. Starting at some vertex and some initial edge, a tuple of labels representing the reflection view map is obtained. Since this tuple repeats itself, only take the first half is taken. The

lexicographically smallest half-tuple is chosen as the unique representation of $\mathrm{Rmp}_{\mathcal{S}}(x)$. In the situation of Fig. 3.11 this gives:

$$\mathrm{Rmp}_{\mathcal{S}}(x) \quad = \quad (0_s, 1_p, 1_s, 2_p, 0_s, 3_p, 4_s, 2_s).$$

The reflection visibility partition is obtained by identifying points $x, y \in D$ if their reflection view maps $\mathrm{Rmp}_{\mathcal{S}}(x)$ and $\mathrm{Rmp}_{\mathcal{S}}(y)$ are equal. When moving x, if the reflection view map changes, then the set of reflection visible endpoints (relative to x) changes. This is seen as follows. If x moves, either an endpoint identifier, a segment identifier, or an intersection identifier, appears or disappears. In case of an endpoint identifier, the result is immediate. In case of a segment identifier, a change occurs only if one endpoint starts to occlude another one. If an intersection identifier disappears it is replaced by an endpoint identifier. Since a change in the set of endpoints also implies a change in the reflection view map, the the following can be concluded: each class in the reflection visibility partition is a maximal connected subset of D in which a fixed set of endpoints is reflection visible.

Next, a collection of event segments for the reflection view map is defined. Consider an endpoint p_i. Consider $\mathcal{P}_i$, $\mathcal{B}_i$ and $\mathcal{X}_i$ exactly as for trans visibility. Consider a scaling transformation f that leaves p_i fixed and which has a scaling factor of $1/2$. That is, all coordinates relative to p_i are multiplied by $1/2$. To obtain the desired reflection view map, replace all segments in $\mathcal{P}_i$, $\mathcal{B}_i$ and $\mathcal{X}_i$ by their images under f. Like before, construct unions of the segment collections over $i = 1, \ldots, v$.

Theorem 3.5.5. *The boundaries in the reflection visibility partition are formed by the event segments:*

$$\bigcup_{Q \in \mathcal{Q}_r(\mathcal{S})} \mathrm{Bd}\,(Q) = \bigcup \mathcal{P} \cup \bigcup \mathcal{B} \cup \bigcup \mathcal{X}.$$

Figure 3.12 shows a reflection visibility partition. The elements of $\mathcal{P}$, $\mathcal{B}$, and $\mathcal{X}$ are shown dotted, dashed, and coarse dashed, respectively.

The complexity of the reflection visibility partition is at most $O(n^4)$. An $\Omega(n^4)$ worst case lower bound is achieved by copying the construction for trans visibility and making sure that the new segments lie far enough from the original segments.

To compute the reflection visibility, the techniques used to compute the visibility and trans visibility partitions can be applied again.

Theorem 3.5.6. *The reflection visibility partition of n segments has worst case complexity $\Theta(n^4)$. Using randomisation, it can be computed in $O((n + e)\log(n)+w)$ time, where e is the number of visibility edges, and $w = O((n+e)^2)$ is the number of vertices in the arrangement.*

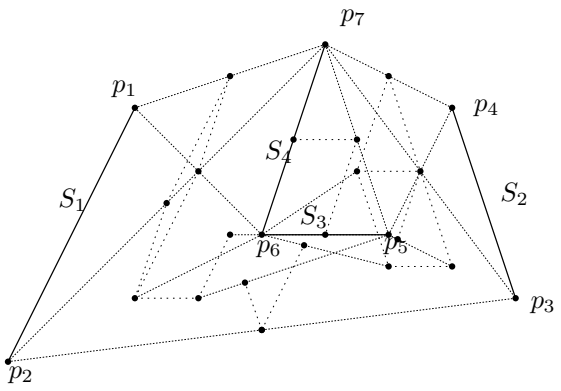


Figure 3.12: A reflection visibility partition

Computing the reflection visibility distance

The previously defined structures can be applied in the computation of reflection visibility based distances. Recall that in two dimensions, $\boldsymbol{r}_A(x)$ is the volume of the reflection visibility star $\mathrm{R}^\star_A(x)$ in $x \in \mathbb{R}^2$. If A and B are finite unions of line segments, the reflection visibility based distances $\mathsf{r}(A, B)$ and $\mathsf{r}^*(A, B)$ are defined by applying the pseudometrics on functions of Section 2.3 to the pair of reflection visibility surfaces $\boldsymbol{r}_A$ and $\boldsymbol{r}_B$. The value $\mathsf{r}(A, B)$ is the integral over $\mathbb{R}^2$ of the pointwise absolute difference of $\boldsymbol{r}_A$ and $\boldsymbol{r}_B$. If $\mathsf{r}(A, B)$ is known, the normalised distance $\mathsf{r}^*(A, B)$ is found by the equation

$$\mathsf{r}^*(A, B) = 2\mathsf{r}(A, B)/(\|\boldsymbol{r}_A\|_1 + \|\boldsymbol{r}_B\|_1 + \mathsf{r}(A, B)). \tag{3.31}$$

Now, the results for reflection visibility partitions are applied to compute the distances. The structure of the function $\boldsymbol{r}_A$ is determined by the reflection visibility partition: for each cell in this partition, $\boldsymbol{r}_A$ is a rational function of degree $O(n)$ in $\mathbb{R}^2$. Similarly, $\boldsymbol{r}_B$ consists of a finite number of rational patches of degree $O(m)$. The pointwise absolute difference $|\boldsymbol{r}_A - \boldsymbol{r}_B|$ is therefore a piecewise rational of degree $O(n + m)$. A model of computation is assumed in which the absolute value of a rational function in two variables, can be integrated over a triangular domain in $\Theta(s)$ time, where s is the maximum degree of the polynomial numerator and denominator.

The computation of the integrals of the rational function $|\boldsymbol{r}_A - \boldsymbol{r}_B|$ proceeds as follows. Let e and e' denote the number of visibility edges corresponding to A and B respectively. First, the visibility graphs of A and B, are computed taking times $O(n \log(n)+e)$ and $O(m \log(m)+e')$, respectively. Then, the event segments that correspond to the reflection visibility partition are computed in

$O(h \log(n))$ and $O(h' \log(m))$ time, where $h = \Theta(e)$ and $h' = \Theta(e')$ are the number of event segments for A and B, respectively. Then, a trapezoidal decomposition for the union of both event segment collections is computed in time $O((h + h') \log(n + m) + w)$, where w is the number of intersections. Integrate $|\boldsymbol{r}_A - \boldsymbol{r}_B|$ by summing the partial integrals over all trapezoids (each trapezoid is a union of two triangles). In our model of computation, this takes $\Theta(n+m)$ time for each trapezoid. Since the summation of the partial integrals dominates the overall complexity, the next result follows.

Theorem 3.5.7. *Let A and B each be unions of n and m segments, respectively. Using randomisation, the distances $\mathsf{r}(A, B)$ and $\mathsf{r}^*(A, B)$ can be computed in $O(c(n + m))$ time, where c $(= O((n + m)^4))$ is the complexity of the overlay of the reflection visibility partitions of A and B.*

3.6 Discussion

In this chapter, combinatorial properties of similarity measures were discussed. In the previous chapter, qualitative properties of similarity measures were discussed. There seems to be a relation between the two types of properties. For example, consider the Hausdorff metric. It was observed in Section 2.4 that this similarity measure is not very robust, because of its max-min type of definition the "worst point" dominates its value. The same property allows efficient minimisation because the resulting min-max-min expression fits nicely into the theory of Davenport-Schinzel sequences. The volume of symmetric difference was shown to be quite robust in Section 2.5. However, for comparable problems, the combinatorial and computational bounds for the volume of symmetric difference exceed those of the Hausdorff metric. It is interesting to investigate whether these observations hold in more generality: might one say that the more robust a similarity measure is, the harder it becomes to compute and minimise this measure? Perhaps, using the axioms of Section 2.2 similar statements may formalised. For example, it might be possible to provide bounds not for a specific similarity measure, but for any similarity measure that satisfies a given set of precisely defined properties.

The evidence of Chapter 3 indicates that for interesting similarity measures (that is, not like the discrete or indiscrete examples), minimisation under a non-trivial transformation group is hard. For example, for the relatively simple problem of minimising the Hausdorff metric under two-dimensional translations, there is no algorithm that has a subquadratic worst case performance. For rigid transformations, no algorithm is known that works in time below $O(n^5)$. Of course, these are upper bounds on the worst case performance, the algorithms might be a lot faster for a lot of inputs. The question remains

how such a statement may be formalised. A possible way out might be to describe the complexity in terms of the size of intermediate constructions (such as Voronoi diagrams in computing the Hausdorff metric). For example, in minimisation algorithms, something meaningful might be said if the complexity of the distance as a function of transformation is incorporated as a variable in the "big-oh".

In Section 3.1, three strategies for minimising similarity measures as a function of transformations were discussed. Many of the existing algorithms are applications of these strategies. The algorithms dealt with in this chapter work in polynomial time. However, for the more "complex" transformation groups such as isometries, homotheties and affine transformations, the degree of these polynomials, although constant in the input size, is quite high. Perhaps, lower asymptotic time bounds may be achieved through different strategies. Below, some suggestions for alternative strategies are given.

A way to "reduce" the complexity of algorithms is through assumptions on the input. Typical examples of this approach assume that the input patterns are convex. This allows faster algorithms for both the Hausdorff metric and the volume of symmetric difference [19, 46]. In many applications of pattern matching, the convexity assumption is too severe. Perhaps, less stringent input assumptions, for example fatness [123], may also result in lower time bounds. This a topic for future research.

Another way to achieve lower asymptotic time bounds may be through randomised algorithms [100]. Section 3.2 mentions a randomised algorithm for the problem of exact congruence matching. The algorithm that computes the reflection visibility distance, presented in Section 3.5, also uses randomised techniques. The application of randomised algorithms in pattern matching problems is a topic for future research.

Chapter 4

Approximation of the minimum distance

This chapter presents algorithms for the following approximate minimisation problem: find lower and upper bounds for the infimum of the values of a similarity measure under a transformation group that differ no more than some given $\epsilon > 0$. The basic approach is inspired by the work of Huttenlocher and Rucklidge [86, 110]. They solve the following approximate *decision* problem: determine if the infimum of the values of the (partial) Hausdorff distance under a transformation group is smaller than some given $\tau > 0$. The algorithm by Huttenlocher and Rucklidge recursively subdivides transformation space into cells. Here, bounds on the maximum change of the (partial) Hausdorff distance over the transformations in a cell are used to reject cells that do not contain some transformation with a (partial) Hausdorff distance less than τ. These bounds are based on bounds for the distance that a point can move under all transformations in a cell.

Together with Veltkamp I adapted the algorithm by Huttenlocher and Rucklidge by using tighter bounds [73]. The adapted method computes bounds using a constructs that are called "traces". A trace is a region that contains the images of a point under all transformations in a cell. The tightening of the bounds results earlier cell rejections, which results in lower execution times. The traces approach has been applied to several transformation groups. These include translations, isometries, stretch transformations and affine transformations [67].

The traces approach can be modified to solve the approximate minimisation problem [71]. The resulting method can be applied for the Hausdorff distance, the volume of symmetric difference, and several other similarity mea-

sures. Mount, Netanyahu, and le Moigne independently devised a similar infimum approximation method [98]. They also elaborate on the multi-resolution method of Huttenlocher and Rucklidge, classifying the basic technique as "geometric branch-and-bound". Mount et al. use so-called uncertainty regions, which are equivalent to our traces. In addition, they speed up the branch-and-bound process using a technique called bounded alignment. This uses the uncertainty regions to sample a number of correspondences between triples of points in both patterns.

Section 4.1 presents the geometric branch-and-bound algorithm in its purest form. This algorithm can be applied to each objective function f that allows uniformly converging bounds. The geometric branch-and-bound algorithm finds, for each given $\epsilon > 0$, lower and upper bounds for the infimum of f that differ at most ϵ. The results given for this algorithm generalise the previous termination results [71]. Most important, it is so general that there are no restrictions on the shape of cells. Theorem 4.1.1 shows that convergence of the algorithm is guaranteed when the "cell bounds" satisfy a uniform convergence condition. Sections 4.2 and 4.3 present two radically different applications of the algorithm; both of them satisfy the conditions of Theorem 4.1.1. Section 4.2 applies the geometric branch-and-bound algorithm using the traces approach, where cells are subsets of a representation space $\mathbb{R}^r$. Section 4.3 applies the algorithm for cells that represent sets of transformations in which the "match" between two pattern simplifications is combinatorially constant. Section 4.4 presents experimental results obtained using implementations of the GBB algorithm following the traces approach.

4.1 The geometric branch-and-bound algorithm

This section presents an algorithm that approximates the infimum of a function f from G to $\mathbb{R}$. The algorithm is called GBB, abbreviating geometric branch-and-bound. Given a precision $\epsilon > 0$, this algorithm finds a lower bound and an upper bound for $\inf f(G)$ such that the difference between the bounds is less than ϵ. First, the fundamental operations for the GBB algorithm are introduced. After that, Theorem 4.1.1 says that GBB terminates if the fundamental operations satisfy a uniform convergence condition.

GBB uses coverings of the domain G. Each set in the covering is represented by a *cell*. The collection of all cells is denoted by $\mathcal{C}$. The subset of G represented by a cell C in $\mathcal{C}$ is written as $G(C)$. The distinction between subsets of G and the cells representing them is convenient in Sections 4.2 and 4.3, where the same definition of cell is used for a variety of domains G.

A cell C is called feasible if $G(C)$ is nonempty. GBB assumes that for each feasible cell C, there exist lower and upper bounds for the infimum of f over

$G(C)$.

Definition 4.1.1. *Cell bounds* are fundamental operations $\mathtt{l}$ and $\mathtt{u}$ satisfying for all feasible $C \in \mathcal{C}$,

$$\mathtt{l}(C) \leq \inf f(G(C)) \leq \mathtt{u}(C). \tag{4.1}$$

Observe that $\mathtt{l}(C)$, being a lower bound for the infimum of the image set $f(G(C))$, is a lower bound of f restricted to $G(C)$. In contrast, $\mathtt{u}(C)$ is only an upper bound for the infimum of f restricted to $G(C)$, not an upper bound for all values in $f(G(C))$.

A cell collection $\mathcal{S}$ is said to cover a subset K of G if the collection of $G(C)$ with C in $\mathcal{S}$ covers K. Each finite feasible cell collection $\mathcal{S}$ covering G results in lower and upper bounds for the global infimum $\inf f(G)$. These global bounds are defined in terms of cell bounds. For each finite feasible cell collection $\mathcal{S}$, the global bounds $\mathtt{lGlobal}$ and $\mathtt{uGlobal}$ are given by

$$\mathtt{lGlobal}(\mathcal{S}) = \min\{\, \mathtt{l}(C) \mid C \in \mathcal{S} \,\}, \tag{4.2}$$
$$\mathtt{uGlobal}(\mathcal{S}) = \min\{\, \mathtt{u}(C) \mid C \in \mathcal{S} \,\}. \tag{4.3}$$

If $\mathcal{S}$ is a finite feasible cell collection covering a subset K of G, then $\mathtt{lGlobal}(\mathcal{S})$ and $\mathtt{uGlobal}(\mathcal{S})$ are lower and upper bounds for $\inf f(K)$, respectively.

Figure 4.1 visualises cell bounds and the resulting global bounds for an example where G is the real line. Here, several cells in $\mathcal{S}$ overlap. The example shows clearly that each upper cell bound $\mathtt{u}(C)$ is only required to be an upper bound for the infimum of f over C: sometimes parts of $f|_C$ lie above $\mathtt{u}(C)$. The example also shows that the minimum upper bound value determining the value of the global upper bound $\mathtt{uGlobal}(\mathcal{S})$ may be smaller than the upper bound $\mathtt{u}(C)$ corresponding to a cell C that minimises $\mathtt{l}(C)$.

The idea behind GBB is that better global bounds can be found by repeated refinement of a cell collection. In each refinement, a cell with minimum lower cell bound is chosen, after which it is removed and replaced by a finite collection of new cells.

Definition 4.1.2. The fundamental operation $\mathtt{cover}$ assigns to each $C \in \mathcal{C}$, a finite collection $\mathtt{cover}(C)$ satisfying

$$G(C) \subseteq \bigcup\{\, G(C') \mid C' \in \mathtt{cover}(C) \,\}. \tag{4.4}$$

Observe that this definition does not demand that the subsets of G represented by cells in $\mathtt{cover}(C)$ are contained in $G(C)$. In addition, the elements of $\mathtt{cover}(C)$ do not have to be feasible. Furthermore, the subsets of G corresponding to a cell collection are allowed to overlap.

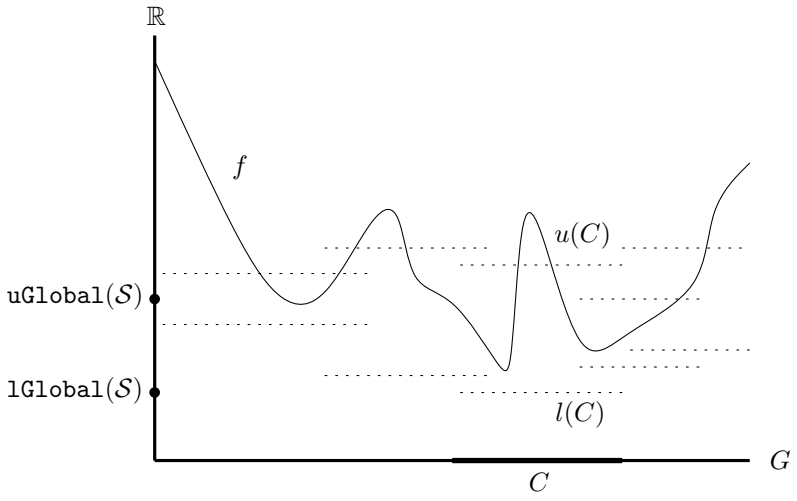


Figure 4.1: Bounds formed by a collection of cells $\mathcal{S}$

The operation `refine`, whose exact definition follows shortly, constructs a refined cell collection given a cell collection. For the refinement, a member C of the cell collection is chosen and replaced by elements of $\texttt{cover}(C)$. Choosing the cells is done by the operation `select`, which for each finite cell collection $\mathcal{S}$, determines a single member of $\mathcal{S}$ minimising `l`:

$$\texttt{l}(\texttt{select}(\mathcal{S})) = \min\{\, \texttt{l}(C) \mid C \in \mathcal{S} \,\}. \tag{4.5}$$

It must be ensured that the current cell collection only includes feasible cells. The operation `refine` is defined in terms of `cover` and `select`: for each finite feasible cell collection $\mathcal{S}$,

$$\texttt{refine}(\mathcal{S}) = (\mathcal{S} - \{\, \texttt{select}(\mathcal{S}) \,\}) \cup \{\, C \in \texttt{cover}(\texttt{select}(\mathcal{S})) \mid G(C) \neq \varnothing \,\}. \tag{4.6}$$

Define $\texttt{refine}^t$ as the t-fold composition of `refine`. Each refined cell collection $\texttt{refine}^t(\mathcal{S})$ covers the set represented by $\mathcal{S}$. If $\mathcal{S}$ covers a subset K of G satisfying $\inf f(K) = \inf f(G)$, then for all $t \geq 0$,

$$\texttt{lGlobal}(\texttt{refine}^t(\mathcal{S})) \leq \inf f(K) \leq \texttt{uGlobal}(\texttt{refine}^t(\mathcal{S})). \tag{4.7}$$

For GBB, presented in Algorithm 2, these inequalities express the main invariant. The operation `refine` uses the operation `select` which determines a cell that has minimum lower bound. This can be implemented efficiently using a

priority queue ordered on the lower cell-bound `l`. This queue maintains the collection of cells $\mathcal{S}$: each call of `refine` removes a least lower bound cell, say C, from the queue, after which the cells in $\texttt{cover}(C)$ are inserted in $\mathcal{S}$. The current global lower bound $\texttt{lGlobal}(\mathcal{S})$ is simply the front of the queue. The current global upper bound $\texttt{uGlobal}(\mathcal{S})$ can be maintained during refinement using a red-black tree.

Algorithm 2 Geometric branch-and-bound

Input: An objective function $f : G \to \mathbb{R}$ and a precision $\epsilon > 0$.

Output: A pair (α, β) satisfying $\alpha \leq \inf f(G) \leq \beta$ and $\beta - \alpha < \epsilon$.

1: Let K be a subset of G such that $\inf f(K) = \inf f(G)$.
2: Compute a cell collection $\mathcal{S}$ covering K.
3: **while** $\texttt{uGlobal}(\mathcal{S}) - \texttt{lGlobal}(\mathcal{S}) \geq \epsilon$ **do**
4: $\quad \mathcal{S} := \texttt{refine}(\mathcal{S})$.
5: **end while**
6: Return $(\texttt{lGlobal}(\mathcal{S}), \texttt{uGlobal}(\mathcal{S}))$.

Let $\texttt{cover}^t$ denote the t-fold composition of `cover`. The following definition gives conditions on the fundamental operations that guarantee correctness of Algorithm 2.

Definition 4.1.3. The fundamental operations `l`, `u` and `cover` *converge uniformly* if for each feasible cell C in $\mathcal{C}$, and each $\epsilon > 0$, there is an r such that $\texttt{u}(C') - \texttt{l}(C') < \epsilon$ for all feasible C' in $\texttt{cover}^r(C)$.

The next theorem proves that uniform convergence is sufficient for Algorithm 2 to be correct.

Theorem 4.1.1. *If the fundamental operations* `l`, `u`, *and* `cover` *converge uniformly, then GBB (Algorithm 2) terminates.*

Proof. It is shown that for each finite feasible cell collection $\mathcal{S}$ covering G, and for each $\epsilon > 0$, there is a number t such that

$$\texttt{uGlobal}(\texttt{refine}^t(\mathcal{S})) - \texttt{lGlobal}(\texttt{refine}^t(\mathcal{S})) < \epsilon.$$

For each C in $\mathcal{S}$, choose $r(C)$ as in the uniform convergence condition. For each C in $\mathcal{S}$, let $n(C)$ be the number of cells derived from C in at least one and strictly less than $r(C)$ cover operations:

$$n(C) = \sum_{n=1}^{r(C)-1} |\texttt{cover}(C)| \,. \tag{4.8}$$

Let $t_{\max}$ be sum of all these sizes increased by one:

$$t_{\max} = 1 + \sum_{C \in \mathcal{S}} n(C). \tag{4.9}$$

Choose $t \leq t_{\max}$ such that $\texttt{select}(\texttt{refine}^t(\mathcal{S}))$ is a member of $\texttt{cover}^{r(C)}(C)$, where $C \in \mathcal{S}$. This choice is valid: Suppose, for a contradiction, that for all $t \leq t_{\max}$, $\texttt{select}(\texttt{refine}^t(\mathcal{S}))$ is a member of $\texttt{cover}^s(C)$, where C in $\mathcal{S}$ and $s < r(C)$. This implies that the total collection of cells created from each cell C in $\mathcal{S}$ in more than 1 and strictly less than $r(C)$ iterations is one larger than $n(C)$. This contradicts the definition of $n(C)$.

By a choice of t, the difference of the cell bounds `u` and `l` corresponding to $\texttt{select}(\texttt{refine}^t(\mathcal{S}))$ must be strictly smaller than ϵ. By definition of `select`, the cell $\texttt{select}(\texttt{refine}^t(\mathcal{S}))$ minimises `l` over $\texttt{refine}^t(\mathcal{S})$. As a result, the difference of global bounds `lGlobal` and `uGlobal` for $\texttt{refine}^t(\mathcal{S})$ can be bounded by ϵ.

$$\begin{aligned}
&\texttt{uGlobal}(\texttt{refine}^t(\mathcal{S})) - \texttt{lGlobal}(\texttt{refine}^t(\mathcal{S})) \\
&= \texttt{uGlobal}(\texttt{refine}^t(\mathcal{S})) - \texttt{l}(\texttt{select}(\texttt{refine}^t(\mathcal{S}))) \\
&\leq \texttt{u}(\texttt{select}(\texttt{refine}^t(\mathcal{S}))) - \texttt{l}(\texttt{select}(\texttt{refine}^t(\mathcal{S}))) \\
&< \epsilon.
\end{aligned}$$

□

Section 4.2 and 4.3 present two different applications of Theorem 4.1.1 to minimise a similarity measure between a transformed pattern and another fixed pattern. The two approaches define the fundamental operations `l`, `u`, and `cover` differently. In Section 4.2, cells are Cartesian products of closed intervals in a finite dimensional Euclidean space that represents the transformation group. In Section 4.3, cells are regions in transformation space that correspond to the same combinatorial intersection of two partitions of the space containing the patterns.

4.2 The traces approach

Theorem 4.1.1 states that for fundamental operations `l`, `u`, and `cover` that converge uniformly, GBB terminates. This section presents fundamental operations satisfying these conditions. The approach is applied to approximate the infimum distance between two patterns A and B for various transformation groups G and various similarity measures d. The choices of d include the Hausdorff distance `h`, and the normalised volume of symmetric difference $\texttt{s}^*$.

Cell bounds based on traces

Perhaps most difficult is finding a lower cell bound `l`. Under conditions discussed below, a similarity measure d has simple lower cell bound. A similarity measure $d^{\triangleright} : \mathcal{P} \times \mathcal{P} \to \mathbb{R}$ *decreases in its second argument* if $B \subseteq B'$ implies $d^{\triangleright}(A, B) \geq d^{\triangleright}(A, B')$. A binary operator $\sqcap : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is increasing if $\alpha \leq \alpha'$ and $\beta \leq \beta'$ implies $\alpha \sqcap \beta \leq \alpha' \sqcap \beta'$.

This section presents a generic lower cell bound that applies to each similarity measure d that can be decomposed symmetrically by applying an increasing operator $\sqcap$ to a similarity measure $d^{\triangleright}$ that decreases in its second argument.

Definition 4.2.1. The pair $(d^{\triangleright}, \sqcap)$ is a *monotone decomposition* of a similarity measure d on $\mathcal{P}$ if for all A and B in $\mathcal{P}$,

$$d(A, B) = d^{\triangleright}(A, B) \sqcap d^{\triangleright}(B, A), \tag{4.10}$$

where $\sqcap$ is uniformly continuous and increasing, and $d^{\triangleright}$ decreases in its second argument.

The general form of the lower cell bound, given shortly, uses constructs called forward and backward traces. For now, suppose $\mathcal{C}$ is just any collection of cells, specific examples are given shortly. For each cell C in $\mathcal{C}$ the forward trace $\overrightarrow{\tau}(C, A)$ and the backward trace $\overleftarrow{\tau}(C, B)$ are defined by

$$\overrightarrow{\tau}(C, A) = \{\, g(a) \mid g \in G(C) \text{ and } a \in A \,\}, \tag{4.11}$$

$$\overleftarrow{\tau}(C, B) = \{\, g^{-1}(b) \mid g \in G(C) \text{ and } b \in B \,\}. \tag{4.12}$$

From here on, assume that the forward and backward traces belong to the collection of patterns $\mathcal{P}$. Considering the set of transformations represented by a cell, the forward trace is the union of all images of a pattern, the backward trace is the union of all inverse images of a pattern. The next results says that under the conditions previously discussed, the forward and backward traces provide a lower cell bound.

Lemma 4.2.1. *Suppose d has a monotone decomposition $(d^{\triangleright}, \sqcap)$. If $d^{\triangleright}$ is invariant under G, then a lower cell bound for f is given by:*

$$\mathtt{l}(C) = d^{\triangleright}(A, \overleftarrow{\tau}(C, B)) \sqcap d^{\triangleright}(B, \overrightarrow{\tau}(C, A)). \tag{4.13}$$

Proof. It must be shown that $\mathtt{l}(C) \leq d(g(A), B)$ for all g in $G(C)$. By definition of traces, $g(A)$ is a subset of $\overrightarrow{\tau}(C, A)$ and $g^{-1}(B)$ is a subset of $\overleftarrow{\tau}(C, B)$. Since $d^{\triangleright}$ decreases in its second element it follows that $d^{\triangleright}(A, \overleftarrow{\tau}(C, B)) \leq$

$d^{\triangleright}(A, g^{-1}(B))$ and $d^{\triangleright}(B, \overrightarrow{\tau}(C, A)) \leq d^{\triangleright}(B, g(A))$. Using that $\sqcap$ is increasing, and that $d^{\triangleright}$ is invariant for G, desired result is found:

$$\begin{aligned} \mathtt{l}(C) &= d^{\triangleright}(A, \overleftarrow{\tau}(C, B)) \sqcap d^{\triangleright}(B, \overrightarrow{\tau}(C, A)) \\ &\leq d^{\triangleright}(A, g^{-1}(B)) \sqcap d^{\triangleright}(B, g(A)) \\ &= d^{\triangleright}(g(A), B) \sqcap d^{\triangleright}(B, g(A)) \\ &= d(g(A), B). \end{aligned}$$

□

Still, an upper cell bound `u` is needed. For this, a technique analogous to that used for the lower cell bound could be applied. However, recall that the upper cell bound is an upper bound for the infimum of f over a cell, and that the upper cell bound is not required to be an upper bound for all values of f over the cell. Therefore, a tighter upper cell bound `u` is given by

$$\mathtt{u}(C) = d(\sigma_C(A), B), \tag{4.14}$$

where σ_C is an element of $G(C)$ for each feasible cell C in $\mathcal{C}$.

Uniform convergence using traces

Throughout the remainder of this section it is assumed that `l` and `u` are defined by Equations 4.13 and 4.14, respectively. The operation `cover` is not yet defined. Below, conditions under which the fundamental operations `l`, `u`, and `cover` are uniformly convergent are discussed.

Here, the focus lies on topological transformation groups G that can be represented in $\mathbb{R}^r$ for some dimension r. Topological transformation groups and various examples of them are discussed in Appendix B. In what follows, the topological transformation group G is considered a subspace of the set of all continuous transformations. The set of all continuous transformations in $\mathbb{R}^r$ has the compact-open topology (which is defined in Appendix A). A transformation representation for G in $\mathbb{R}^r$ is a continuous function

$$\rho : \mathbb{R}^r \to H, \tag{4.15}$$

where H is a set of continuous functions containing G. Cells are restricted to be subsets C of $\mathbb{R}^r$ such that for each $c \in C$, both $\rho(c)$ and $\rho(c)^{-1}$ are members of G. For each such cell C, the corresponding set of transformations is given by

$$G(C) = \{\, \rho(c) \in G \mid c \in C \,\}. \tag{4.16}$$

The following theorem gives sufficient conditions under which the traces approach results in uniformly converging fundamental operations. Recall that

$\mathrm{N}(P, \epsilon)$ stands for the union of all open balls centred at points of P and having radius ϵ. The diameter of a bounded set P, denoted $\mathrm{diam}\,(P)$, is the supremum of $\|p - q\|$ over all point pairs in P.

Theorem 4.2.2. *Let $\mathcal{P}$ be a collection of bounded subsets of $\mathbb{R}^k$; let d be a similarity measure on $\mathcal{P}$. Let G be a topological transformation group on $\mathbb{R}^k$. Let $\rho : \mathbb{R}^r \to G$ be a representation function. Let the cell class $\mathcal{C}$ consist of bounded subsets of $\mathbb{R}^r$. Suppose that the pseudometric d has a monotone decomposition $(d^{\triangleright}, \sqcap)$, where $d^{\triangleright}$ satisfies the triangle inequality and is invariant for G. Furthermore, suppose the following conditions hold:*

1. *For each $P \in \mathcal{P}$ and each $\epsilon > 0$, there is a $\delta > 0$ such that $d^{\triangleright}(P, P') < \epsilon$ for all $P' \in \mathcal{P}$ satisfying $P \subseteq P' \subseteq \mathrm{N}(P, \delta)$.*
2. *For each $C \in \mathcal{C}$, there is a compact $K \subseteq \mathbb{R}^r$ such that $\rho(k), \rho(k)^{-1} \in G$ for all $k \in K$ and $C' \subseteq K$ for all $C' \in \mathtt{cover}^r(C)$ and $r \geq 0$.*
3. *For each $\gamma > 0$ and each C in $\mathcal{C}$, there is an r such that $\mathrm{diam}\,(C') < \gamma$ for all $C' \in \mathtt{cover}^r(C)$.*

Then `l`, `u`, *and* `cover` *converge uniformly.*

Proof. The claim is that for $C \in \mathcal{C}$ and each $\epsilon > 0$, there is a t such that $\mathtt{u}(C') - \mathtt{l}(C') < \epsilon$ for all $C' \in \mathtt{cover}^t(C)$. Observe that this difference is rewritten as

$$\begin{aligned} \mathtt{u}(C') - \mathtt{l}(C') = {} & d^{\triangleright}(A, \sigma_{C'}^{-1}(B)) \sqcap d^{\triangleright}(B, \sigma_{C'}(A)) \\ & - d^{\triangleright}(A, \overleftarrow{\tau}(C', B)) \sqcap d^{\triangleright}(B, \overrightarrow{\tau}(C', A)). \end{aligned} \tag{4.17}$$

The latter expression must be proven strictly smaller than ϵ. By uniform continuity of $\sqcap$, this is accomplished for some $\epsilon' > 0$:

$$d^{\triangleright}(A, \sigma_{C'}^{-1}(B)) - d^{\triangleright}(A, \overleftarrow{\tau}(C', B)) < \epsilon', \tag{4.18}$$

$$d^{\triangleright}(B, \sigma_{C'}(A)) - d^{\triangleright}(B, \overrightarrow{\tau}(C', A)) < \epsilon'. \tag{4.19}$$

Using the triangle inequality (Axiom 2.1.4), the previous inequalities are implied by

$$d^{\triangleright}(\sigma_{C'}^{-1}(B), \overleftarrow{\tau}(C', B)) < \epsilon', \tag{4.20}$$

$$d^{\triangleright}(\sigma_{C'}(A), \overrightarrow{\tau}(C', A)) < \epsilon'. \tag{4.21}$$

Using invariance this is rewritten as:

$$d^{\triangleright}(B, \sigma_{C'} \overleftarrow{\tau}(C', B)) < \epsilon', \tag{4.22}$$

$$d^{\triangleright}(A, \sigma_{C'}^{-1} \overrightarrow{\tau}(C', A)) < \epsilon'. \tag{4.23}$$

Observe that

$$\sigma_{C'}\overleftarrow{\tau}(C', B) \subseteq \bigcup_{g,h \in G(C')} gh^{-1}(B), \tag{4.24}$$

$$\sigma_{C'}^{-1}\overrightarrow{\tau}(C', A) \subseteq \bigcup_{g,h \in G(C')} g^{-1}h(A). \tag{4.25}$$

Choose $\delta > 0$ satisfying Condition 1 for ϵ', and $P = A$ and $P = B$. Now, it is sufficient that

$$\bigcup_{g,h \in G(C')} gh^{-1}(B) \subseteq \mathrm{N}(B, \delta), \tag{4.26}$$

$$\bigcup_{g,h \in G(C')} g^{-1}h(A) \subseteq \mathrm{N}(A, \delta). \tag{4.27}$$

Now, the previous pair of inequalities is implied by

$$\forall_{g,h \in \rho(C')} \forall_{b \in B}\ \|b - gh^{-1}(b)\| < \delta, \tag{4.28}$$

$$\forall_{g,h \in \rho(C')} \forall_{a \in A}\ \|a - g^{-1}h(a)\| < \delta. \tag{4.29}$$

By Condition 2, all descendants C' of C are contained in a compact K. In addition, both $\mathrm{Cl}(A)$ and $\mathrm{Cl}(B)$ are compact. From the continuity of the evaluation mapping (Appendix A), it follows that the following two functions are continuous:

$$K \times K \times \mathrm{Cl}(B) \to \mathbb{R}^k : (c_1, c_2, x) \mapsto \rho(c_1)\rho(c_2)^{-1}(x), \tag{4.30}$$

$$K \times K \times \mathrm{Cl}(A) \to \mathbb{R}^k : (c_1, c_2, x) \mapsto \rho(c_1)^{-1}\rho(c_2)(x). \tag{4.31}$$

Since the domain of these functions is compact, the two functions are uniformly continuous. As a consequence, the diameter $\gamma > 0$ can be chosen such that $\mathrm{diam}(C') < \gamma$ implies (4.28) and (4.29). Using Condition 3, t may be chosen such that the diameter of all cells in $\texttt{cover}^t(C)$ is strictly smaller than γ. □

Observe that Condition 1 only pertains to the decomposition of the similarity measure d using $d^{\triangleright}$, and that Conditions 2 and 3 only pertain to the operation `cover`. Next, decompositions satisfying Condition 1 are discussed for various similarity measures. After that, operations `cover` satisfying Conditions 2 and 3 are presented for various transformation groups. This shows how generic the traces approach is: the decomposition (depending on the similarity measure) and the operation `cover` (depending on the transformation group) can be chosen independently of each other as long as each in isolation satisfies the required properties.

Application to various similarity measures

Most of the metrics discussed in this thesis can be put in the form demanded by Lemma 4.2.1. Therefore, for each of them, the previously introduced cell bounds (Equations 4.13 and 4.14) may be applied. Below, it is shown how the metrics can be decomposed as in Lemma 4.2.1.

The Euclidean Hausdorff metric

Recall that the *directed* Hausdorff distance $\mathsf{h}^{\triangleright}$, is defined on nonempty compact subsets of $\mathbb{R}^k$ by

$$\mathsf{h}^{\triangleright}(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\|. \tag{4.32}$$

In words, this is the maximum Euclidean distance over all points a in A to the point in B closest to a. Using $\mathsf{h}^{\triangleright}$ the Hausdorff metric can be defined as

$$\mathsf{h}(A, B) = \max\{\, \mathsf{h}^{\triangleright}(A, B), \mathsf{h}^{\triangleright}(B, A) \,\}. \tag{4.33}$$

Clearly, as a binary operator, max is increasing. In addition, it is not difficult to see that $\mathsf{h}^{\triangleright}$ decreases in its second element. Finally, $\mathsf{h}^{\triangleright}$ is invariant for all transformations that preserve Euclidean distances. Substituting $d^{\triangleright} = \mathsf{h}^{\triangleright}$, Equation 4.2.1 can be used for a lower cell bound for the Hausdorff distance under each transformation group G that preserves Euclidean distances.

The normalised volume of symmetric difference

A directed version of the normalised volume of symmetric difference is defined as

$$\mathsf{s}^{*\triangleright}(A, B) = \mathrm{vol}\,(A - B)/\mathrm{vol}\,(A \cup B). \tag{4.34}$$

The normalised volume of symmetric difference is defined equivalently as a symmetric sum

$$\mathsf{s}^{*}(A, B) = \mathsf{s}^{*\triangleright}(A, B) + \mathsf{s}^{*\triangleright}(B, A).$$

Clearly, addition is an increasing binary operator. The directed distance $\mathsf{s}^{*\triangleright}$ is decreasing in its second argument: $B \subseteq B'$ implies

$$\begin{aligned} \mathsf{s}^{*\triangleright}(A, B) &= \mathrm{vol}\,(A - B)/\mathrm{vol}\,(A \cup B) \\ &\geq \mathrm{vol}\,(A - B')/\mathrm{vol}\,(A \cup B) \\ &\geq \mathrm{vol}\,(A - B')/\mathrm{vol}\,(A \cup B') \\ &= \mathsf{s}^{*\triangleright}(A, B'). \end{aligned}$$

Other similarity measures

The traces approach applies to many other similarity measures. In fact, the traces approach applies to any similarity measure that can be constructed using the methods of Section 2.1. For example, the approach applies to the (normalised) reflection visibility distance defined in Section 2.6. Another example is the absolute difference defined in Section 2.5. The idea is that the real valued functions used in the construction of similarity measures can be turned into sets of one dimension higher than the domain. This reduction is sketched below.

For each pattern $P \subseteq \mathbb{R}^k$, the corresponding function $\boldsymbol{n}_P$ is formed. The function $\boldsymbol{n}_P$ in turn, can be converted into a subset $P' \subseteq \mathbb{R}^{k+1}$ by the equation

$$P' = \{\, (x_1, \ldots, x_k, \alpha) \mid \ x \in \mathbb{R}^k \text{ and } 0 \leq \alpha \leq \boldsymbol{n}_P(x) \ \}. \tag{4.35}$$

Now, monotone decompositions for s and s^* can be used on the subsets of $\mathbb{R}^{k+1}$. The reduction is finished by defining the action of the a subgroup G of UDif^k for $\mathbb{R}^{k+1}$. Each transformation in G acts on the first k coordinates as usual. The last coordinate is multiplied by the inverse of the absolute value of the Jacobi determinant of g.

4.3 The partition combination approach

This section presents a new method for approximating the minimum volume of symmetric difference: it is an instance of GBB that subdivides transformation space in a radically different way. Its fundamental operations converge uniformly, implying by Theorem 4.1.1 that GBB terminates. Instead of considering cells as subsets of some parameter space of finite dimension, cells represent transformation sets that correspond to combinatorially constant intersections of pattern simplifications.

Strategy

The technique of this section uses pointwise bounds for the restriction of the objective function to a cell in the definition of the cell bounds. Throughout the section it is assumed that A and B are both nonempty finite unions of k-simplices in $\mathbb{R}^k$. Consider the function f from a group G that consists of affine transformations to $\mathbb{R}$ given by

$$f(g) = \mathrm{vol}\,(g(A) \triangle B)/\mathrm{vol}\,(g(A) \cup B). \tag{4.36}$$

For each cell C, consider functions $\underline{f}_C$ and $\overline{f}_C$ that are pointwise lower and upper bounds for the restriction $f_C = f|_{G(C)}$ under g in $G(C)$:

$$\underline{f}_C(g) \leq f_C(g) \leq \overline{f}_C(g). \tag{4.37}$$

A lower cell bound $\mathtt{l}$ and an upper cell bound $\mathtt{u}$ are now provided by:

$$\mathtt{l}(C) = \inf\{\, \underline{f}_C(g) \mid g \in G(C) \,\}, \tag{4.38}$$

$$\mathtt{u}(C) = \inf\{\, \overline{f}_C(g) \mid g \in G(C) \,\}. \tag{4.39}$$

The pointwise bounds $\underline{f}_C$ and $\overline{f}_C$ are determined by the cell C and the patterns A and B. A cell consists of two collections of k-dimensional intervals together with a relation between the collections. One collection partitions a superset of A, the other collection partitions a superset of B. The relation on the collections determines a set of transformations; namely those transformations for which the images of elements of the partition for A intersect exactly those elements of the partition for B that are related.

Figure 4.2 depicts an an overlay of two partitions $\mathcal{K}_C$ and $\mathcal{L}_C$ corresponding to A and B, respectively. It shows the image of $\mathcal{K}_C$ under an affine transformation g. The cell that g belongs to is determined by the intersection relation between $g(\mathcal{K}_C)$ and $\mathcal{L}_C$; this cell represents all transformations for which the intersection relation is constant.

The definition of a cell as two collections with a "must-intersect" relation allows a hierarchical subdivision of transformation space in terms of hierarchical descriptions of the patterns. If one of the collections of the cell is refined (that is, some interval is replaced by a number of smaller intervals), the result is a finite number of subcells, each corresponding to a new possibility in the must-intersect relation formed by the new intervals. Thus, the subdivision of transformation space follows naturally from subdivisions of the two spaces containing the patterns.

The method of this section has a number of clear advantages. First of all, it is an application of GBB with uniformly converging operations, ensuring that it always terminates by Theorem 4.1.1. Second, the method of this section is, as much as possible, independent of topology and representation of the patterns; the patterns A and B are only accessed by measurement of "local volume" expressed as $\mathrm{vol}\,(A \cap K)$ and $\mathrm{vol}\,(B \cap L)$ where K and L are k-dimensional intervals. This allows for robust, reliable implementations. Third, the efficiency of the problem depends more on the "geometric complexity" of the patterns as subsets of $\mathbb{R}^k$ than the complexity of the topology of the patterns and their representations. The algorithm ignores details of the patterns until the point that these details become relevant in the computation.

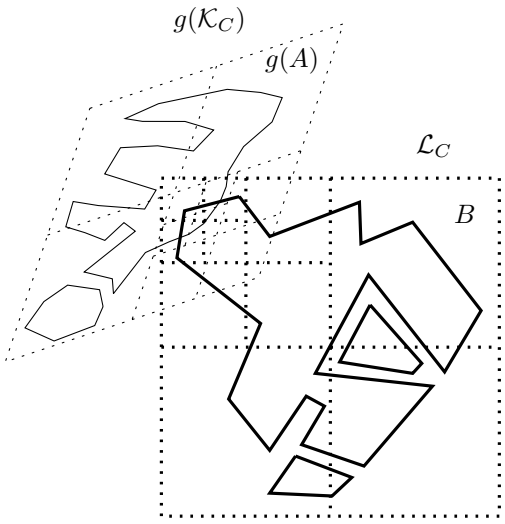


Figure 4.2: Overlay of partitions

Definition of fundamental operations

First, the precise definition of a cell is given. A cell C has two associated collections of k-dimensional closed intervals $\mathcal{K}_C$ and $\mathcal{L}_C$ that correspond to A and B, respectively. The union of $\mathcal{K}_C$ always contains A and the union of $\mathcal{L}_C$ always contains B. The intervals in each partition have disjoint interiors. A cell consists of the two collections $\mathcal{K}_C$ and $\mathcal{L}_C$ together with the must-intersect relation $\mathcal{I}_C$ which is a subset of $\mathcal{K}_C \times \mathcal{L}_C$. Therefore, a cell can be written as a triple $C = (\mathcal{K}_C, \mathcal{L}_C, \mathcal{I}_C)$. The set of transformations $G(C)$ described by a cell is the set of all $g \in G$ such that for all $(K, L) \in \mathcal{K}_C \times \mathcal{L}_C$:

$$(K, L) \in \mathcal{I}_C \Leftrightarrow g(K) \cap L \neq \varnothing \tag{4.40}$$

The objective function f can be expressed as in terms of a number of "local" functions, one for each pair (K, L) in $\mathcal{K}_C \times \mathcal{L}_C$. Each local function is the volume of symmetric difference of $g(A)$ and B within an intersection $g(K) \cap L$, where K is in $\mathcal{K}_C$ and L is in $\mathcal{L}_C$. Pointwise lower and upper bounds for these local functions automatically lead to the desired pointwise lower and upper bounds for f.

Before bounding expressions are given for the local functions, subcollections of $\mathcal{K}_C \times \mathcal{L}_C$ are identified for which the local functions can be computed exactly. Let $\mathcal{K}_C^0$ be the subcollection of $\mathcal{K}_C$ consisting of those K such that $\mathrm{vol}\,(A \cap K)$

equals 0 or $\mathrm{vol}\,(K)$. Let $\mathcal{L}_C^0$ be defined analogously in terms of $\mathcal{L}_C$ and B. The complements are denoted by $\mathcal{K}_C^+ = \mathcal{K}_C - \mathcal{K}_C^0$ and $\mathcal{L}_C^+ = \mathcal{L}_C - \mathcal{L}_C^0$. Let $\mathcal{K}_C^{\rightarrow}$ be the subcollection of $\mathcal{K}_C^+$ consisting of those K such that $\mathrm{vol}\,(B \cap L) = 0$ for all L with $(K, L) \in \mathcal{I}_C$ or $\mathrm{vol}\,(B \cap L) = \mathrm{vol}\,(L)$ for all L with $(K, L) \in \mathcal{I}_C$. Let $\mathcal{L}_C^{\leftarrow}$ be defined analogously. The bounds, introduced shortly, are summations over pairs of elements from $\mathcal{K}_C$ and $\mathcal{L}_C$. For convenience, let $\mathcal{R}_C$ denote the Cartesian product, the set of all such pairs, of $\mathcal{K}_C$ and $\mathcal{L}_C$. The local functions can be computed exactly for the three disjoint subcollections of $\mathcal{R}_C$ given by

$$\mathcal{R}_C^{00} = \mathcal{K}_C^0 \times \mathcal{L}_C^0, \tag{4.41}$$

$$\mathcal{R}_C^{\rightarrow} = \mathcal{K}_C^{\rightarrow} \times \mathcal{L}_C, \tag{4.42}$$

$$\mathcal{R}_C^{\leftarrow} = \mathcal{K}_C \times \mathcal{L}_C^{\leftarrow}. \tag{4.43}$$

Next, the relations are used to form a decomposition of f into local functions. Let $\mathcal{R}_C^{\xi}$ be the union of the three collections $\mathcal{R}_C^{00}$, $\mathcal{R}_C^{\rightarrow}$, and $\mathcal{R}_C^{\leftarrow}$. The complement of $\mathcal{R}_C^{\xi}$ in $\mathcal{R}_C$ is denoted $\mathcal{R}_C^{\iota}$. Over this subcollection, the bounds that are generally inexact are summed. In terms of these subcollections of $\mathcal{R}$, the normalised volume of symmetric difference is decomposed into local functions as

$$f_C(g) = \frac{2\sigma_C(g)}{\mathrm{vol}\,(g(A)) + \mathrm{vol}\,(B) + \sigma_C(g)}, \tag{4.44}$$

$$\sigma_C(g) = \sum_{(K,L)\in\mathcal{R}_C^{\xi}} \xi_{KL}(g) + \sum_{(K,L)\in\mathcal{R}_C^{\iota}} \iota_{KL}(g), \tag{4.45}$$

$$\xi_{KL}(g) = \iota_{KL}(g) = \mathrm{vol}\,((g(A) \triangle B) \cap g(K) \cap L). \tag{4.46}$$

Each ξ_{KL} stands for a local function that can be computed exactly. For the remaining local functions ι_{KL} lower and upper bounds are found below. Pointwise bounds $\underline{f}_C$ and $\overline{f}_C$ for f_C follow from the decomposition by substituting lower and upper bounds for each ι_{KL}.

Next, the pointwise bounds on ι_{KL} are derived. The volumes of $g(K) \cap L$ intersected with $g(A)$ and B are denoted by

$$\alpha_{KL}(g) = \mathrm{vol}\,(g(A) \cap g(K) \cap L), \tag{4.47}$$

$$\beta_{KL}(g) = \mathrm{vol}\,(B \cap g(K) \cap L). \tag{4.48}$$

The volume of intersection of $g(K)$ and L is abbreviated as

$$\gamma_{KL}(g) = \mathrm{vol}\,(g(K) \cap L). \tag{4.49}$$

The bounds for ι_{KL} are defined in terms of bounds for α_{KL} and β_{KL}. These bounds for α_{KL} and β_{KL} are defined below. For each affine transformation g,

let δ_g denote the absolute value of the Jacobi determinant of g in all points. This equals the ratio of volumes of the image of a set under g and the original set. Observe that $\operatorname{vol}(g(A) \cap g(K)) = \delta_g \operatorname{vol}(A \cap K)$. The bounds for $\alpha_{KL}(g)$ and $\beta_{KL}(g)$ are

$$\begin{aligned}
\underline{\alpha}_{KL}(g) &= \max\{\, 0, \delta_g \operatorname{vol}(A \cap K) + \gamma_{KL}(g) - \delta_g \operatorname{vol}(K) \,\}, && (4.50)\\
\overline{\alpha}_{KL}(g) &= \min\{\, \gamma_{KL}(g), \delta_g \operatorname{vol}(A \cap K) \,\}, && (4.51)\\
\underline{\beta}_{KL}(g) &= \max\{\, 0, \operatorname{vol}(B \cap L) + \gamma_{KL}(g) - \operatorname{vol}(L) \,\}, && (4.52)\\
\overline{\beta}_{KL}(g) &= \min\{\, \gamma_{KL}(g), \operatorname{vol}(B \cap L) \,\}. && (4.53)
\end{aligned}$$

Finally, bounds for ι_{KL} are defined in terms γ_{KL} and the bounds for α_{KL} and β_{KL}. The lower bound $\underline{\iota}_{KL}$ is derived by

$$\begin{aligned}
\iota_{KL}(g) &= \operatorname{vol}((g(A) \cap g(K) \cap L) \triangle (B \cap g(K) \cap L))\\
&\geq |\operatorname{vol}(g(A) \cap g(K) \cap L) - \operatorname{vol}(B \cap g(K) \cap L)|\\
&= |\alpha_{KL}(g) - \beta_{KL}(g)|\\
&= \max\{\, 0, \alpha_{KL}(g) - \beta_{KL}(g), \beta_{KL}(g) - \alpha_{KL}(g) \,\}\\
&\geq \max\{\, 0, \underline{\alpha}_{KL}(g) - \overline{\beta}_{KL}(g), \underline{\beta}_{KL}(g) - \overline{\alpha}_{KL}(g) \,\}\\
&= \underline{\iota}_{KL}(g).
\end{aligned} \tag{4.54}$$

The upper bound $\overline{\iota}$ is derived by

$$\begin{aligned}
\iota_{KL}(g) &= \operatorname{vol}((g(A) \cap g(K) \cap L) \triangle (B \cap g(K) \cap L))\\
&\leq \min\{\, \gamma_{KL}(g), \alpha_{KL}(g) + \beta_{KL}(g), 2\gamma_{KL}(g) - (\alpha_{KL}(g) + \beta_{KL}(g)) \,\}\\
&\leq \min\{\, \gamma_{KL}(g), \overline{\alpha}_{KL}(g) + \overline{\beta}_{KL}(g), 2\gamma_{KL}(g) - (\underline{\alpha}_{KL}(g) + \underline{\beta}_{KL}(g)) \,\}\\
&= \overline{\iota}_{KL}(g).
\end{aligned} \tag{4.55}$$

Substitution of $\underline{\iota}_{KL}$ and $\overline{\iota}_{KL}$ for ι_{KL} in the decomposition of Equation 4.44, results in the desired pointwise bounds $\underline{f}$ and $\overline{f}$, respectively. The bounds have a number of symmetries: Replacing both A and B with their respective complements leads to equivalent bounds. Exchanging the roles of A and $\mathcal{K}$ with B and $\mathcal{L}$ results in pointwise bounds for $f \circ \cdot^{-1}$, where $\cdot^{-1} : G \to G$ is the inversion mapping.

It is useful to recall the objectives. The method of this section is an application of the GBB algorithm of Section 4.1. This means that the three fundamental operations `l`, `u`, and `cover` need to be defined. The first two of these follow by Equations 4.38 and 4.39 using the bounds $\underline{f}$ and $\overline{f}$. This means the operation `cover` must still be specified. This is done below.

The operation `cover` is defined in terms of the collections $\mathcal{K}_C$ or $\mathcal{L}_C$. For simplicity, it is assumed that each of $\mathcal{K}_C$ or $\mathcal{L}_C$ has a unique maximum volume element (this can be achieved by choosing the maximum volume element whose lowest left corner coordinates are lexicographically smallest). The collection of cells that results from applying the operation `cover` on a cell C is computed as follows. First, select the largest volume interval of $\mathcal{K}_C^+ - \mathcal{K}_C^{\rightarrow}$. Replace this interval with the 2^k subintervals obtained by splitting in the middle of the interval for each dimension. Do the same for $\mathcal{L}_C^+ - \mathcal{L}_C^{\leftarrow}$. Now, determine which intersection relations are possible between the two resulting interval collections for g in $G(C)$. Each of these relations together with the two new partitions forms a cell of $\texttt{cover}(C)$. This finishes the description of the third and final fundamental operation `cover`.

Uniform convergence

At this point, the operations `l`, `u` and `cover` are fully defined. To show that Algorithm 2 terminates if it is based on these operations, it is sufficient to prove that uniform convergence holds (Theorem 4.1.1). Below, it is shown that the difference of bounds on σ_C can be bounded from above using an expression that depends only on the number of iterations, and not on some particular path in the cell hierarchy. First, the bounding expression is given. This is followed by two lemmas. The first lemma proves that the expression is indeed an upper bound for each number of iterations. The second lemma shows that the expression converges to zero if the number of iterations increases. Finally, a theorem proves that `l`, `u` and `cover` converge uniformly using the two lemmas.

Use the notation $\tilde{\tau}$ to be the difference $\overline{\tau} - \underline{\tau}$, where τ is either ι_{KL}, σ_C or f. A pointwise upper bound for $\tilde{\sigma}_C$ is introduced as follows. First, recursively define $\mathcal{K}_C[t]$ and $\mathcal{L}_C[t]$. Set $\mathcal{K}_C[0] = \mathcal{K}_C$. Define $\mathcal{K}_C[t+1]$ to be $\mathcal{K}_C[t]$ where the maximum volume element of $\mathcal{K}_C^+$ is replaced with equally-sized k-dimensional intervals in $\mathcal{K}_C$. Define $\mathcal{L}_C[t]$ analogous.

The pointwise bound on $\tilde{\sigma}_C$ is defined in terms of $\mathcal{K}_C[t]$ and $\mathcal{L}_C[t]$ by

$$\tilde{\epsilon}(t, g) = \delta_g \operatorname{vol}\left(\bigcup \mathcal{K}_C^+[t]\right) + \operatorname{vol}\left(\bigcup \mathcal{L}_C^+[t]\right). \tag{4.56}$$

The following lemma states that $\tilde{\epsilon}(t, g)$ is indeed a pointwise upper bound for $\tilde{\sigma}_{C_t}$.

Lemma 4.3.1. *For each infinite sequence of cells $(C_t)_{t=0}^{\infty}$ satisfying $C_{t+1} \in \texttt{cover}(C_t)$ for all $t \geq 1$, the inequality $\tilde{\sigma}_{C_t}(g) \leq \tilde{\epsilon}(t, g)$ holds for $g \in G(C_t)$.*

Proof. Life is made much easier by the abbreviations

$$\mathcal{K}_{C_t}^{\iota} = \mathcal{K}_{C_t}^+ - \mathcal{K}_{C_t}^{\rightarrow}, \qquad \mathcal{L}_{C_t}^{\iota} = \mathcal{L}_{C_t}^+ - \mathcal{L}_{C_t}^{\leftarrow}. \tag{4.57}$$

The proof is based on the following two claims:

1. For each $K \in \mathcal{K}^{\iota}_{C_t}$ there is a $K^* \in \mathcal{K}^{+}_{C}[t]$ such that $K \subseteq K^*$.

2. For each $L \in \mathcal{L}^{\iota}_{C_t}$ there is a $L^* \in \mathcal{L}^{+}_{C}[t]$ such that $L \subseteq L^*$.

Because of symmetry, it is sufficient to prove the first claim. This is done using induction. For $t = 0$, the claim holds trivially. Now, the induction step is made. Suppose the claim holds for t. It is shown that the claim holds for $t + 1$ using contradiction. Assume that the claim does not hold for $t+1$. This means there is a $K \in \mathcal{K}^{\iota}_{C_{t+1}}$ such that $\mathcal{K}^{+}_{C}[t + 1]$ contains no superset of K. Let K' be the element of $\mathcal{K}_{C_t}$ containing K. It is not difficult to see that $K' \in \mathcal{K}^{\iota}_{C_t}$ By the induction hypothesis, there is a $K'' \in \mathcal{K}^{+}_{C}[t]$ that contains K'. Our assumption that $\mathcal{K}^{+}_{C}[t + 1]$ contains no superset of K implies that $K = K' = K''$ and K'' is the maximum volume element of $\mathcal{K}^{+}_{C}[t]$. As a consequence, K' must be the maximum volume element of $\mathcal{K}^{\iota}_{C_t}$. By the definition of the operation `cover`, K is not an element of $\mathcal{K}^{\iota}_{C_{t+1}}$. Contradiction.

It follows from the first claim that the union of $\mathcal{K}^{\iota}_{C_t}$ is contained in the union of $\mathcal{K}^{+}_{C}[t]$. From the second claim it follows that the union of $\mathcal{L}^{\iota}_{C_t}$ is contained in the union of $\mathcal{L}^{+}_{C}[t]$.

These observations are now used to derived the inequality $\tilde{\sigma}_{C_t}(g) \leq \tilde{\epsilon}(t, g)$ for $g \in G(C_t)$. The expression $\tilde{\sigma}_{C_t}(g)$ is the sum of $\tilde{\iota}_{KL}(g)$ over all (K, L) in $\mathcal{R}^{\iota}_{C_t}$. This is bounded from above by the sum of $\tilde{\iota}_{KL}(g)$ over all (K, L) in the union of $\mathcal{K}^{\iota}_{C_t} \times \mathcal{L}_{C_t}$ and $\mathcal{K}_{C_t} \times \mathcal{L}^{\iota}_{C_t}$. Over the first of these two collections the sum is bounded by the first term of $\tilde{\epsilon}(t, g)$ in Equation 4.56:

$$\begin{aligned}\sum_{(K,L)\in\mathcal{K}^{\iota}_{C_t}\times\mathcal{L}_{C_t}} \tilde{\iota}_{KL}(g) &\leq \sum_{(K,L)\in\mathcal{K}^{\iota}_{C_t}\times\mathcal{L}_{C_t}} \gamma_{KL}(g) \\ &= \text{vol}\left(g(\bigcup \mathcal{K}^{\iota}_{C_t}) \cap \bigcup \mathcal{L}_{C_t}\right) \\ &\leq \text{vol}\left(g(\bigcup \mathcal{K}^{\iota}_{C_t})\right) \\ &\leq \text{vol}\left(g(\bigcup \mathcal{K}^{+}_{C}[t])\right).\end{aligned} \tag{4.58}$$

Similarly, it is found that the second term of $\tilde{\epsilon}(t, g)$ in Equation 4.56 is an upper bound for the sum of $\tilde{\iota}_{KL}(g)$ over all (K, L) in $\mathcal{K}_{C_t} \times \mathcal{L}^{\iota}_{C_t}$:

$$\sum_{(K,L)\in\mathcal{K}_{C_t}\times\mathcal{L}^{\iota}_{C_t}} \tilde{\iota}_{KL}(g) \leq \text{vol}\left(\bigcup \mathcal{L}^{+}_{C}[t]\right). \tag{4.59}$$

The proof is finished by summing the two inequalities. □

The class $\mathcal{C}$ is restricted to cells C for which the value of δ_g over g in $G(C)$ is bounded.

Theorem 4.3.2. *Let the cell bounds* $\mathtt{l}$ *and* $\mathtt{u}$ *be defined in terms of Equations 4.38 and 4.39, respectively. These cell bounds converge uniformly.*

Proof. It must be shown that for each feasible cell C in $\mathcal{C}$, and each $\epsilon > 0$, there is an r such that $\mathtt{u}(C') - \mathtt{l}(C') < \epsilon$ for all feasible C' in $\mathtt{cover}^r(C)$.

Observe that the difference between $\mathtt{u}(C)$ and $\mathtt{l}(C)$ is bounded by the maximum difference between $\overline{f}_C(g)$ and $\underline{f}_C(g)$ over $g \in G(C)$. Recall that $\overline{f}_C$ and $\underline{f}_C$ are defined by substitution of $\overline{\sigma}_C$ and $\underline{\sigma}_C$, respectively. This means it is sufficient to prove that for each feasible $C \in \mathcal{C}$ and $\epsilon > 0$, there is an $r \geq 1$ such that for all feasible C' in $\mathtt{cover}^r(C)$, and all $g \in G(C')$, $\overline{\sigma}_{C'}(g) - \underline{\sigma}_{C'}(g) < \epsilon$. The desired result follows by continuity of f_C as a function of σ_C (like in Equation 4.44).

Let γ be an upper bound for δ_g over $g \in G(C)$. Let $\epsilon > 0$ be given. Since A and B are finite unions of k-simplices, $r \geq 1$ may be chosen so that

$$\mathrm{vol}\left(\bigcup \mathcal{K}_C^+[r]\right) < \epsilon/(2\gamma), \qquad \mathrm{vol}\left(\bigcup \mathcal{L}_C^+[r]\right) < \epsilon/2. \tag{4.60}$$

Now, the result follows directly from Lemma 4.3.1. □

4.4 Experimental results

The first of the two specialisations of the GBB algorithm discussed in this chapter is the traces approach. I implemented the traces approach for two different similarity measures, namely, the directed Hausdorff distance $\mathsf{h}^{\triangleright}$ and the directed absolute difference $\mathsf{a}^{\triangleright}$ [71]. These similarity measures are applied to perform partial pattern matching (explained in Section 1.1). In this section, some experimental results obtained using the implementations are given.

The directed Hausdorff distance $\mathsf{h}^{\triangleright}$ is defined by Equation 4.32. The directed absolute difference is given by

$$\mathsf{a}^{\triangleright}(A, B) = \int_{\mathbb{R}^k} 0 \vee (\boldsymbol{n}_A(x) - \boldsymbol{n}_B(x))\, dx. \tag{4.61}$$

The relations between the directed distances and the "undirected" distances are $\mathsf{h}(A, B) = \max\{\, \mathsf{h}^{\triangleright}(A, B), \mathsf{h}^{\triangleright}(B, A) \,\}$ and $\mathsf{a}(A, B) = \mathsf{a}^{\triangleright}(A, B) + \mathsf{a}^{\triangleright}(B, A)$.

Figures 4.3–4.7 show the inputs and the outcomes of the experimental results. In each figure, the upper row contains the two original grey-scale images. Using image processing, geometric patterns A and B are obtained. These patterns are superimposed upon the images. The white dots in the upper left subfigure form the pattern A, while the black dots in the upper right subfigure

form the pattern B. The middle and lower rows contain matches. The matches on the left side in the middle and lower row were obtained by minimisation of the directed Hausdorff distance $\mathsf{h}^{\triangleright}$. The matches on the right side were obtained using the directed absolute difference $\mathsf{a}^{\triangleright}$. The middle row shows the image of the pattern A under the minimising transformation, superimposed on the pattern B. The lower row shows the grey scale image corresponding to A, transformed under the same transformation, averaged with the grey scale image corresponding to B. The transformation is always found using the feature patterns A and B only, and then applied to the grey scale images to allow visual verification.

Figure 4.3 shows a plane with markings. The markings were cut and cropped manually, resulting in the subimage shown in the top left subfigure. The original image of the plane shown in the top right subfigure. Figures 4.4 and 4.5 depict three-view drawings of plane designs. In each of the two figures, the top left images are rotated and translated versions of the top view from the right image. Figure 4.6 shows two frontal views of the same plane taken from different distances. Figure 4.7 shows two images of the Mir space station obtained from viewpoints with different orientations. The image in the top right subfigure contains a part that is not visible in the image of the top left subfigure.

In some cases the patterns consists of edges, in others they consist of corners. In Figures 4.3–4.7 the patterns are obtained using edge detection and thresholding. In Figure 4.7 the patterns are obtained using manual corner detection. The patterns are presented to the directed Hausdorff matching algorithm as point patterns. For matching under the directed absolute difference, neighbourhoods of the point sets were formed. More precisely, each point in a pattern is converted to a square. The input patterns for the directed absolute difference are formed by the unions of the squares.

Table 4.1 shows, for each test, the transformation group which was used for matching, and the sizes of the patterns. The table contains test statistics for both the directed Hausdorff distance $\mathsf{h}^{\triangleright}$ and the directed absolute difference $\mathsf{a}^{\triangleright}$. This table is included only to give an indication of the number of cells and the processing times. Comparing the statistics has limited meaning since, in principle, the implementations solve two different problems.

Implementation details

Both the Hausdorff metric version and the absolute difference version of the traces approach were implemented in C++. Both programs were compiled using the SGI Delta/C++ compiler. The tests were executed on a Silicon Graphics Indy workstation having a MIPS R5000 processor and 64 MB of memory.

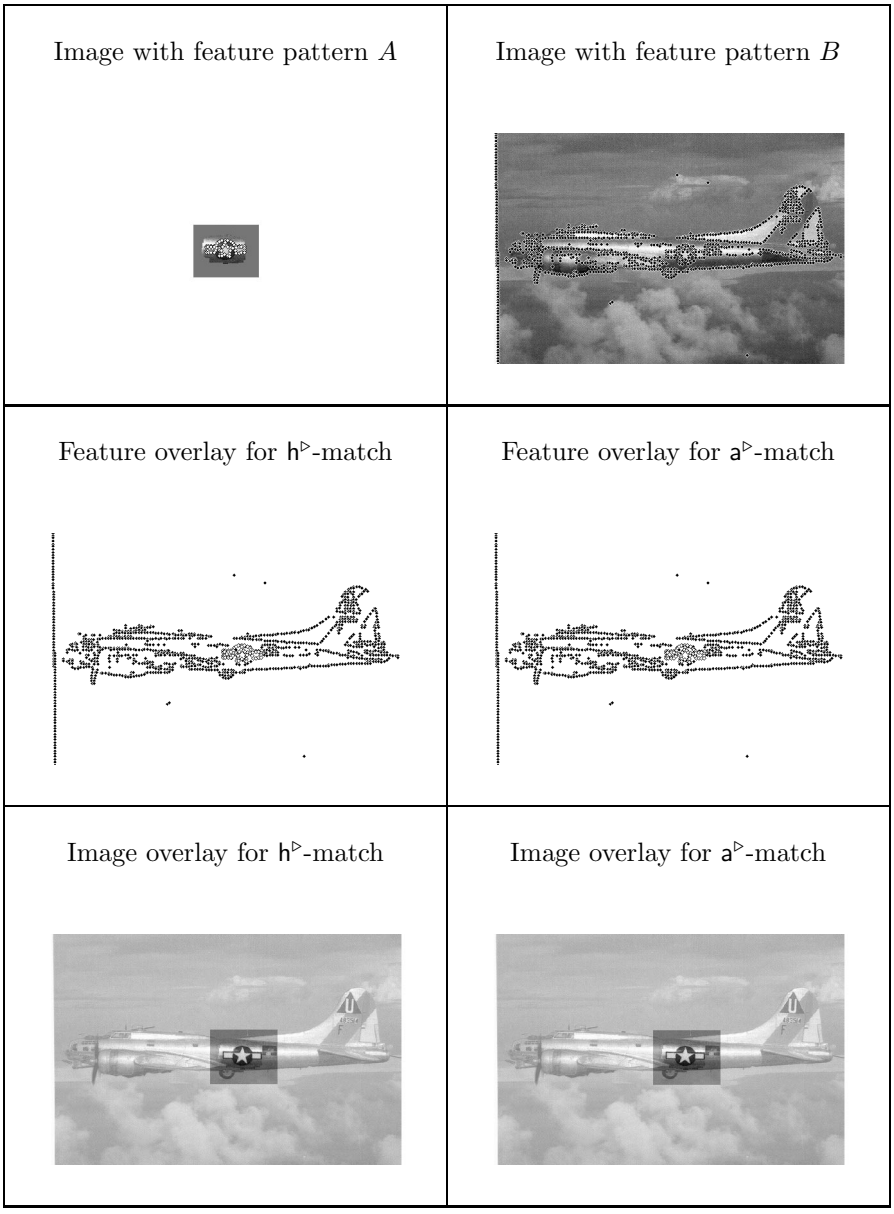


Figure 4.3: Image with a translated subimage

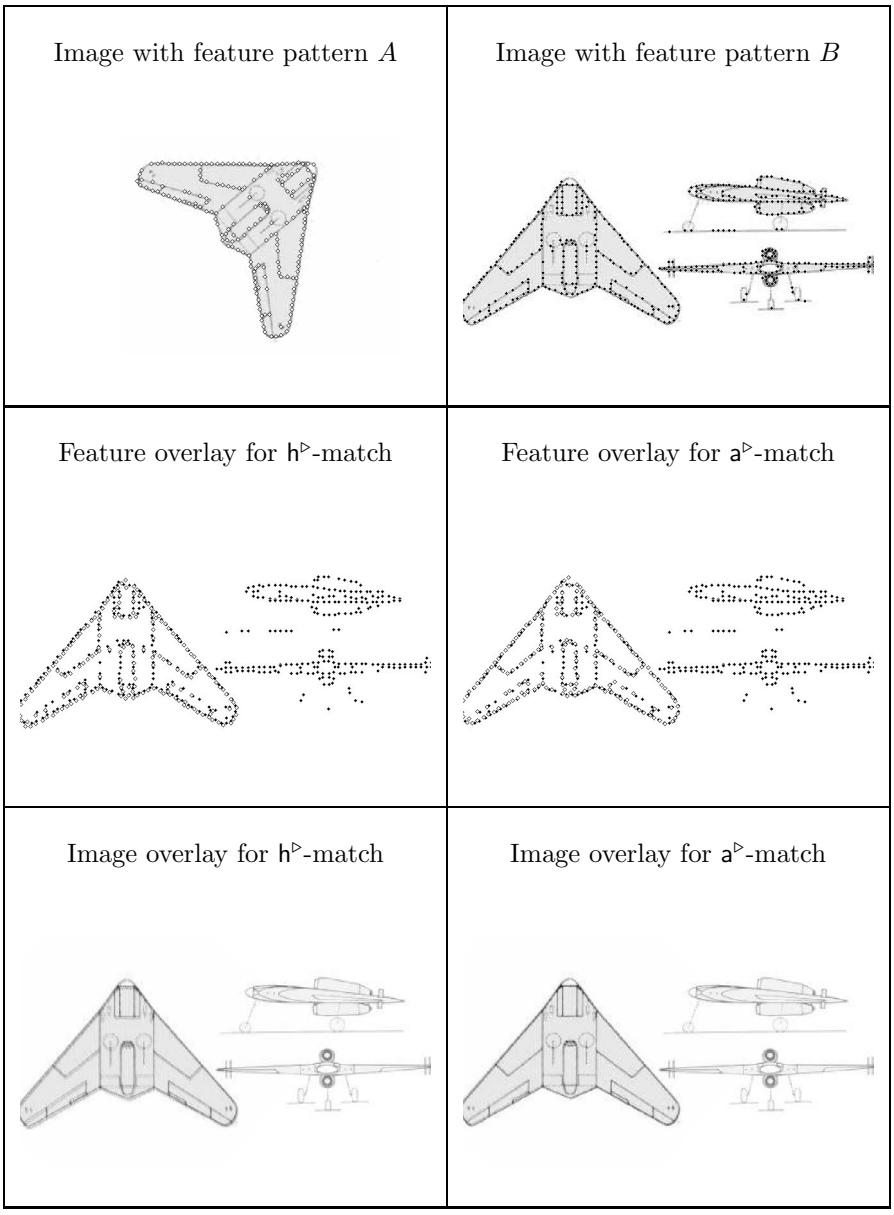


Figure 4.4: Image with a rotated and translated subimage

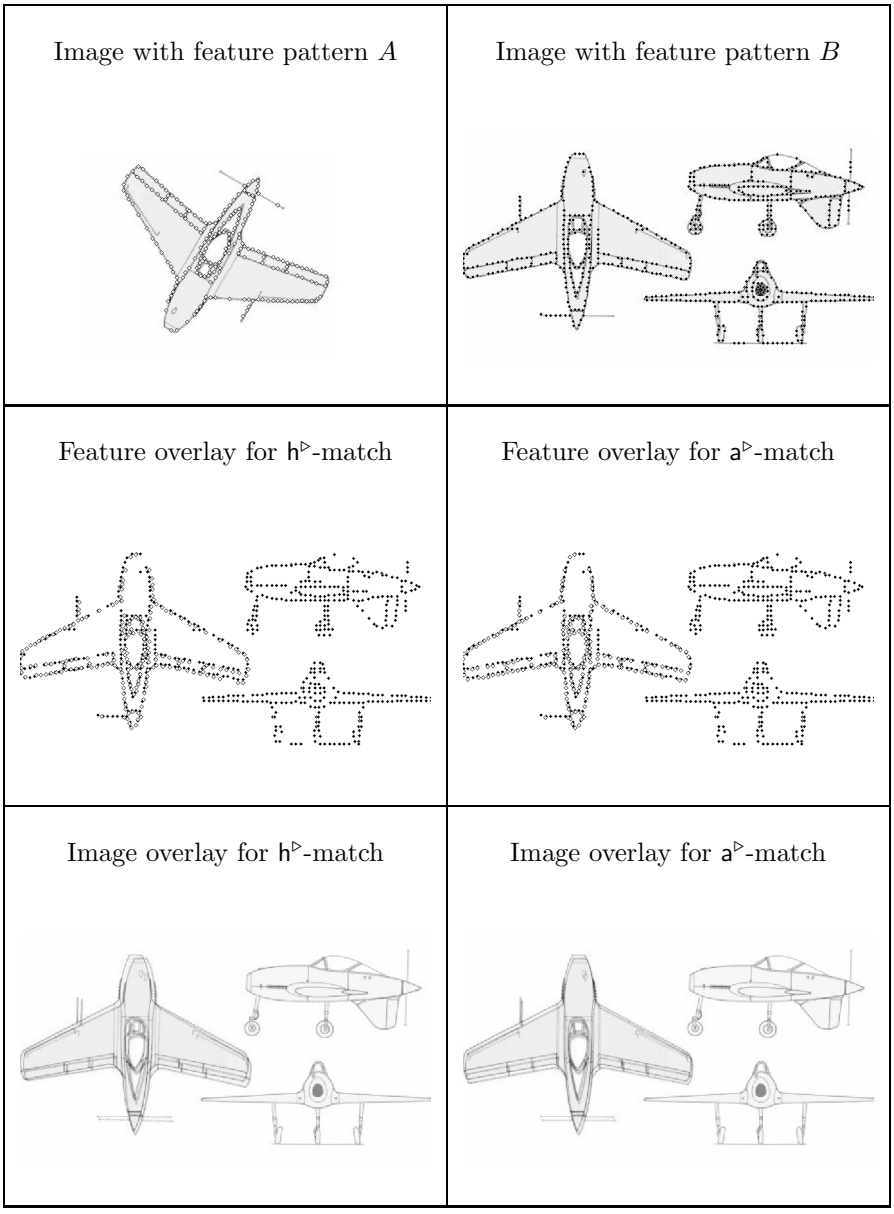


Figure 4.5: Image with a rotated and translated subimage

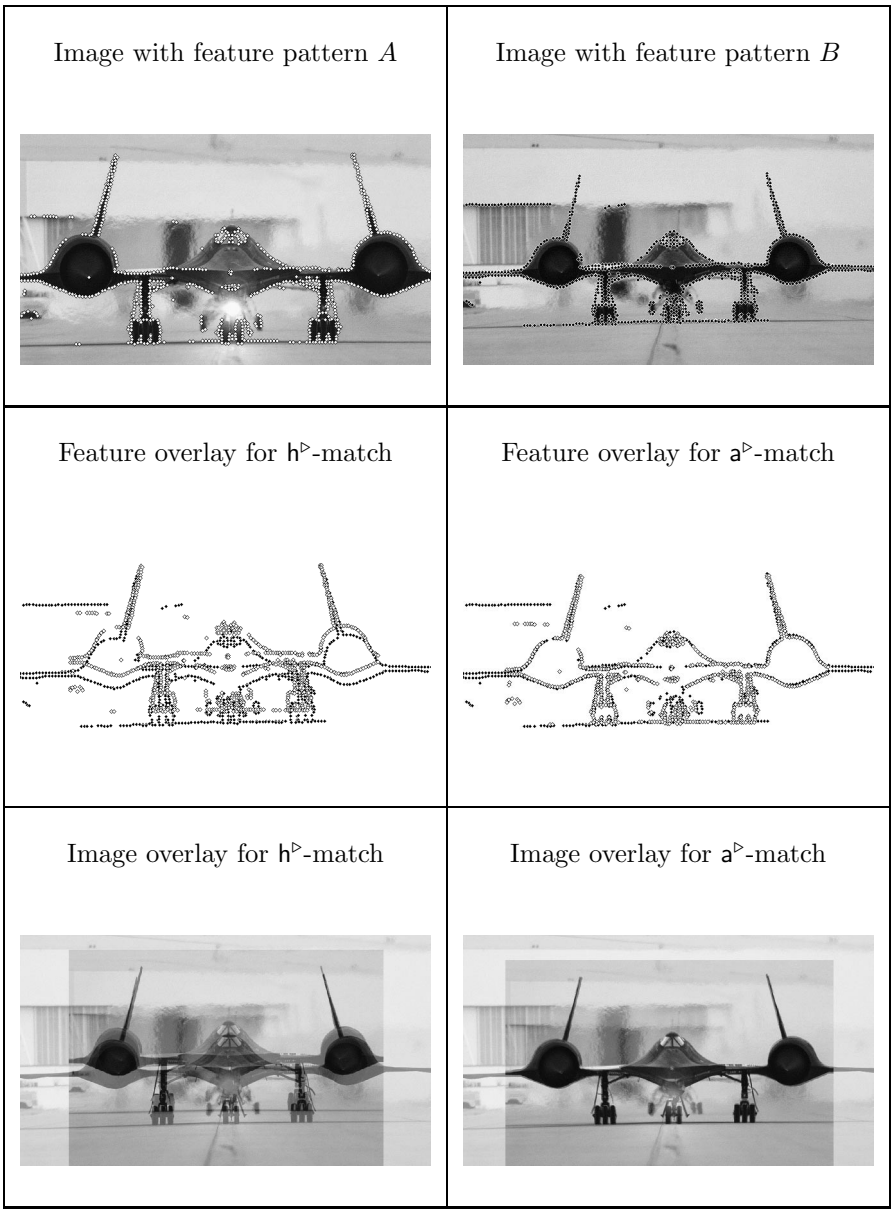


Figure 4.6: Two different frontal views of a plane

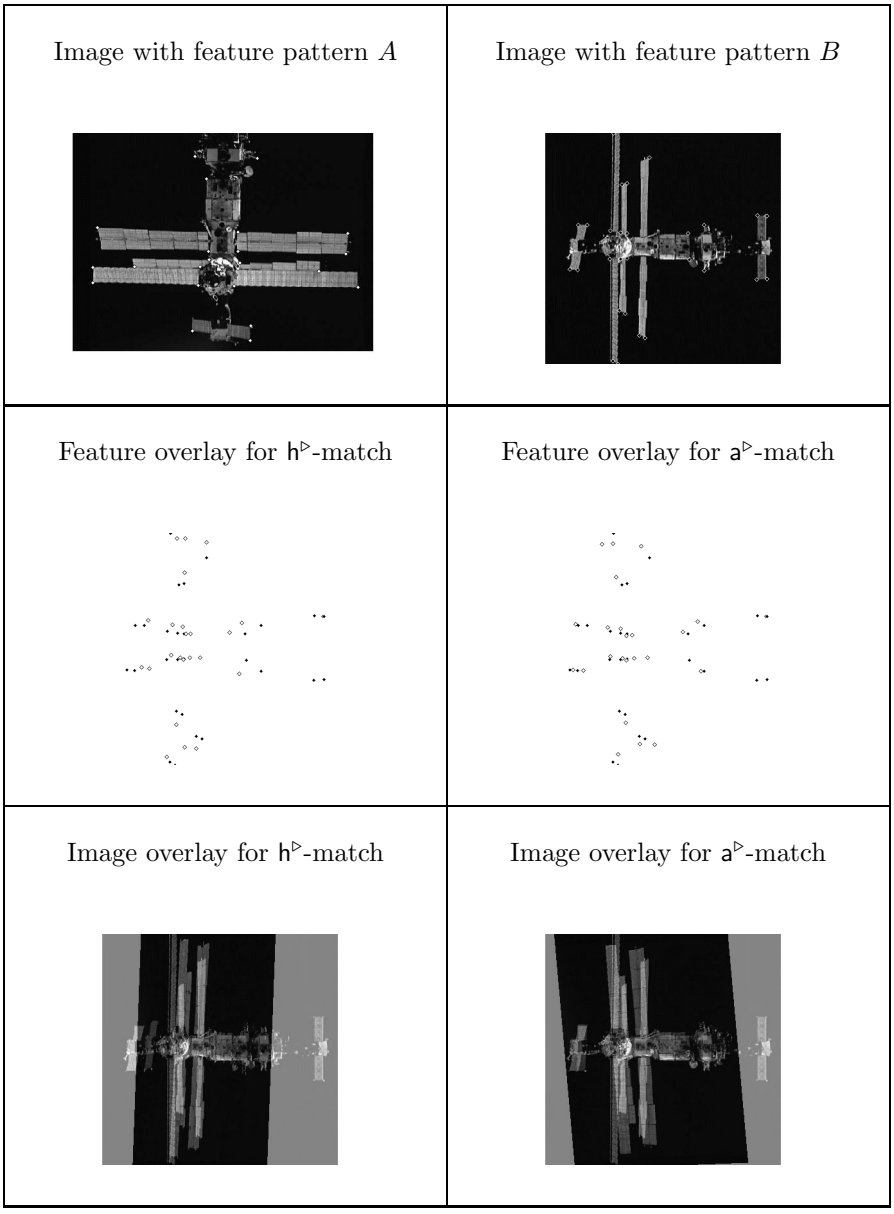


Figure 4.7: Two different views of the Mir space station

Test figure	Transformation group	Size of A	Size of B
4.3	Lat^2	60	1084
4.4	Iso^2	189	439
4.5	Iso^2	203	637
4.6	Stret^2	717	740
4.7	Af^2	23	29

Table 4.1: Overview of the test input

Test figure	Cell evaluations		Total time (sec.)	
	$\mathsf{h}^{\triangleright}$	$\mathsf{a}^{\triangleright}$	$\mathsf{h}^{\triangleright}$	$\mathsf{a}^{\triangleright}$
4.3	573	791	53.8	7.2
4.4	641	7275	127.4	233.1
4.5	961	29137	250.2	1020.0
4.6	641	8413	568.2	533.5
4.7	375673	177997	2928.0	290.4

Table 4.2: Statistics for the test results

4.5 Discussion

The GBB algorithm is designed to be efficient if the global minima are contained in a limited number of small regions. The difference with hill climbing techniques is that the GBB algorithm cannot get stuck in local minima. The GBB algorithm guarantees a solution that is strictly within the specified accuracy. It was proven that the instances of the GBB algorithm given in this chapter find a solution in finite time for every input.

The efficiency of a GBB algorithm depends much on the tightness of the cell bounds. It is hard to say something about the computational complexity of instances of the GBB algorithm in terms of the input size. This is not surprising, since the flow of control of the GBB algorithm is independent of the representations of the input patterns. The flow of control only depends on properties of the geometric properties of the patterns as subsets of a space.

In the previous sections, two applications of the GBB algorithm were discussed: the traces approach and the partition combination approach. The traces approach has been implemented for the Hausdorff metric and another similarity measure. Experimental results obtained using these implementations were presented. The execution times in these experiments are good, considering the complexity of the problems. An implementation of the partition com-

bination approach is under construction. I expect the partition combination approach to be faster than the traces approach because the former approach works at different levels of detail and has good cell bounds at the same time.

Appendix A

Point set topology

The concepts from point set topology that are used in this thesis are summarised below. The definitions and results are adapted from Kelley [89], Hocking and Young [79], and Munkres [101].

A topology on a set S is a collection $\mathcal{T}$ of subsets of S having the following properties.

1. The empty set and S are in $\mathcal{T}$.
2. The intersection of each finite subcollection of $\mathcal{T}$ is an element of $\mathcal{T}$.
3. The union of each subcollection of $\mathcal{T}$ is an element of $\mathcal{T}$.

If $\mathcal{T}$ is a topology on S, then $(S, \mathcal{T})$ is a topological space. If the topology on S is understood, it is said that S is a topological space.

A subset U of S is defined to be open if U is an element of the topology of S. A neighbourhood of a point x in a topological space S is a subset of S that contains an open set containing x.

Consider two topologies $\mathcal{T}$ and $\mathcal{T}'$ on the same set. If $\mathcal{T}$ is a subcollection of $\mathcal{T}'$ it is said that $\mathcal{T}'$ is finer than $\mathcal{T}$ and $\mathcal{T}$ is coarser than $\mathcal{T}'$.

A basis for a topology on S is a collection $\mathcal{B}$ of subsets of S having the following properties.

1. For each point x in S, there is an element of $\mathcal{B}$ containing x.
2. If $x \in B_1 \cap B_2$ for some $B_1, B_2 \in \mathcal{B}$, then $x \in B_3 \subseteq B_1 \cap B_2$ for some $B_3 \in \mathcal{B}$.

The topology generated by a basis $\mathcal{B}$ consists of all finite unions of elements of $\mathcal{B}$. A subbasis for S is a collection of subsets of S whose union equals S.

The topology generated by a subbasis $\mathcal{S}$ is the collection of all unions of finite intersections of elements of $\mathcal{S}$. This is the coarsest topology containing $\mathcal{S}$.

A function $f : S \to T$ for topological spaces S and T is continuous if for each open subset V of T, the inverse image $f^{-1}(V)$ is open in S. The function f is continuous at the point x in S if for each open neighbourhood V of $f(x)$, there is a open neighbourhood U of x such that $f(U)$ is contained in V. The function f is continuous if and only if it is continuous at each point in S. A function $f : X \to Y$ is called a homeomorphism if f is a bijection and both f and f^{-1} are continuous.

If T is a subset of a topological space S, then the subspace topology for T consists of all intersections of open sets of S with T. In this topology, it is said that T is a (topological) subspace of S. If S is a subspace of T, then the topology on S is called the relative topology induced by T.

Given two topological spaces S and T, the product topology for $S \times T$ has as a basis the collection of all $U \times V$ where U is open in S and V is open in T.

Let S be a topological space and let $f : S \to T$ be a function. Define $\mathcal{Q}$ as the collection of all subsets Q of T such that $f^{-1}(Q)$ is open in S. This the quotient topology for T relative to f and S. It is the finest topology for which the mapping f is continuous.

A subset C of a topological space S is said to be closed if the complement $S - C$ is open in S. A point x of S is called a limit point of a subset P of S if every open neighbourhood of x intersects P in a point other than x. A subset of S is closed if and only if it contains all its limit points.

The interior of a subset P of S is the union of all open sets of S contained in P. The closure of a subset P of S is the intersection of all closed sets of S containing P. A subset P of S is said to be dense in S if $\mathrm{Cl}\,(P) = S$. The boundary of P in S consists of all points that are in the closure of P and not in the interior of P. A point x in S is in the boundary of P if and only if each open neighbourhood of x intersects both P and $S - P$. The interior, the closure and the boundary of P are denoted by $\mathrm{Int}\,(P)$, $\mathrm{Cl}\,(P)$, and $\mathrm{Bd}\,(P)$, respectively.

A collection $\mathcal{S}$ of subsets of S is a cover for S if the union of $\mathcal{S}$ equals S. An open cover consists only of open sets. A space S is compact if every open cover for S contains a finite subcollection that is also a cover for S. If S is a subset of $\mathbb{R}^k$, then S is compact if and only if S is closed and bounded.

A topological space S is called a Hausdorff space if for each pair of points p and q in S, there are disjoint open neighbourhoods U and V, respectively.

Let S and T be topological spaces and let $\mathcal{C}(S,T)$ denote the set of all continuous functions mapping S into T. The set of functions $\mathcal{C}(S,T)$ is endowed with a topology as follows. For each compact subset C of S and each open subset U of T, define the subset $S(C,U)$ of $\mathcal{C}(S,T)$ by

$$S(C,U) = \{\, f \mid f \in \mathcal{C}(X,Y) \text{ and } f(C) \subseteq U \,\}.$$

The set of all such $S(C,U)$ is a subbasis for the compact-open topology on $\mathcal{C}(S,T)$. If $\mathcal{C}(S,T)$ has the compact-open topology, then the following function is continuous:

$$e : S \times \mathcal{C}(S,T) \to T : (x,f) \mapsto f(x).$$

This is called the evaluation map.

Let S be a topological space and (T,d) be a pseudometric space. Let T^S be the set of all functions $f : S \to T$. The topology of compact convergence on T^S is given by the basis defined below. Each basis element is a set of functions g in T^S given by

$$\sup\{\, d(f(x),g(x)) \mid x \in C \,\},$$

where f is an element of T^S, C is a compact subset of S and $\epsilon > 0$. The compact-open topology on $\mathcal{C}(S,T)$ and the topology of compact convergence on $\mathcal{C}(S,T)$ (as a subspace of T^S) are the same.

Appendix B

Topological transformation groups

This section summarises the theory of topological transformation groups that is relevant for this thesis. The discussion uses concepts from from point set topology (explained in Appendix A). For more elaborate discussions of topological transformation groups can be found in the books by Montgomery and Zippin [97] and Bredon [32].

A group is a set G so that for each x and y in G there is a unique product xy in G satisfying the following axioms:

1. There is a unique element e_G in G such that $xe_G = x = e_G x$ for all x in G.

2. For each x in G there is an x^{-1} in G such that $xx^{-1} = e_G = x^{-1}x$.

3. For all x, y and z in G, $x(yz) = (xy)z$.

A group G that is also a topological space is called a topological group if $x \mapsto x^{-1}$ is a continuous function from G to G and $(x, y) \mapsto xy$ is a continuous function from $G \times G$ to G.

If G is a group, then a subset H of G is a subgroup of G if H is itself a group (using the same product). The subset H is a subgroup of G if and only if xy^{-1} belongs to H for all x and y in H. A subset H of a If G is a topological group, and H is a subgroup of G, then H is a topological group, where H has the subspace topology.

Each element g of a group G defines two functions from G onto G. The left translation by g is the function $h \mapsto gh$ from G to itself. The right translation

by g is the function $h \mapsto hg^{-1}$ from G to itself. In the case of a topological group, both the left and right translations are continuous functions.

Let G be a group, let X be a set, and let f be a function from $G \times X$ to X whose values are denoted by $g(x) = f(g, x)$. Then, G is a transformation group on X if it satisfies the following axioms:

1. For each x in X, $e_G(x) = x$.
2. For each g and each h in G, $(gh)(x) = g(h(x))$.

In this case, for each x in X, the function $x \mapsto g(x)$ is a bijection from X onto itself.

Let G be a topological group that is a transformation group on a topological space X. Then, G is is a topological transformation group on X if $(g, x) \mapsto g(x)$ is a continuous function from $G \times X$ to X. In this case, for each x in X, the function $x \mapsto g(x)$ is a homeomorphism from X onto itself, Below, several topological transformation groups that are important in this thesis are discussed.

In this thesis, the focus lies mostly on transformation groups that consist of differentiable functions from $\mathbb{R}^k$ onto itself. Concepts are used that are related to the total derivative of a function in a point [51]. A function f from $\mathbb{R}^k$ onto itself is called differentiable in a point $x \in \mathbb{R}^k$ if there exists a linear function l from $\mathbb{R}^k$ onto itself such that

$$\lim_{v \to o, v \neq o} \|f(x+v) - f(x) - l(v)\| / \|v\| = 0 \tag{B.1}$$

where o is the origin of $\mathbb{R}^k$, and v ranges over $\mathbb{R}^k$. If it exists, the linear function l is uniquely defined, and is called the total derivative of f in x, denoted by $\mathrm{D}_f(x)$. A function f is called differentiable if it is differentiable in each point of its domain. Expressed in the standard basis of $\mathbb{R}^k$, the total derivative can be expressed as a $k \times k$ matrix of partial derivatives of coordinate functions of f, denoted by $\partial_j f_i$. The determinant of a linear function l from $\mathbb{R}^k$ onto itself is denoted by $\det(l)$. The Jacobi determinant of a function $f : \mathbb{R}^k \to \mathbb{R}^k$ in a point $x \in \mathbb{R}^k$ is the determinant of the total derivative of f in x.

The group of *diffeomorphisms* in $\mathbb{R}^k$, denoted by Dif^k, consists of all homeomorphisms g on $\mathbb{R}^k$ onto $\mathbb{R}^k$ such that both g and g^{-1} have total derivatives that are continuous as functions of $\mathbb{R}^k$. The subgroup CDif^k of Dif^k consists of all diffeomorphisms for which the Jacobi determinant is constant. These transformations preserve the ratio of volumes of each two sets. The subgroup UDif^k of CDif^k consists of all diffeomorphisms for which the absolute value of the Jacobi determinant equals 1. These transformations are volume preserving.

The affine transformations Af^k from a subgroup of Dif^k, consisting of all diffeomorphisms from $\mathbb{R}^k$ onto itself that can be written as

$$x \mapsto (Lx) + t, \tag{B.2}$$

where L is a $k \times k$ matrix over $\mathbb{R}$ having a nonzero determinant and t is an element of $\mathbb{R}^k$. The matrix L represents a linear transformation, the vector t represents a translation. The affine transformations are a proper subgroup of CDif^k. The affine transformations map simplices to simplices. Several subgroups of the affine transformations are discussed below.

The volume preserving affine transformations, denoted with UAf^k are those affine transformations that can be expressed as in Equation B.2, where L is a matrix whose determinant has absolute value 1. The volume-preserving affine transformations are a proper subgroup of UDif^k.

The following affine subgroup has no official name, therefore it is named here: the stretch transformations, denoted by Stret^k, are affine transformations for which the matrix L in Equation B.2 is restricted to be a diagonal matrix, that is, a matrix in which elements off the diagonal are zero. The transformations map k-dimensional intervals onto k-dimensional intervals.

The group of *homotheties* (or homothetic transformations), denoted Thet^k, is a subgroup of the stretch transformations. It consists of all affine transformations as in Equation B.2, where L is a diagonal matrix in which all nonzero elements have the same value. Such matrices represent uniform scaling. Under each homothety, the image of each line is parallel to the original line.

The similarity transformations, denoted Sim^k, consist of all affine transformations as in Equation B.2, where L is a scalar multiple of an orthogonal matrix. These transformations preserve angles.

The (Euclidean) *isometries* in $\mathbb{R}^k$, denoted by Iso^k, consist of affine transformations as in Equation B.2, where L is an orthogonal matrix, that is, $LL^T = I$. Here, L^T denotes the transpose of the matrix L. The isometries preserve Euclidean distances in $\mathbb{R}^k$.

The *translations* Lat^k are all affine transformations that can be put in Equation B.2, where L is the identity. These transformations preserve vectors between points.

Finally, the *identity group* Id^k (sometimes called trivial group) is the group consisting only of the identity transformation $x \mapsto x$ from $\mathbb{R}^k$ to itself.

Bibliography

[1] P. K. Agarwal and M. Sharir. On the number of views of polyhedral terrains. *Discrete & Computational Geometry*, 12:177–182, 1994.

[2] P. K. Agarwal and M. Sharir. Arrangements and their applications. In *Handbook of Computational Geometry*. North Holland, Amsterdam, 1998.

[3] P. K. Agarwal, M. Sharir, and S. Toledo. Applications of parametric searching in geometric optimization. *Journal of Algorithms*, 17(3):292–318, November 1994.

[4] A. V. Aho, J. E. Hopcroft, and J. D. Ullman. *The Design and Analysis of Computer Algorithms*. Addison-Wesley, Reading, Mass., 1974.

[5] T. Akutsu. On determining the congruence of point sets in d dimensions. *Computational Geometry: Theory and Applications*, 9(4):247–256, March 1998.

[6] H. Alt, O. Aichholzer, and G. Rote. Matching shapes with a reference point. *International Journal of Computational Geometry and Applications*, 7:349–363, 1997.

[7] H. Alt, B. Behrends, and J. Blömer. Approximate matching of polygonal shapes. *Annals of Mathematics and Artificial Intelligence*, 13:251–266, 1995.

[8] H. Alt, J. Blömer, M. Godeau, and H. Wagener. Approximation of convex polygons. In *Proceedings 17th International Colloq. Automata Lang. Program.*, volume 443 of *Lecture Notes in Computer Science*, pages 703–716. Springer-Verlag, 1990.

[9] H. Alt, U. Fuchs, G. Rote, and G. Weber. Matching convex shapes with respect to the symmetric difference. *Algorithmica*, 21(1):89–103, 1998.

[10] H. Alt and M. Godau. Computing the Fréchet distance between two polygonal curves. *International Journal of Computational Geometry and Applications*, 5:75–91, 1995.

[11] H. Alt and L. Guibas. Discrete geometric shapes: Matching, interpolation, and approximation. In J. Sack and J. Urrutia, editors, *Handbook of Computational Geometry*, pages 121–153. North Holland, Amsterdam, 1999.

[12] H. Alt, K. Mehlhorn, H. Wagener, and E. Welzl. Congruence, similarity and symmetries of geometric objects. *Discrete & Computational Geometry*, 3:237–256, 1988.

[13] M. Alzina, W. Szpankowski, and A. Grama. 2D-Pattern matching image and video compression. internet. http://www.cs.purdue.edu/homes/spa/Compression/2D-PMC.html.

[14] N. Amenta. Bounded boxes, Hausdorff distance, and a new proof of an interesting Helly theorem. In *10th Annual ACM Symposium on Computational Geometry*, pages 340–347. ACM Press, 1994.

[15] M. A. Armstrong. *Groups and Symmetry*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, N. Y., 1998.

[16] D. Arnaud and W. Szpankowski. Pattern natching image compression with prediction loop: preliminary and experimental results. In *Proceedings Data Compression Conference*, 1996. also Purdue University CSD-TR-96-069.

[17] B. Aronov, L. J. Guibas, M. Teichmann, and L. Zhang. Visibility queries in simple polygons and applications. In *International Symposium on Algorithms and Computation*, pages 357–366, 1998.

[18] M. Atallah, Y. Génin, and W. Szpankowski. A pattern matching image compression. In *Proceedings International Conference on Image Processing*, volume 2, pages 349–352, 1996. also Purdue University CSD-TR-95-083.

[19] M. J. Atallah. A linear time algorithm for the Hausdorff distance between convex polygons. *Information Processing Letters*, 17:207–209, 1983.

[20] M. J. Atallah. Checking similarity of planar figures. *International Journal of Computer and Information Sciences*, 13(4):279–290, August 1984.

[21] M. D. Atkinson. An optimal algorithm for geometrical congruence. *J. Algorithms*, 8:159–172, 1987.

[22] A. J. Baddeley. An error metric for binary images. In W. Förstner and S. Ruwiedel, editors, *Robust Computer Vision: Quality of Vision Algorithms - Proceedings International Workshop on Robust Computer Vision*, pages 59–78, 1992.

[23] A.J. Baddeley. Errors in binary images and an L^p version of the Hausdorff metric. *Nieuw Archief voor Wiskunde*, 10:157–183, 1992.

[24] D. H. Ballard. Generalizing the hough transforming to detect arbitrary shapes. *Pattern recognition*, 13:111–122, 1981.

[25] H. G. Barrow, J. M. Tenenbaum, R. C. Bolles, and H. C. Wolf. Parametric correspondence and chamfer matching: Two new techniques for image matching. In *Proc. Int. Joint Conference on Artificial Intelligence*, volume 2, pages 659–663, 1977.

[26] M. J. Bishop and E. A. Thompson. Maximum likelihood alignment of DNA sequences. *Journal of Molecular Biology*, 190:159–165, 1986.

[27] I. Bloch. On fuzzy distances and their use in image processing under imprecision. *Pattern Recognition*, 32:1873–1895, 1999.

[28] R. C. Bolles and P. Horaud. 3dpo: a three-dimensional part orientation system. *The International Journal of Robotics Research*, 5(3):3–26, 1986. 3DPO system, see Sueten, Fua and Hanson for correct bib.

[29] P. Bose, A. Lubiw, and I. Munro. Efficient visibility queries in simple polygons. In *4th Canadian Conference on Computational Geometry*, 1992. To appear in Computational Geometry: Theory and Applications.

[30] K. W. Bowyer and C. R. Dyer. Aspect graphs: An introduction and survey of recent results. *Int. J. of Imaging Systems and Technology*, 2:315–328, 1990.

[31] P. Braß and C. Knauer. Testing the congruence of d-dimensional point sets. Technical Report B 99-18, Freie Universität Berlin, Berlin, November 1999.

[32] G. E. Bredon. *Introduction to Compact Transformation Groups*, volume 46 of *Pure and Applied Mathematics*. Academic Press, New York, N. Y., 1972.

[33] S. Brin. Near neighbour search in very large databases. In *21st International Conference on Very Large Data Bases*, pages 574–584. Morgan Kaufmann Publishers, September 1995.

[34] H. Bunke. A fast algorithm for finding the nearest neighbor of a word in a dictionary. Technischer Bericht IAM 93-025, Institut für Informatik, Universität Bern, Schweiz, November 1993. Filed: yes, PSFile: yes.

[35] Jochen Burghardt. A tree pattern matching algorithm with reasonable space requirements. In *Proc. 13th Colloquium on Trees in Algebra and Programming*, volume 299 of *LNCS*, pages 1–15, March 1988.

[36] C. C. Chen. Improved moment invariants for shape discrimination. *Pattern Recognition*, 26(5):683–686, 1993.

[37] J. S. Chen and G. G. Medioni. Parallel multiscale stereo matching using adaptive smoothing. In *ECCV90*, pages 99–103, 1990.

[38] L. P. Chew, D. Dor., A. Efrat, and K. Kedem. Geometric pattern matching in d-dimensional space. *Discrete & Computational Geometry*, 21(2):257–274, 1999.

[39] L. P. Chew, M. T. Goodrich, D. P. Huttenlocher, K. Kedem, J. M. Kleinberg, and D. Kravets. Geometric pattern matching under Euclidean motion. *Computational Geometry: Theory and Applications*, 7(1):113–124, 1997.

[40] L. P. Chew and K. Kedem. Getting around a lower bound for the minimum hausdorff distance. *International Journal of Computational Geometry and Applications*, 10:197–202, 1998.

[41] K. L. Clarkson. Nearest neighbor queries in metric spaces. In *29th Annual ACM Symposium on the Theory of Computing*, pages 609–617, 1997.

[42] E. T. Copson. *Metric spaces.* Cambridge University Press, 1968.

[43] G. Cox, G. de Jager, and B. Warner. A new method of rotation, scale and translation invariant point pattern matching applied to the target acquisition and guiding of an automatic telescope. In *2nd South African Workshop on Pattern Recognition*, pages 167–72, November 1991.

[44] H. S. M Coxeter. *Regular Polytopes.* Dover Publications, New York, N. Y., 1974.

[45] R. H. Davis and J. Lyall. Recognition of handwritten characters-a review. *Image and Vision Computing*, 4:208–218, 1986.

[46] M. de Berg, O. Cheong, O. Devillers, M. J. van Kreveld, and M. Teillaud. Computing the maximum overlap of two convex polygons under translation. *Theory of Computing Systems*, 31(5):613–628, 1998.

[47] M. de Berg, D. Halperin, M. H. Overmars, and M. van Kreveld. Sparse arrangements and the number of views of polyhedral scenes. *International Journal of Computational Geometry and Applications*, 7(3):175–195, 1997.

[48] P. J. de Rezende and D. T. Lee. Point set pattern matching in d-dimensions. *Algorithmica*, 13(4):387–404, April 1995.

[49] M. Drumheller and T.A. Poggio. On parallel stereo. In *CRA87*, pages 527–538, 1987.

[50] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern classification*. Wiley, New York, N. Y., 2000.

[51] J. J. Duistermaat. Differentiaalrekening op variëteiten. Course Notes, Mathematics Department, revised version, 1995.

[52] H. Edelsbrunner, J. O'Rourke, and R. Seidel. Constructing arrangements of lines and hyperplanes with applications. *SIAM Journal on Computing*, 15:341–363, 1986.

[53] A. Efrat and A. Itai. Improvements on bottleneck matching and related problems using geometry. In *12th Annual ACM Symposium on Computational Geometry*, pages 301–310, 1996.

[54] A. Efrat and M. J. Katz. Computing fair and bottleneck matchings in geometric graphs. *International Symposium on Algorithms and Computation*, pages 115–125, 1996.

[55] R. Fagin and L. Stockmeyer. Relaxing the triangle inequality in pattern matching. *International Journal of Computer Vision*, 28(3):219–231, 1998.

[56] L. M. G. Fonseca and M. H. M. Costa. Automatic registration of satellite images. In *X Brazilian Symposium of Computer Graphic and Image Processing, Campos de Jordao*, pages 219–226, October 1997.

[57] K. Fukunaga. *Introduction to statistical pattern recognition*. Academic Press, San Diego, 2nd edition, 1990.

[58] S. K. Ghosh and D. M. Mount. An output-sensitive algorithm for computing visibility graphs. *SIAM Journal on Computing*, 20:888–910, 1991.

[59] Z. Gigus, J. Canny, and R. Seidel. Efficiently computing and representing aspect graphs of polyhedral objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13:542–551, 1991.

[60] M. Godeau. *On the complexity of measuring the similarity between geometric objects in higher dimensions*. PhD thesis, Freie Universität Berlin, 1998.

[61] R. C. Gonzalez and M. G. Thomason. *Syntactic Pattern Recognition*. Addison-Wesley, Reading, Mass., 1978.

[62] N. Grimal, J. Hallof, and D. van der Plas. Hieroglyphica, sign list. http://www.ccer.theo.uu.nl/hiero/hiero.html, 1993.

[63] W.E.L. Grimson. Computational experiments with a feature based stereo algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7(1):17–34, January 1985.

[64] B. Grünbaum. *Convex Polytopes*. Interscience, London, 1967.

[65] L. J. Guibas, Rajeev Motwani, and Prabhakar Raghavan. The robot localisation problem. *SIAM J. Computing*, 26(4), 1997.

[66] M. Hagedoorn, M. H. Overmars, and R. C. Veltkamp. New visibility partitions with applications in affine pattern matching. Technical Report UU-CS-1999-21, Utrecht University, Padualaan 14, 3584 CH Utrecht, the Netherlands, July 1999. http://www.cs.uu.nl/docs/research/publication/TechRep.html.

[67] M. Hagedoorn and R. C. Veltkamp. A general method for partial point set matching. In *13th Annual ACM Symposium on Computational Geometry*, pages 406–408. ACM Press, 1997.

[68] M. Hagedoorn and R. C. Veltkamp. Reliable and efficient pattern matching using an affine invariant metric. Technical Report UU-CS-1997-33, Utrecht University, Padualaan 14, 3584 CH Utrecht, the Netherlands, October 1997. Revision published in International Journal of Computer Vision.

[69] M. Hagedoorn and R. C. Veltkamp. Measuring resemblance of complex patterns. In L. Perroton G. Bertrand, M. Couprie, editor, *Discrete Geometry for Computer Imagery*, Lecture Notes in Computer Science, pages 286–298, Berlin Heidelberg, 1999. Springer Verlag.

[70] M. Hagedoorn and R. C. Veltkamp. Metric pattern spaces. Technical Report UU-CS-1999-03, Utrecht University, Padualaan 14, 3584 CH Utrecht, the Netherlands, January 1999. http://www.cs.uu.nl/docs/research/publication/TechRep.html.

[71] M. Hagedoorn and R. C. Veltkamp. Reliable and efficient pattern matching using an affine invariant metric. *International Journal of Computer Vision*, 31(2, 3):103–115, 1999.

[72] M. Hagedoorn and R. C. Veltkamp. A robust affine invariant metric on boundary patterns. *International Journal of Pattern Recognition and Articifial Intelligence*, 13(8):1151–1164, 1999. Special issue on invariants for pattern recognition and classification.

[73] Michiel Hagedoorn. Efficient pattern matching using traced volumes. Master's thesis, Department of Computer Science, Utrecht University, the Netherlands, August 1996.

[74] P. A. V. Hall and G. R. Dowling. Approximate string matching. *ACM Computing Surveys*, 12(4):381–402, 1980.

[75] N. Hata. *Rigid and deformable medical image registration for image-guided surgery.* PhD thesis, University of Tokyo, 1997.

[76] F. Hausdorff. *Mengenlehre.* Springer, Berlin, 3rd edition, 1927.

[77] U. Heckmanns. On the topology of ultrametric spaces. In *Proceedings of the 8th Prague Topology Symposium*, pages 149–156, 1996.

[78] P. J. Heffernan and S. Schirra. Approximate decision algorithms for point set congruence. *Computational Geometry: Theory and Applications*, 4(3):137–156, 1994.

[79] J. G. Hocking and G. S. Young. *Topology.* Addison-Wesley, Reading, Mass., 1961.

[80] J. S. Huang and M. L. Chung. Separating similar complex chinese characters by walsh transform. *Pattern Recognition*, 20:425–428, 1987.

[81] D. P. Huttenlocher and K. Kedem. Computing the minimum Hausdorff distance for point sets under translation. In *6th Annual ACM Symposium on Computational Geometry*, pages 340–349, 1990.

[82] D. P. Huttenlocher, K. Kedem, and M. Sharir. The upper envelope of Voronoi surfaces and its applications. *Discrete & Computational Geometry*, 9:267–291, 1993.

[83] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge. Comparing images using the Hausdorff distance. Technical Report 91-1211, Cornell University, 1991.

[84] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge. Comparing images using the Hausdorff distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15:850–863, 1993.

[85] D. P. Huttenlocher, J. J. Noh, and W. J. Rucklidge. Tracking non-rigid objects in complex scenes. Technical report, Cornell University, December 1992.

[86] D. P. Huttenlocher and W. J. Rucklidge. A multi-resolution technique for comparing images using the Hausdorff distance. Technical Report 92-1321, Cornell University, Ithaca, New York, 1992.

[87] I. M. James. *Introduction to Uniform Spaces*, volume 144 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, 1990.

[88] I. Kaplansky. *Set Theory and Metric Spaces*. Chelsea Publishing Company, New York, N. Y., 1977.

[89] J. L. Kelley. *General Topology*. The University Series in Higher Mathematics. D. van Nostrand Company, Inc, 120 Alexander St. Princeton, New Jersey, March 1955.

[90] A. Khotanzad and Y. H. Hong. Invariant image recognition by Zernike moments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(5):489–497, 1990.

[91] D. J. Kriegman and J. Ponce. Computing exact aspect graphs of curved objects: Solids of revolution. *International Journal of Computer Vision*, 5:119–135, 1990.

[92] T. Luczak and W. Szpankowski. A suboptimal lossy data compression based in approximate pattern matching. In *IEEE Transactions on Information Theory*, volume 43, pages 1439–1451, 1997.

[93] J. B. A. Maintz. *Retrospective registration of tomographic brain images*. PhD thesis, Universiteit Utrecht, December 1996. ISBN 90-393-1227-3.

[94] G. Matheron. *Random Sets and Integral Geometry*. Wiley series in probability and mathematical statistics. John Wiley & Sons, 1975.

[95] N. Megiddo. Applying parallel computation algorithms in the design of serial algorithms. *Journal of the ACM*, 30(4):852–865, October 1983.

[96] E. Michael. Topologies on spaces of subsets. *Transactions of the American Mathematical Society*, 71:151–182, 1951.

[97] D. Montgomery and L. Zippin. *Topological Transformation Groups*. Interscience tracts in pure and applied mathematics. Interscience Publishers, Inc., 1964.

[98] D. M. Mount, N. S. Netanyahu, and J. Le Moigne. Efficient algorithms for robust feature matching. *Computer Vision and Image Understanding*, 32:17–38, 1999.

[99] D. M. Mount, R. Silverman, and A. Y. Wu. On the area of overlap of translated polygons. *Computer Vision and Image Understanding*, 64:53–61, July 1996.

[100] K. Mulmuley. *Computational Geometry: An introduction through randomized algorithms*. Prentice Hall, Englewood Cliffs, N. J., 1994.

[101] J. R. Munkres. *Topology*. Prentice Hall, Englewood Cliffs, N. J., 1975.

[102] M. E. Munroe. *Measure and Integration*. Addison-Wesley, Reading, Mass., 1953.

[103] H.K. Nishihara. Prism: A practical real-time imaging stereo matcher. In *MIT AI*, 1984.

[104] Y. H. Pao. *Adaptive Pattern Recognition and Neural Networks*. Addison-Wesley, Reading, Mass., 1989.

[105] W. H. Plantinga and C. R. Dyer. Visibility, occlusion and the aspect graph. *International Journal of Computer Vision*, 5(2):137–160, 1990.

[106] M. Pocchiola and G. Vegter. Topologically sweeping visibility complexes via pseudotriangulations. *Discrete & Computational Geometry*, 16(4):419–453, 1996.

[107] R. Pollack, M. Sharir, and S. Sifrony. Separating two simple polygons by a sequence of translations. *Discrete & Computational Geometry*, 3:123–136, 1988.

[108] T. E. Portegys. A search technique for pattern recognition using relative distances. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(9):910–913, September 1995.

[109] Günter Rote. Computing the minimum Hausdorff distance between two point sets on a line under translation. *Information Processing Letters*, 38(3):123–127, 17 May 1991.

[110] W. J. Rucklidge. *Efficient Visual Recognition Using the Hausdorff Distance*. Number 1173 in Lecture Notes in Computer Science. Springer-Verlag, 1996.

[111] W. J. Rucklidge. Lower bounds for the complexity of the graph of the Hausdorff distance as a function of transformation. *Discrete & Computational Geometry*, 16(2):135–153, September 1996.

[112] Y. Rui, A. She, and T. S. Huang. *Image Databases and Multimedia Search*, volume 8 of *Series on Software Engineering and Knowledge Engineering*, chapter A Modified Fourier Descriptor for Shape Matching in MARS, pages 165–180. World Scientific Publishing House, Singapore, 1998.

[113] E. Schulte. Symmetry of polytopes and polyhedra. In J. E. Goodman and J. O'Rourke, editors, *Handbook of Discrete and Computational Geometry*, pages 311–344. CRC Press, Boca Raton, 1997.

[114] C. G. Small. *The Statistical Theory of Shapes*. Springer Series in Statistics. Springer-Verlag, New York, N. Y., 1996.

[115] D. Spotts Fry. *Shape recognition using Metrics on the space of shapes*. PhD thesis, Harvard University, 1993.

[116] P. Suetens, P. Fua, and A. J. Hanson. Computational strategies for object recognition. *ACM Computing Surveys*, 24(1), March 1992.

[117] J. Uhlmann. Metric trees. *Applied Mathematics Letters*, 4(5), 1991.

[118] P. M. Vaidya. Geometry helps in matching. *SIAM J. Computing*, 18(6):1201–1224, December 1989.

[119] P. A. van den Elsen. *Multimodality matching of brain images*. PhD thesis, Universiteit Utrecht, February 1993. ISBN 90-71546-02-0.

[120] R. C. Veltkamp and M. Hagedoorn. State-of-the-art in shape matching. Technical Report UU-CS-1999-27, Utrecht University, Padualaan 14, 3584 CH Utrecht, the Netherlands, September 1999.

[121] R. C. Veltkamp and M. Hagedoorn. Shape similarity measures, properties and constructions. In *Visual 2000*, 2000. To appear.

[122] S. Venkatasubramanian. *Geometric Shape Matching and Drug Design*. PhD thesis, Stanford University, 1999.

[123] J. Vleugels. *On Fatness and Fitness–Realistic Input Models for Geometric Algorithms*. PhD thesis, Department of computer science, Utrecht University, March 1997.

[124] Jules Vleugels and Remco C. Veltkamp. Efficient image retrieval through vantage objects. In D. P. Huijsmans and A. W. M. Smeulders, editors, *Visual Information and Information Systems – Proceedings of the Third International Conference VISUAL'99, Amsterdam, The Netherlands, June 1999*, LNCS 1614, pages 575–584. Springer, 1999.

[125] J. Weng, N. Ahuja, and T. S. Huang. Matching two perspective views. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(8):806–825, August 1992.

[126] S. White. Stereo using the displacement representation. In *DARPA92*, pages 391–399, 1992.

[127] H. J. Wolfson. Model based object recognition by geometric hashing. In *Proc. 1st European Conference on Computer Vision*, pages 526–536, 1990.

[128] H. J. Wolfson and I. Rigoutsos. Geometric hashing: An overview. *IEEE Computational Science & Engineering*, 4(4), October 1997.

[129] P. N. Yianilos. Data structures and algorithms for nearest neighbour search in general metric spaces. In *4th ACM-SIAM Symposium on Discrete Algorithms*, pages 311–321, 1993.

[130] A.L. Yuille and T.A. Poggio. A generalized ordering constraint for stereo correspondence. In *MIT AI Memo*, 1984.

Acknowledgements

First of all, I thank my parents and my brother for their unconditional support. I thank my parents in particular because they got me a Commodore 64 on my twelfth birthday. Within a year I became a member of the computer club STACK in my hometown Vught. After my C64, I got an Amiga 2000. In my first year as a computer science student I got a PC with an Intel 386 processor. During the last the year of my studies the use of the 386 machine motivated me to develop really fast algorithms. Actually, this oldtimer is still in use: I wrote large parts of this thesis using it.

Remco Veltkamp has been my supervisor from the moment I started the research for my master's thesis. Remco has an open minded attitude. He encourages his students to study how research problems are approached by different scientific communities. Remco enthusiastically supported my research ideas, allowing me to develop them with success. When discussing new ideas with him, he proved to have the special ability of making many unconscious choices explicit. Remco's accurate reading much helped to improve this thesis.

I thank Mark Overmars, my supervising professor, for his appraisal of my work. It can not be overemphasised how important this is for a PhD student.

Bram Verwey helped me with a variety of things, including advanced LaTeX stuff. C++ programming, image processing and theoretical issues. Especially from the moment that his own thesis was finished, Bram has been spending many hours helping me out with nasty details. In addition, he installed the "reanimator" desktop theme on my computer, allowing me to work much more efficiently. If it were not for Bram, this thesis would have had a much different appearance.

Many other people from the department of computer science helped me on a regular basis. Mark de Berg, Marc van Kreveld, and Mark Overmars shared their extensive knowledge of computational geometry with me. Whenever I had a question, they generously sacrificed their time to help me. Geert-Jan Giezeman, Dima Pasechnik, and Wieger Wesselink patiently helped me whenever I had problems implementing my algorithms. I thank Job Smeltink, Tycho

Strijk and Bram Verwey for sharing their thoughts about several problems with me. With his mathematical background, Tycho Strijk was able to provide me with exactly the right literature.

I am much indebted to Klara Kedem, Jan van Leeuwen and Mark Overmars for the the useful comments they provided on the manuscript of this thesis. I am also indebted to Geert-Jan Giezeman and Hans Tangelder. Their attentiveness during a talk allowed me to make a few improvements.

A few people made life at the University more fun, especially my friends and collegues Bram Verwey and Robert-Paul Berretty. Their continuous mental coaching got me through the years as a PhD student. Sometimes their support went extremely far. For example, Bram and RP always let me win when playing Doom over the network, even when cooperating.

Samenvatting

Computers kunnen gebruikt worden om geometrische vormen te herkennen. Dit heeft toepassingen zoals het automatisch lezen van handgeschreven tekst, het zelfstandig oppakken van objecten door robots en het vinden van het meest gelijkende plaatje op internet, gegeven een zoekplaatje. Gelijkenismaten zijn een solide basis voor zulke technieken.

Dit proefschrift behandelt de wiskundige en algoritmische aspecten van gelijkenismaten. Het eerste aspect heeft te maken met de vraag hoe de gelijkenis tussen geometrische patronen zou moeten worden gemeten. Het tweede aspect heeft te maken met de berekening van een gelijkenismaat en de minimalisatie van de waarde van een gelijkenismaat onder geometrische transformaties.

In Hoofdstuk 2 presenteer ik een nieuwe theorie die gebruikt wordt voor de analyse van verschillende gelijkenismaten. De aandacht ligt bij gelijkenismaten die pseudometrieken zijn op een collectie van deelverzamelingen van een ruimte. Zoals de "grote-oh" notatie kan worden gebruikt om uitspraken te doen over de efficientie van algoritmen, zo kan de nieuwe theorie in dit proefschrift worden gebruikt om zinvolle uitspraken te doen over de robuustheid van gelijkenismaten. The theorie voor gelijkenismaten wordt toegepast in de analyse van zowel bekende gelijkenismaten als nieuwe gelijkenismaten die worden geintroduceerd in dit proefschrift. De bestaande gelijkenismaten zijn o.a. de Hausdorff metriek en het volume van het symmetrische verschil. De nieuwe gelijkenismaten zijn o.a. het genormaliseerde volume van het symmetrische verschil en de reflectie-zichtbaarheids afstand.

Eerst bespreek ik de theorie van algemene pseudometrische ruimten. Deze verhandeling gaat onder meer over de transformatiegroep waaronder een pseudometrische ruimte invariant is, de topologie behorende bij een pseudometrische ruimte en de operaties die kunnen worden toegepast op een pseudometrische ruimte. Ook laat ik zien hoe de minimalisatie van een pseudometriek onder een transformatiegroep tot een nieuwe pseudometriek leidt die onafhankelijk is van transformaties. Verder laat ik zien hoe een pseudometrische ruimte kan worden uitgebreid met een nieuw element zonder dat de essentiële eigenschappen van

de oorspronkelijke ruimte verloren gaan. Deze techniek kan worden toegepast om het domein van een gelijkenismaat uit te breiden met de lege verzameling.

Vervolgens introduceer ik een nieuwe structuur: de pseudometrische patroonruimte. Deze structuur is rijker dan pseudometrische ruimten in het algemeen. In het bijzonder maken pseudometrische ruimten de formalisatie van verschillende soorten robuustheid mogelijk. In de vakliteratuur wordt het belang van robuustheidseigenschappen voor een gelijkenismaat bevestigd. Echter, zover ik weet, zijn dit soort eigenschappen tot nu toe nooit precies gemaakt. In de vorm van vier axioma's druk ik vier soorten robuustheid uit. Deze vormen van robuustheid heten vervormings robuustheid, vervagings robuustheid, barst robuustheid en ruis robuustheid. Ik bewijs dat de axioma's zich netjes gedragen onder de toepassing van verschillende standaard operaties op pseudometrische patroonruimten.

Hierna geef ik een nieuwe methode waarmee verschillende pseudometrieken op een collectie patronen kunnen worden geconstrueerd. De constructie methode is gebaseerd op de toekenning van reëlwaardige functies aan patronen. Ik bewijs dat eenvoudige voorwaarden op deze toekenning voldoende zijn om de invariantie onder een gegeven transformatiegroep te garanderen. De constructie methode wordt op verschillende plaatsen in dit proefschrift toegepast, resulterend in nieuwe gelijkenismaten.

De nieuw ontwikkelde theorie wordt eerst toegepast om de Hausdorff metriek te analyseren. Dit resulteert in een aantal nieuwe resultaten voor de Hausdorff metriek. Ik breid een bestaan resultaat van Matheron [94] uit door te laten zien dat de met de lege verzameling uitgebreide Hausdorff metriek precies de "bijziende" topologie definieert op de collectie van alle compacte deelverzamelingen van een metrische "basisruimte". Deze topologie wordt vervolgens gebruikt om eenvoudig te bewijzen dat de Hausdorff metriek vervormings, vervagings, en barst robuust is. Verder toon ik aan dat in een groot aantal gevallen de Hausdorff metriek niet ruis robuust is.

Vervolgens analyseer ik een andere gelijkenismaat, het volume van het symmetrische verschil. In de analyse komen de volgende nieuwe resultaten voort. Ik bewijs dat het volume van het symmetrische verschil een topologie voortbrengt die onvergelijkbaar is met die van de Hausdorff metriek. Ook introduceer ik een nieuwe gelijkenismaat die het genormaliseerde volume van het symmetrische verschil wordt genoemd. Ik laat zien dat deze afstandmaat invariant is onder ratio-van-volume behoudende transformaties. Verder bewijs ik dat het volume van het symmetrische verschil en de genormaliseerde versie daarvan aan elk van de vier robuustheidsaxioma's voldoen.

Verder introduceer ik een nieuwe gelijkenismaat: de reflectie-zichtbaarheids afstand. De definitie hiervan is gebaseerd op de bovengenoemde constructiemethode. De reflectie-zichtbaarheids afstand is een metriek op een vrij al-

gemene klasse van $(k-1)$-dimensionale deelverzamelingen van $\mathbb{R}^k$. De genormaliseerde versie van de reflectie-zichtbaarheids afstand is invariant under de groep van affiene transformaties. Ik bewijs dat zowel de gewone als de genormaliseerde versie van de nieuwe gelijkenismaat aan elk van de vier robuustheidsaxioma's voldoen.

Ik heb computer programma's geschreven die de Hausdorff metriek en de genormaliseerde reflectie-zichtbaarheids afstand kunnen berekenen. Hiermee heb ik experimenten uitgevoerd. De resultaten hiervan suggereren dat de de reflectie-zichtbaarheids afstand ook in praktische gevallen robuuster is dan de Hausdorff metriek.

In Hoofdstuk 3 worden de combinatorische en computationele eigenschappen van gelijkenismaten besproken. Een belangrijk voorbeeld is het aantal minima dat een gelijkenismaat kan hebben under de werking van een transformatiegroep. Dit type resultaat toont aan hoe moeilijk patroonherkenning met behulp van gelijkenismaten is gezien vanuit het oogpunt van de computationele geometrie. Verder bespreekt Hoofdstuk 3 nieuwe en bestaande algoritmen voor het minimaliseren van een gelijkenismaat onder transformaties. Onder meer de discrete metriek, de Hausdorff metriek, het volume van het symmetrische verschil en de reflectie-zichtbaarheids afstand komen aan bod.

Eerst worden algemene computationele strategiën voor het minimaliseren van gelijkenismaten besproken. Er wordt aangetoond dat voor een groot aantal gelijkenismaten het minimalisatie probleem kan worden opgelost door het construeren van een "arrangement".

Het probleem van het minimaliseren van de discrete metriek onder transformaties is ook bekend als "exact congruence matching". Het maximaal aantal minima hierbij hangt samen met het maximaal aantal symmetriën dat een eindige puntenverzameling hebben kan. Ik bewijs nieuwe asymptotische onder en bovengrenzen voor dit aantal symmetriën.

Vervolgens bespreek ik het minimaliseren van de Hausdorff metriek. Hierbij worden voornamelijk reeds bekende combinatorische resultaten en algoritmen besproken.

Hierna bekijk ik de minimalisatie van het volume van het symmetrische verschil. Hierbij worden de volgende nieuwe resultaten besproken. Strakke onder en bovengrenzen voor het aantal minima worden gegeven voor verbonden polygonen in het vlak. Verder worden strakke onder en bovengrenzen gegeven voor verenigingen van simplices in hogere dimensies. Ook wordt de lokale beschrijving van het volume van de doorsnede van twee polygonen als functie van transformaties gegeven. Specifiek wordt bewezen dat voor polygonen met as-parallele zijden deze lokale beschrijving een rationale functie is die gelijk is aan een quotiënt van polynomen met begrensde graad.

Vervolgens wordt het eerste algoritme gepresenteerd waarmee de reflectie-

zichtbaarheids afstand kan worden berekend. Hierbij worden gerandomiseerde technieken toegepast. De berekening is gebaseerd op de constructie van structuren die reflectie-zichtbaarheids partities worden genoemd. Deze structuren zijn nauw verwant aan zichtbaarheids partities. Verder worden strakke onder en bovengrenzen op de complexiteit van reflectie-zichtbaarheids partities gegeven.

In Hoofdstuk 4 introduceer ik nieuwe algoritmen waarmee het infimum van de waarden van een gelijkenismaat onder de werking van een transformatiegroep met elke gewenste nauwkeurigheid kunnen worden berekend. Eerst wordt het "GBB algoritme" gepresenteerd. Daarna geef ik een stelling die de voorwaarden duidelijk maakt waaronder het GBB algoritme termineert voor iedere invoer. Er worden twee toepassingen van het GBB algoritme gegeven: de "sporen aanpak" en de "partitie combinatie" aanpak. Van beide aanpakken bewijs ik de correctheid.

De sporen aanpak kan worden toegepast op een grote groep gelijkenismaten en verschillende transformatiegroepen. Deze aanpak berekent het "spoor" dat een punt achterlaat als een verzameling transformaties op het punt wordt losgelaten.

De partitie combinatie aanpak is gespecialiseerd voor het genormaliseerde volume van het symmetrisch verschil en kan goed worden toegepast op groepen van affiene transformaties. Bij deze aanpak worden vereenvoudigingen van patronen gebruikt om de berekening te versnellen. Als het nodig is worden de vereenvoudigingen op een hiërachische wijze verfijnd.

Uiteindelijk worden enige experimentele resultaten gepresenteerd die verkregen zijn met behulp van een implementatie van de sporen aanpak.

Vita

Michiel Hagedoorn

13 juli 1972	Geboren te Renkum
1985-1991	Atheneum, Maurick-college te Vught
1991-1996	Studie Informatica, Universiteit Utrecht
1996-2000	Assistent in Opleiding, Informatica, Universiteit Utrecht

Symbols

Notation

$r, s, t, \ldots$	integers
$\alpha, \beta, \gamma, \ldots$	real numbers
$a, b, c, \ldots$	points in a space
$A, B, C, \ldots$	sets of points
$\mathcal{A}, \mathcal{B}, \mathcal{C}, \ldots$	collections of point sets
$\mathfrak{A}, \mathfrak{B}, \mathfrak{C}, \ldots$	families of point set collections
$\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \ldots$	real valued functions on $\mathbb{R}^k$

Real valued functions

$\boldsymbol{n}_P$	real valued function associated with pattern P	42
$\mathbf{1}_P$	characteristic function of set P	41
$\boldsymbol{v}_P$	Voronoi surface of P	49
$\boldsymbol{r}_P$	reflection visibility surface of P	61
$\boldsymbol{u}_{A,B}$	pointwise difference of visibility stars of A and B	67
$\boldsymbol{a} + \boldsymbol{b}$	pointwise sum of functions $\boldsymbol{a}$ and $\boldsymbol{b}$	41
$\boldsymbol{a} - \boldsymbol{b}$	pointwise difference of functions $\boldsymbol{a}$ and $\boldsymbol{b}$	41
$\boldsymbol{a} \wedge \boldsymbol{b}$	pointwise minimum of functions $\boldsymbol{a}$ and $\boldsymbol{b}$	41
$\boldsymbol{a} \vee \boldsymbol{b}$	pointwise maximum of functions $\boldsymbol{a}$ and $\boldsymbol{b}$	41
$\int_{\mathbb{R}^k} f(x)\, dx$	Lebesgue integral of f over $\mathbb{R}^k$	40
$\boldsymbol{I}_p^k$	special class of real valued functions	40

Operations on subsets

$\mathrm{vol}\,(P)$	k-dimensional volume of set P	54
$\mathrm{diam}\,(P)$	diameter of set P	129
$\mathrm{Int}\,(P)$	interior of set P	150

$\mathrm{Cl}\,(P)$	closure of set P	150
$\mathrm{Bd}\,(P)$	boundary of set P	150
$P \cap Q$	intersection of P and Q	
$P \cup Q$	union of P and Q	
$P - Q$	set difference of P and Q	
$P \triangle Q$	symmetric set difference of P and Q	54

Linear and differential geometry

M^T	transpose of a matrix M	155
$\det(M)$	determinant of square matrix M	154
$\mathrm{D}_f(x)$	total derivative of f in x	154
$\delta_f(x)$	absolute value of determinant of $\mathrm{D}_f(x)$	42
δ_f	value for constant $\delta_f(x)$	44

Geometric constructions

$\mathrm{Line}(x, y)$	line through x and y	63
$\mathrm{Seg}(x, y)$	open line segment between x and y	60
$\mathrm{Ray}(x, y)$	open ray emanating from x passing through y	68
$\mathrm{B}_d(x, \epsilon)$	open ball centred at x with radius ϵ	24
$\mathrm{N}_d(P, \epsilon)$	ϵ neighbourhood of P	46
$\mathrm{B}(x, \epsilon)$	Euclidean open ball centred at x with radius ϵ	63
$\mathrm{N}(P, \epsilon)$	Euclidean ϵ neighbourhood of P	68
$\mathrm{C}(P)$	smallest Euclidean ball containing P	68
$R_\epsilon(A, B)$	relation in definition Hausdorff metric	47
$\mathrm{S}_P(x)$	generic set for point x relative to P	70
$\mathrm{V}_P(x)$	points visible from x relative to P	61
$\mathrm{V}^\star_P(x)$	visibility star for x relative to P	61
$\mathrm{T}^\star_P(x)$	trans visibility star for x relative to P	111
$\mathrm{R}^\star_P(x)$	reflection visibility star for x relative to P	61
$\mathrm{Vmp}_{\mathcal{S}}(x)$	view map for x in $\mathcal{S}$	106
$\mathrm{Tmp}_{\mathcal{S}}(x)$	trans view map for x in $\mathcal{S}$	112
$\mathrm{Rmp}_{\mathcal{S}}(x)$	reflection view map for x in $\mathcal{S}$	115

Transformation groups

$\mathrm{Hom}(X)$	homeomorphisms from X onto itself	31
$\mathrm{Iso}(X)$	isometries in general metric space X	71
$\mathrm{Clos}(X, \mathcal{P})$	closure group	31

Dif^k	diffeomorphisms from $\mathbb{R}^k$ onto itself	154
CDif^k	ratio of volume preserving diffeomorphisms	154
UDif^k	volume preserving diffeomorphisms	154
Af^k	affine transformations	154
UAf^k	volume preserving affine transformations	155
Stret^k	stretch transformations	155
Thet^k	homotheties	155
Sim^k	similarity transformations	155
Iso^k	Euclidean isometries	155
Lat^k	translations	155
Rot^k	rotations around the origin	91
Id^k	identity group, trivial group	155

Congruence, symmetry and normalisation

$G_{A:B}$	congruences of A and B in G	85
G_P	symmetries of P in G	85
$\langle\cdot\rangle$	normalisation function	86

Collections of subsets

$\wp(X)$	all subsets of X	71
$\mathcal{F}(X)$	finite subsets of X	86
$\mathcal{C}(X)$	closed and bounded subsets of X	46
$\mathcal{C}'(X)$	nonempty elements of $\mathcal{C}(X)$	46
$\mathcal{K}(X)$	compact subsets of X	48
$\mathcal{K}'(X)$	nonempty elements of $\mathcal{K}'(X)$	48
$\mathcal{K}^+(\mathbb{R}^k)$	nonzero volume elements of $\mathcal{K}(\mathbb{R}^k)$	54
$\mathcal{K}_P(X)$	elements of $\mathcal{K}(X)$ intersecting P	50
$\mathcal{K}^P(X)$	elements of $\mathcal{K}(X)$ disjoint with P	50
$\mathcal{K}^C_{V_1,\ldots,V_n}(X)$	equals $\mathcal{K}^C(X) \cap \mathcal{K}_{V_1}(X) \cap \cdots \cap \mathcal{K}_{V_n}(X)$	50
$\mathcal{S}(X)$	solid subsets of X	55
$\mathcal{S}'(X)$	nonempty elements of $\mathcal{S}(X)$	55
$\mathcal{M}(\mathbb{R}^k)$	compact sets with simplex union boundary	62
$\mathcal{T}(\mathbb{R}^k)$	elements of $\mathcal{M}(\mathbb{R}^k)$ with empty interior	62

Distances

d	abstract similarity measure, or (pseudo)metric	22

$d^{\triangleright}$	directed similarity measure	127
ρ	base metric	31
$\|\cdot\|_p$	norm for $\mathbb{R}^k$	14
$\|\cdot\|$	Euclidean norm for $\mathbb{R}^k$	14
l_p	metric for $\mathbb{R}^k$ based on $\|\cdot\|_p$	14
$\|\cdot\|_p$	pseudonorm for $\boldsymbol{I}_p^k$	41
$\boldsymbol{l}_p$	pseudometric for $\boldsymbol{I}_p^k$	41
$\boldsymbol{l}_p^*$	normalised pseudometric for $\boldsymbol{I}_p^k$	43
z	indiscrete pseudometric	23
c	discrete metric	23
b_ρ	Bottleneck distance based on ρ	12
f_ρ	Fréchet distance based on ρ	13
h_ρ	Hausdorff metric based on ρ	46
$\mathsf{h}_\rho^{\triangleright}$	directed Hausdorff distance based on ρ	47
$\mathsf{h}_\rho^{\varnothing}$	Hausdorff metric extended for $\varnothing$	50
s	volume of symmetric difference	54
s^*	normalised volume of symmetric difference	54
a	absolute difference	59
$\mathsf{a}^{\triangleright}$	directed absolute difference	139
$\mathsf{s}^{*\triangleright}$	directed version of s^*	131
r	reflection visibility distance	65
r^*	normalised reflection visibility distance	65
$d^{\mathcal{S}}$	quotient metric for d and partition $\mathcal{S}$	27
d^G	orbit pseudometric for d under G	25
$d\|_{S\times S}$	restriction of d to $S \times S$	29
d^o	extension of d with element o	29
d^c	complementation of d	39

Geometric branch-and-bound

$\mathcal{C}$	class of cells	122
$G(C)$	subset of G represented by cell C	122
`l`	lower cell bound	123
`u`	upper cell bound	123
`cover`	the cover operation	123
`lGlobal`	lower global bound	123
`uGlobal`	upper global bound	123
`select`	selection operation for cell collection	124
`refine`	refinement operation for cell collection	124
$\texttt{cover}^r$	r times iterated cover operation	125
$\texttt{refine}^r$	r times iterated refine operation	124

$\sqcap$	increasing binary operator	127
$\overrightarrow{\tau}(C, A)$	forward trace of A under C	127
$\overleftarrow{\tau}(C, B)$	backward trace of B under C	127
σ_C	transformation represented by C	128
$\underline{\underline{f}}_C$	pointwise lower bound for $f\|_{G(C)}$	133
$\overline{\overline{f}}_C$	pointwise upper bound for $f\|_{G(C)}$	133

Index