COMPUTER VISION AND
MATHEMATICAL MORPHOLOGY

J.B.T.M. Roerdink †
Institute for Mathematics and Computing Science
University of Groningen
P.O. Box 800, 9700 AV Groningen, The Netherlands

Abstract

Mathematical morphology as originally developed by Matheron and Serra is a theory of set mappings, modeling binary image transformations, which are invariant under the group of Euclidean translations. This framework turns out to be too restricted for many applications, in particular for computer vision where group theoretical considerations such as behavior under perspective transformations and invariant object recognition play an essential role. So far, symmetry properties have been incorporated by assuming that the allowed image transformations are invariant under a certain commutative group. This can be generalized by dropping the assumption that the invariance group is commutative. To this end we consider an arbitrary homogeneous space (the plane with the Euclidean translation group is one example, the sphere with the rotation group another), i.e. a set X on which a transitive but not necessarily commutative transformation group Γ is defined. As our object space we then take the Boolean algebra P(Γ) of all subsets of this homogeneous space. Generalizations of dilations, erosions, openings and closings are defined and several representation theorems can be proved. We outline some of the limitations of mathematical morphology in its present form for computer vision and discuss the relevance of the generalizations discussed here.

Keywords: Mathematical morphology, image processing, transitive group action, non-commutative transformation group, Minkowski operations, invariance, computer vision.


†Tel. +31-50-3633931; Fax +31-50-3633800; Email roe@cs.rug.nl
1 Introduction

Mathematical morphology was originally developed at the Paris School of Mines as a set-theoretical approach to image analysis [10,21]. It has a strong algebraic component, studying image transformations with a simple geometrical interpretation and their decomposition and synthesis in terms of set operations. Although the main object of our present study is the algebraic approach we emphasize that our primary motivation comes from the geometrical side, in the sense that various image transformations used in mathematical morphology today (dilations, erosions, openings, closings) have a straightforward geometrical analogue in a more general context. It is then a natural question to ask whether a corresponding algebraic description can be found.

In the original approach a two-dimensional image of, let us say, a planar section of a porous material is modeled as a subset \( X \) of the plane. In order to reveal the structure of the material, the image is probed by translating small subsets \( B \), called structuring elements, of various forms and sizes over the image plane and recording the locations \( h \) where certain relations (e.g., \( B_h \) included in \( X \), \( B_h \) hits \( X \), etc.) between the image \( X \) and the translate \( B_h \) of the structuring element \( B \) over the vector \( h \) are satisfied. In this way one can construct a large class of image transformations which are invariant under the Euclidean translation group. The underlying idea here is that the form or shape of objects in the image does not depend on the relative location with respect to an arbitrary origin and that therefore the transformations performed on the image should respect this. Notice that the basic object of study, the ‘object space’, is not the reference space (the plane in our example) itself, but the collection of subsets of this reference space, and the transformations defined on this collection of subsets.

In practice one encounters various situations where this framework is too restrictive. Certain images show a clear radial symmetry with an intrinsic origin. In this case we need image transformations which are adapted to the symmetries of this polar structure. Now one obtains a straightforward generalization of Euclidean morphology by replacing the Euclidean translations by an arbitrary abelian (commutative) group [5,18]. In the case of the example mentioned above, this would be the group generated by rotations and multiplications with respect to the origin. Here the size of the structuring element increases with increasing distance from the origin. Another example occurs in the analysis of traffic scenes, where the goal is to recognize the shape of automobiles with a camera on a bridge overlooking a highway [2]. In this case the size of the structuring element has to be adapted according to the law of perspective. In Sect. 3 we will show that in this case there is again invariance under a commutative group. Notice that in the two examples just mentioned we have a variable structuring element as a function of position. In fact we will argue that without a concept of invariance (under a group, or otherwise), one cannot even give a meaningful answer to the question when sets at different locations are ‘of the same shape’ or not.

Instead of changing the symmetry group of the object space one may generalize
the object space itself to complete lattices. This is the approach initiated by Serra and Matheron [21,22], as well as Heijmans [5]. A general study of this topic has recently been made by Heijmans and Ronse, see [6,7,20]. A lattice formulation also is in order for studying grey-level images.

One may drop the assumption that the invariance group is commutative [15]. To this end we consider an arbitrary homogeneous space, i.e. a set $X$ on which a transitive but not necessarily commutative group $\Gamma$ of invertible transformations is defined. Here transitive means that for any pair of points in the set there is a transformation in the group which maps one point on the other. If this mapping is unique we say that the transformation group is simply transitive or regular. As the object space of interest from a morphological point of view we take here the Boolean algebra of all subsets of this homogeneous space.

We present some examples for basic motivation. First of all one may extend Euclidean morphology in the plane by including rotations. This case has been extensively discussed in [13]. In that case it is appropriate to use the full Euclidean group of motions (the group generated by translations and rotations) as (non-commutative) invariance group. This is for example the basic assumption made in integral geometry to give a complete characterization (Hadwiger’s Theorem) of functionals of compact, convex sets in $\mathbb{R}^n$ [4]. As our second basic example we mention the sphere with its symmetry group of three-dimensional rotations, again a non-abelian group. A detailed investigation of this case is presented in [14]. A third example is the area of path planning or motion planning. Here the problem is to find a path for an object, say a robot or a car, moving in a space (called ‘workspace’) with obstacles [16].

In the area of computer vision, mathematical morphology has so far not made a breakthrough. This is the case for several reasons. First, there is the limitation that mathematical morphology considers a 2D image as a direct representation of the structure which is analyzed. In particular the question of how to take the projective geometry of the imaging process into account has not been discussed. This is a serious shortcoming, since clearly the symmetry of a two-dimensional plane is not the same as the symmetry of the three-dimensional world of which it is a projection. It is here that we try to make some progress by the methods discussed in this paper.

Another, even more serious problem, which will not be addressed here, is that of occlusion of objects in 3D. The basic ingredients of mathematical morphology are ordering of sets by inclusion and transformations adapted to this ordering. So far it is not clear at all how mathematical morphology can be modified to deal with this problem and related ones, such as influence of lightning conditions, surface properties, etc. This situation was pregnantly summarized by Ronse [19] in the aphorism:

\[ \text{Mathematical morphology is flat.} \]

A first attempt towards morphological analysis of patterns on curved surfaces has been made in [17] using techniques from differential geometry. It remains to be
seen to what extent this will become a useful tool for computer vision too. For general questions of invariance in computer vision, see for example Mundy et al. [11].

The organization of the paper is as follows. First we introduce the general framework of mathematical morphology on Euclidean space, followed by its generalization to arbitrary homogeneous spaces (Section 2). Then the application to computer vision is studied. In Sect. 3 we take the projective geometry of the imaging process into account and explain how to define morphological operations on images of a planar patch produced by perspective projection, where the allowed 3D motion is restricted to translation within the object plane. Finally in Sect. 4 a problem from robot vision for path planning is approached by morphological methods.

2 Preliminaries

In this section we first outline some elementary concepts and results from classical Euclidean morphology. Then we generalize this to homogeneous spaces.

2.1 Euclidean morphology

Consider the space $E$, where $E = \mathbb{R}^n$ or $E = \mathbb{Z}^n$. Denote by $\mathcal{P}(E)$ the power set (set of all subsets) of $E$. The classical Minkowski addition and subtraction for subsets $X, A$ of $E$ are given by

$$X \oplus A = \bigcup_{a \in A} X_a, \quad (1)$$

$$X \ominus A = \bigcap_{a \in A} X_{-a}, \quad (2)$$

where $X_a = \tau_a(X) = \{x + a : x \in X\}$, is the translate of $X$ over the vector $a \in E$, $x + y$ is the sum of $x$ and $y$, and $-x$ the reflection of $x$. It can be shown that

$$X \oplus A = \{h \in E : \tilde{A}_h \uparrow X\}, \quad (3)$$

where $\tilde{A} = \{-a : a \in A\}$ is the reflection of $A$ and $A \uparrow B$ ($A$ ‘hits’ $B$) is a general notation for $A \cap B \neq \emptyset$.

Two characteristic properties of dilation are:

$$\text{Distributivity w.r.t. union :} \quad (\bigcup_{i \in I} X_i) \oplus A = \bigcup_{i \in I} (X_i \oplus A) \quad (4)$$

$$\text{Translation invariance :} \quad (X \oplus A)_h = X_h \oplus A. \quad (5)$$

Similar properties hold for the erosion with intersection instead of union. A consequence of the distributivity property is that dilation and erosion are increasing
mappings. (A mapping \( \psi \) is called increasing when for all \( X, Y \in \mathcal{P}(E) \), \( X \subseteq Y \) implies that \( \psi(X) \subseteq \psi(Y) \).)

Other important increasing transformations are the opening and closing by a structuring element \( A \):

\[
Opening : \quad X \circ A := (X \ominus A) \oplus A = \bigcup_{h \in E} \{ A_h : A_h \subseteq X \} \tag{6}
\]

\[
Closing : \quad X \bullet A := (X \oplus A) \ominus A = \bigcap_{h \in E} \{ (A^c)_h : (A^c)_h \supseteq X \}. \tag{7}
\]

The opening is the union of all the translates of the structuring element which are included in the set \( X \). Opening and closing are related by Boolean duality: \((X^c \circ A)^c = X \bullet A^c\). A more general definition of dilations, erosions, openings and closings will be given in the next subsection in the framework of complete lattices.

2.2 Generalized Minkowski operators

On any group \( \Gamma \) one can define generalizations of the Minkowski operations \([15] \). Recall that a dilation (erosion) is a mapping commuting with unions (intersections). For any subsets \( G, H \) of \( \Gamma \) define the dilation

\[
\delta(G) := G \hat{\oplus} H := \bigcup_{h \in H} G h = \bigcup_{g \in G} g H, \tag{8}
\]

which generalizes the Minkowski addition to non-commutative groups. Here

\[
gH := \{ gh : h \in H \}, \quad Gh := \{ gh : g \in G \},
\]

with \( gh \) the group product of \( g \) and \( h \). Similarly, define the erosion \((h^{-1} \text{ is the group inverse of } h)\)

\[
\epsilon(G) := G \hat{\ominus} H := \bigcap_{h \in H} G h^{-1},
\]

which generalizes the Minkowski subtraction. Both mappings are left-invariant, e.g.

\[
\delta(gG) = g \delta(G), \quad \forall g \in \Gamma.
\]

This is the reason for the superscript ‘\( \lambda \)' on the ‘\( \oplus \)' symbol. For later use we also define the inverted set \( G^{-1} \) of \( G \) by

\[
G^{-1} = \{ g^{-1} : g \in G \}. \tag{9}
\]

Duality by complementation is expressed by the formula \((G \hat{\oplus} H)^c = G^c \hat{\ominus} H^{-1}\).
<table>
<thead>
<tr>
<th></th>
<th>simply transitive</th>
<th>multi-transitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>commutative</td>
<td>Euclidean morphology</td>
<td>--</td>
</tr>
<tr>
<td>non-commutative</td>
<td>group morphology</td>
<td>group action morphology</td>
</tr>
</tbody>
</table>

Figure 1: Classification of transformation groups and the associated morphologies.

2.3 Group actions

Let \( \mathcal{X} \) be a non-empty set, \( \Gamma \) a transformation group on \( \mathcal{X} \), that is, each element \( g \in \Gamma \) is a mapping \( g : \mathcal{X} \rightarrow \mathcal{X} \), satisfying

\[(i) \quad gh(x) = g(h(x)) \quad (ii) \quad e(x) = x,\]

where \( e \) is the unit element of \( \Gamma \), and \( gh \) denotes the product of two group elements \( g \) and \( h \). The inverse of an element \( g \in \Gamma \) will be denoted by \( g^{-1} \). Instead of \( g(x) \) we will also write \( gx \). We say that \( \Gamma \) is a group action on \( \mathcal{X} \) \([1,12,23]\).

The group \( \Gamma \) is called transitive on \( \mathcal{X} \) if for each \( x, y \in \mathcal{X} \) there is a \( g \in \Gamma \) such that \( gx = y \), and simply transitive when this element \( g \) is unique. A homogeneous space is a pair \((\Gamma, \mathcal{X})\) where \( \Gamma \) is a group acting transitively on \( \mathcal{X} \). Any transitive abelian permutation group \( \Gamma \) is simply transitive. If \( \Gamma \) acts on \( \mathcal{X} \), the stabilizer of \( x \in \mathcal{X} \) is the subgroup \( \Gamma_x := \{g \in \Gamma : gx = x\} \). Let \( \omega \) be an arbitrary but fixed point of \( \mathcal{X} \), henceforth called the origin. The stabilizer \( \Gamma_\omega \) will be denoted by \( \Sigma \) from now on:

\[\Sigma := \Gamma_\omega = \{g \in \Gamma : g\omega = \omega\}.\] (10)

The set of group elements which map \( \omega \) to a given point \( x \) is called a left coset and denoted by

\[g_x \Sigma := \{g_x s : s \in \Sigma\}.\] (11)

Here \( g_x \) is a representative (an arbitrary element) of this coset.

In Fig. 1 we give a classification of transformation groups and the associated morphologies.

2.4 Examples

In the following we present two examples. In each case \( \Gamma \) denotes the group and \( \mathcal{X} \) the corresponding set.
Example 1 \( \mathcal{X} = \text{Euclidean space } \mathbb{R}^n, \ \Gamma = \text{the Euclidean translation group } \mathbf{T}. \) 
\( \mathbf{T} \) is abelian. Elements of \( \mathbf{T} \) can be parametrized by vectors \( h \in \mathbb{R}^n \), with \( \tau_h \) the translation over the vector \( h \):
\[
\tau_h x = x + h, \quad h, x \in \mathbb{R}^n.
\] (12)

Example 2 \( \mathcal{X} = \text{Euclidean space } \mathbb{R}^n \ (n \geq 2), \ \Gamma = \text{the Euclidean motion group } \mathbf{M} := E^+(3) \) (proper Euclidean group, group of rigid motions), i.e., the group generated by translations and rotations (see [13]). The subgroup leaving a point \( p \) fixed is the set of all rotations around that point. \( \mathbf{M} \) is not abelian. The collection of translations forms the Euclidean translation group \( \mathbf{T} \). The stabilizer, denoted by \( \mathbf{R} \), equals the circle group \( S^1 \) (also commutative) of rotations around the origin. Let \( \tau_h \) denote the translation over the vector \( h \in \mathbb{R}^2 \) and \( \rho^0_\phi \) the rotation over an angle \( \phi \) around the point \( p \). The following relations, whose proof is left to the reader, are needed in the sequel:
\[
\rho^0_\phi = \tau_p \rho^0_\phi \tau_{-p} = \tau_{p-\rho^0_\phi p} \rho^0_\phi,
\] (13)
\[
\rho^0_\phi \tau_h = \tau_{\rho^0_\phi h} \rho^0_\phi.
\] (14)

Let \( \gamma_{h,\phi} \) denote a rotation around the origin followed by a translation:
\[
\gamma_{h,\phi} = \tau_h \rho^0_\phi, \quad h \in \mathbb{R}^2, \ \phi \in S^1.
\] (15)

Any element of \( \mathbf{M} \) can be written in this form. Using the rules (13)-(14) one finds the multiplication rule
\[
\gamma_{h_2,\phi_2} \gamma_{h_1,\phi_1} = \gamma_{h_1+\rho^0_\phi h_2,\phi_1+\phi_2}.
\] (16)

**Geometrical representation**

The following representation is useful in this case [13,15]. Attach a set of unit vectors \( \vec{v} \) with direction varying over the unit circle to each a point in the plane. We call \( p := (x, \vec{v}) \) a *pointer*. Given any pointer \( p = (x, \vec{v}) \), there is a unique element of the group \( \Gamma \) which maps a fixed pointer \( b := (\omega, \vec{e}_1) \), called the *base pointer* to \( p \), where \( \omega \) is the origin and \( \vec{e}_1 \) a unit vector in the \( x \)-direction. So the pointer \( p = (x, \vec{v}) \), where \( \vec{v} = (\cos \phi, \sin \phi) \), represents the motion \( \gamma_{x,\phi} \). In this representation, the rotation group \( \mathbf{R} \) is the set of unit vectors attached to the origin and \( \mathbf{T} \) is represented by the collection of horizontal unit vectors attached to points of \( \mathbb{R}^2 \). In the discrete case we will use a hexagonal grid and replace \( \mathbf{R} \) by a finite group \( \mathbf{H} \) consisting of rotations over multiples of \( 60^\circ \). The coset \( \tau_\omega \mathbf{R} \) is represented on the hexagonal grid by the six unit vectors attached to the point \( x \) [15]. An example is given in Fig. 2, cf. also Fig. 6 below.

**2.5 Morphological operations on homogeneous spaces**

One can construct morphological operations on an arbitrary homogeneous space \( \mathcal{X} \) as follows. Define the 'origin' \( \omega \) to be an arbitrarily chosen point of \( \mathcal{X} \). To each
subset $X$ of $\mathcal{X}$ we associate all elements of the group which map the origin $\omega$ to an element of $X$. We also can go back from the group $\Gamma$ to the space $\mathcal{X}$ by associating to each subset $G$ of $\Gamma$ the collection of all points $g\omega$ where $g$ ranges over $G$. This is summarized in the following definition. As above, the symbol $\mathcal{P}(A)$ denotes the set of all subsets of $A$, $A$ an arbitrary set.

**Definition 3** The lift $\vartheta : \mathcal{P}(\mathcal{X}) \to \mathcal{P}(\Gamma)$ and canonical projection $\pi : \mathcal{P}(\Gamma) \to \mathcal{P}(\mathcal{X})$ are defined by

$$\vartheta(X) = \{g \in \Gamma : g\omega \in X\}, \quad X \subseteq \mathcal{X}$$

$$\pi(G) = \{g\omega : g \in G\}, \quad G \subseteq \Gamma.$$ 

For the case of Example 2, these formulas specialize to

$$\vartheta(X) = \bigcup_{x \in X} \tau_x \mathbb{R} = \tau(X) \oplus \mathbb{R}, \quad (17)$$

where

$$\tau(X) := \{ \tau_x : x \in X \}. \quad (18)$$

In [13,15] a construction was performed of various morphological operators between the distinct lattices $\mathcal{P}(\mathcal{X})$ and $\mathcal{P}(\Gamma)$. Here we restrict ourselves to dilations. That is, consider the mapping --- referred to as the hitting function below --- which associates to a subset $X$ of $\mathcal{X}$ the set of group elements $g \in \Gamma$ for which the translated set $gB := \{gb : b \in B\}$ hits $X$ (cf. (3)):

$$\mathcal{Q}_B(X) := \{g \in \Gamma : gB \uparrow X\}. \quad (19)$$

Then it was shown in [15] that

$$\mathcal{Q}_B(X) = \{g \in \Gamma : g\vartheta(B) \uparrow \vartheta(X)\}$$

$$= \vartheta(X) \oplus \vartheta^{-1}(B). \quad (20)$$

Here $\vartheta(X) \oplus \vartheta^{-1}(B)$ is a generalized Minkowski operation on subsets of $\Gamma$ as defined in (8), with $\vartheta^{-1}(B)$ the inverted set of $\vartheta(B)$, cf. (9). This mapping is
• a dilation $\mathcal{P}(\mathcal{X}) \rightarrow \mathcal{P}(\Gamma)$;

• invariant under $\Gamma$, that is
  
  $Q_B(g(X)) := gQ_B(X), \quad g \in \Gamma.$

More generally, if $A$ is a subset of $\Gamma$ and $x \mapsto g_x$ a function from $\mathcal{X}$ to $\Gamma$, the mapping

$$\delta^\Gamma_A(X) := \bigcup_{x \in X} g_xA,$$

is a dilation $\mathcal{P}(\mathcal{X}) \rightarrow \mathcal{P}(\Gamma)$ which is $\Gamma$-invariant. For this reason we speak sometimes of group dilations, or $\Gamma$-dilations. Dilations (and erosions) from $\mathcal{P}(\Gamma)$ to $\mathcal{P}(\mathcal{X})$ or from $\mathcal{P}(\mathcal{X})$ to $\mathcal{P}(\mathcal{X})$ can be constructed similarly but are not needed in the sequel.

### 2.6 The role of symmetry groups in shape description

We make a short remark concerning the problem how to define ‘shape’, which is known to present great difficulties and is a recurring theme in the image processing and computer vision literature. Often, ‘shape’ is defined as referring to those properties of geometrical figures which are invariant under the Euclidean similarity group [8]. Intuitively spoken, one first has to bring figures to a standard location, orientation and scale before being able to ‘compare’ them. Now it is not necessary to restrict oneself to the similarity group, although in the absence of any form of group invariance there is no way at all for comparing figures. In the present context the following definition seems appropriate.

**Definition 4** Let $\mathcal{X}$ be a set, $\Gamma$ a group acting on $\mathcal{X}$. Two subsets $X, Y$ of $\mathcal{X}$ are said to have the same shape with respect to $\Gamma$, or the same $\Gamma$-shape, if they are $\Gamma$-equivalent, meaning that there is a $g \in \Gamma$ such that $Y = gX$. If no such $g \in \Gamma$ exists, $X$ and $Y$ are said to have different $\Gamma$-shape.

In essence this definition goes back to F. Klein’s Erlanger Program (1872), which considers geometry to be the study of transformation groups and the properties invariant under these groups [9]. So in Euclidean morphology, all translates of a set $X$ by the Euclidean translation group $\mathcal{T}$ have the same $\mathcal{T}$-shape. Adding rotations to get the Euclidean motion group $\mathcal{M}$, rotated versions of $X$ or its translates have the same $\mathcal{M}$-shape as $X$. Extreme cases are (i) $\Gamma = \{\text{id}\}$, so that all sets have different shape, and (ii) $\Gamma = \text{Sym}_X$ (full permutation group), in which case all sets with the same cardinality have the same shape.

### 3 Relevance for computer vision

Above we have outlined how Euclidean morphology can be generalized to arbitrary homogeneous spaces ($\mathcal{X}, \Gamma$), where the group $\Gamma$ acting on $\mathcal{X}$ is not necessarily commutative. The case where $\Gamma$ acts simply transitively on $\mathcal{X}$ leads to the study
of transformations of subsets of an arbitrary group which are invariant under either left or right group translations. The general case where \( \Gamma \) acts transitively on \( \mathcal{X} \) has been treated by (i) mapping the subsets of \( \mathcal{X} \) to subsets of \( \Gamma \); (ii) using the results for the simply transitive case, and (iii) projecting back to the original space. The main result is that the scope of mathematical morphology is widened to situations where a non-commutative group is involved.

3.1 Perspective transformations

The application to computer vision concerns in particular the behavior of morphological operators under perspective transformations. This is a difficult problem and we confine us here to a sketch of a simplified case which nevertheless exhibits some of the essential ingredients of the general case. The emphasis will be on combining invariance concepts as relevant for computer vision in general with the construction of invariant morphological operators as outlined above. In a real 3D situation we will have to consider the full (non-abelian) group of projective transformations. But this only makes sense if at the same time we take the occlusion problem into account (see the introduction).

Consider the perspective transformation of a plane shape lying in a plane \( V \) which we call the object plane, see Fig. 3. As is well known, 3D Euclidean motions of the plane \( V \) induce 2D projection transformations on the image plane under perspective projection. Subgroups of the projection group apply when the motion of the planar object is constrained, for example to rotations and/or translations within the object plane.

Effect of object plane translations

The case to which we confine ourselves here is that of a translation only within the object plane. We will derive the corresponding morphological operators such as dilation and erosion, and show how this reduces to the Euclidean case when the object plane is parallel to the image plane.

Let the object plane \( V \) have equation

\[
AX + BY + CZ + D = 0.
\]  

(22)

We use a camera-centered coordinate system with coordinates \( X, Y, Z \) in object space, so the image plane \( \mathcal{X} \) is parallel to the \( X-Y \) plane at distance \( f \) above it (\( f \) is the focal length) with the viewpoint located at the origin \( (0,0,0) \). We use coordinates \( x, y \) in the image plane, with the \( x \) and \( y \) axes parallel to the \( X \) and \( Y \) axes, respectively. We also assume that \( C \neq 0 \), i.e., the object plane \( V \) has a nonzero slope w.r.t. the \( Z \)-axis.

Then we have that a point \( (X,Y,Z) \) on the object plane is projected to the point \( (x,y) \) on the image plane, where

\[
x = fX/Z, \quad y = fY/Z.
\]
Figure 3: Perspective projection of a planar object ABCD on the image plane.

We will use $\pi$ to denote this projection:

$$(x, y) = \pi(X, Y, Z).$$ (23)

Conversely, a point $\bar{x} = (x, y)$ on the image plane and below the vanishing line is mapped to a unique point $\vartheta(x, y) = (Zx/f, Zy/f, Z)$ on the object plane, where $Z$ is given by (use (22)):

$$Z = \frac{-Df}{Ax + By + Cf}.$$ (24)

Therefore we find

$$\vartheta(x, y) = \frac{-D}{N(\bar{x})} (x, y, f)$$ (25)

where

$$N(\bar{x}) = Ax + By + Cf.$$ (26)

Now consider a translation of the plane shape within the object plane $V$. The translation vector $\bar{\tau}$ has to parallel to $V$, i.e., perpendicular to the normal $(A, B, C')$ of $V$, so

$$A\tau_1 + B\tau_2 + C'\tau_3 = 0.$$ (27)

Using the notation of Sect. 2, the corresponding transformation $g_{\bar{\tau}}$ acting on points of the image plane is given by

$$g_{\bar{\tau}} \bar{x} = \pi [\vartheta(\bar{x}) + \bar{\tau}].$$ (28)
Using (25) and (23) we obtain
\[ g_\tau \vec{x} = \frac{f}{-Df/N(\vec{x}) + \tau_3} (-Dx/N(\vec{x}) + \tau_1, -Dy/N(\vec{x}) + \tau_2). \]  
(29)

It is clear that \( g_\tau \circ g_{\tau'} = g_{\tau + \tau'}, \) that \( g_{\tau}^{-1} = g_{-\tau} \) and that \( g_0 \) is the identity transformation, where \( \vec{0} \) denotes the origin of the image plane.

**Parametrization of the translations**

There is a 1-1 correspondence between translations parallel to \( V \) and points of the image plane below the vanishing line: given such a point \( \vec{x} = (x, y) \), there is a unique translation \( \vec{\tau}(\vec{x}) \) such that
\[ g_{\vec{\tau}(\vec{x})}\vec{0} = \vec{x}. \]  
(30)
So from (29) we see that \( \vec{\tau}(\vec{x}) = (\tau_1, \tau_2, \tau_3) \) has to satisfy
\[ \frac{f}{-D/C + \tau_3} (\tau_1, \tau_2) = (x, y). \]  
(31)
This is a pair of equations which together with (27) has the unique solution
\[
\begin{align*}
\tau_1 &= \frac{-Dx}{Ax + By + Cf} \\
\tau_2 &= \frac{-Dy}{Ax + By + Cf} \\
\tau_3 &= \frac{(D/C)Ax + By}{Ax + By + Cf}
\end{align*}
\]  
(32)
(33)
(34)

Using (26) we can write this more compactly as
\[ \vec{\tau}(\vec{x}) = \frac{D}{N(\vec{x})} (-x, -y, \frac{N(\vec{x})}{C} - f). \]  
(35)

**Group operation on points of the image plane**

Using the parametrization found above we can now define a group operation between points \( \vec{x} = (x, y) \) and \( \vec{x}' = (x', y') \) of the image plane below the vanishing line, which mimics the action of the translation group \( T \) on \( V \):
\[ \vec{x} \ast \vec{x}' := g_{\vec{x}} \circ g_{\vec{x}'} \vec{0} = g_{\vec{\tau}(\vec{x}) + \vec{\tau}(\vec{x}')} \vec{0} \]  
(36)

Here the superscript ‘\( P \)’ in \( \ast \) stands for ‘perspective’.

Using (29) and (35) we obtain after doing the algebra
\[ \vec{x} \ast \vec{x}' = \frac{fC \left( x/N(\vec{x}) + x'/N(\vec{x}'), y/N(\vec{x}) + y'/N(\vec{x}') \right)}{-1 + fC/N(\vec{x}) + fC/N(\vec{x}')}. \]
\[
\frac{fC \left( \bar{x} / N(\bar{x}) + \bar{x}' / N(\bar{x}') \right)}{-1 + fC / N(\bar{x}) + fC / N(\bar{x}')}. \tag{37}
\]

Note that the result does not depend on \( D \); the distance of the object plane \( V \) to the camera is irrelevant.

The identity element under this group operation is \( \bar{0} = (0,0) \) and the inverse of a point \( \bar{x} \) is
\[
\bar{x}^{-1} = g_{-\varphi(x)} \bar{0} = \frac{-Cf}{2N(\bar{x}) - Cf} \bar{x}. \tag{38}
\]

\[
\bar{x}^{-1} = \frac{-Cf}{2N(\bar{x}) - Cf} \bar{x}. \tag{39}
\]

**Remark 5** Is is clear from (37) that the group operation \( \mathcal{P} \) is abelian. This is of course due to the commutativity of the translation group \( \mathbf{T} \) acting on \( V \).

**Remark 6** As a special case we consider an object plane parallel to the image plane: \( A = B = 0 \). In that case we find that
\[
\bar{x} \mathcal{P} \bar{x}' = \bar{x} + \bar{x}'. \tag{40}
\]

This is the ordinary vector addition leading to the classical Minkowski addition and subtraction of Sect. 2.1.

**Perspective Minkowski operations**

Now that we have found a commutative group operation on points of the image plane, it is time to define the corresponding ‘perspective’ Minkowski operations using the standard recipe of Sect. 2.2. We formulate the result as a theorem.

**Theorem 7** Let \( \mathcal{P} \) be the group operation between points \( \bar{x} = (x,y) \) and \( \bar{x}' = (x',y') \) of the image plane induced by translations \( \mathbf{T} \) within the object plane defined by the equation \( AX + BY + CZ + D = 0 \) (\( C \neq 0 \)):
\[
\bar{x} \mathcal{P} \bar{x}' = \frac{fC \left( \bar{x} / N(\bar{x}) + \bar{x}' / N(\bar{x}') \right)}{-1 + fC / N(\bar{x}) + fC / N(\bar{x}')}. \tag{41}
\]

where \( N(\bar{x}) = Ax + By + Cf \).

Then the dilation \( \delta(G) \) and erosion \( \epsilon(G) \) of subsets \( G \) of the image plane \( \mathcal{X} \) invariant w.r.t. the group \( \mathcal{P} \) induced by \( \mathbf{T} \) are given by:
\[
\delta(G) := G \mathcal{P} H := \bigcup_{\bar{x} \in H} \bar{x} \mathcal{P} \bar{x} = \bigcup_{\bar{x} \in G} \bar{x} H \tag{42}
\]
\[
\epsilon(G) := G \mathcal{P} H := \bigcap_{\bar{x}' \in H} \bar{x}' \mathcal{P} \bar{x}'^{-1} \tag{43}
\]

where
\[
\bar{x} H := \{ \bar{x} \mathcal{P} \bar{x}' : \bar{x}' \in H \}, \quad G \bar{x}' := \{ \bar{x} \mathcal{P} \bar{x}' : \bar{x} \in G \}.
\]
Figure 4: Adapting the structuring element (black regions) under perspective projection.

Practical relevance

Suppose one takes pictures by a camera from a scene, such as occurs in the analysis of traffic scenes. If morphological analysis of such pictures is applied in order to recognize the shape of automobiles, the size of the structuring element has to be adapted according to the law of perspective, see Fig. 4. This is exactly what the theorem encapsulates.

4 Robot vision for path planning

Here the problem is to find a path for an object, say a robot or a car, moving in a space with obstacles. The problem falls apart into two distinct subproblems [3]. First, the empty-space problem: find the allowed states of the robot. Any possible configuration of the robot is represented as a point in a configuration space $C$, whose dimensionality equals the number of degrees of freedom of the robot. Points in $C$ such that a robot in that configuration would collide with any of the obstacles in work space are ‘forbidden’. The set of allowed points of $C$ is called ‘empty-space’. The second problem to be solved is the find-path problem: find a trajectory in the empty space, where the definition of ‘trajectory’ has to specify which transitions between allowed states are permissible. An approach based on mathematical morphology is able to solve the empty-space problem. If only translations of the robot are possible, one can find allowed positions of the (arbitrarily chosen) center of the robot by an ordinary erosion of the space outside the obstacles, where the structuring element $B$ is the robot itself. Equivalently, one may perform the dilation by the reflected set $\hat{B}$ of the set of obstacles to find the forbidden positions of the center of the robot. If the robot has rotational degrees of freedom one has to perform dilations with all rotated versions of the robot. The problem becomes even more difficult when the robot has internal degrees of freedom, see for example, for a robot with several rotating joints. For a full discussion see [16]. In
this subsection the allowed motion group equals either the translation group or the translation-rotation group.

4.1 Translations only

Consider a robot moving in the plane $\mathbb{R}^2$ with obstacles. The robot corresponds to a subset $B$ of the plane and the obstacles to another subset, say $X$. The problem is to find the set of forbidden configurations. The state of the robot is parametrized by the location $h$ of an arbitrary point of the robot $B$, initially at the origin; hence the configuration space $\mathcal{C}$ is identical to $\mathbb{R}^2$. The allowed motions form the translation group $T$ which can be identified with $\mathbb{R}^2$, see Example 1 above. The forbidden points can be identified with the set

$$Q_B(X) = \{ h \in \mathbb{R}^2 : \tau_h B \cap X \}. \quad (44)$$

One immediately recognizes this (cf. (3)) as the Euclidean dilation of $X$ by $\bar{B}$. This leads to the first result.

**Proposition 8** If $B \in \mathbb{R}^2$ is a robot with translational degrees of freedom, then the hitting function $Q_B : \mathcal{P}(\mathbb{R}^2) \to \mathcal{P}(\mathcal{C})$ is given by

$$Q_B(X) = X \overset{T}{\uplus} \bar{B}$$

$$=: \delta_B^T(X) = \bigcup_{x \in X} \tau_x \bar{B}. \quad (45)$$

For clarity the dependence of the Minkowski sum on the Euclidean translation group $T$ is explicitly indicated in (45). An example can be found in Fig. 5.
4.2 Translations and rotations

Next the case of a mobile robot with translational and rotational degrees of freedom is considered. The appropriate group is $\mathbf{M}$, the Euclidean motion group, see Example 2. To parametrize the state of the robot, choose two distinct points, say $P_1$ and $P_2$ inside the robot. The configuration space $C$ is 3-dimensional in this case,

$$C := \{(h, \phi) : h \in \mathbb{R}^2, \phi \in S^1\},$$

with $h$ the location of point $P_1$ and $\phi$ the angle of the line segment $[P_1P_2]$ with respect to the $x$-axis. Note that $C$ can be identified with the parameter space of the Euclidean motion group $\mathbf{M}$: to each $(h, \phi)$ in $C$ corresponds a unique motion $\gamma_{h,\phi} = \tau_h \rho^0_\phi$ and vice versa. This identification will tacitly be made below without further comment. An alternative way to represent $C$ is by means of pointers; see Sect. 2.4. Assuming the initial state to be equal to $\{h = (0,0), \phi = 0\}$, the hitting function in this case becomes

$$Q_B(X) = \{(h, \phi) \in C : \gamma_{h,\phi} B \uparrow X\}.$$ 

Now the results of Sect. 2.5 are applicable. From the general result (20) together with (17)–(18) it follows that

$$Q_B(X) = \vartheta(X)^{\mathbf{M}} + \vartheta^{-1}(B) = \vartheta(X)^{\mathbf{M}} \uplus (\mathbb{R}^+ \uplus \tau^{-1}(B))$$

where (cf. (18))

$$\tau^{-1}(B) = \{\tau_b^{-1} : b \in B\} = \tau(\check{B}),$$

with $\check{B}$ the reflected set of $B$. From (48) one obtains the following result.

**Proposition 9** If $B \in \mathbb{R}^2$ is a robot with translational and rotational degrees of freedom, then the hitting function $Q_B : \mathcal{P}(\mathbb{R}^2) \to \mathcal{P}(C)$ is given by

$$Q_B(X) = \bigcup_{x \in X} \tau_x (\mathbb{R}^+ \uplus \tau^{-1}(B))$$

$$= \bigcup_{\phi \in S^1} \bigcup_{x \in X} \tau_x \rho^0_\phi \tau(\check{B})$$

The equality (51) expresses the fact, which is obvious from (47), that $Q_B(X)$ can be found by doing, for each $\phi \in S^1$ an ordinary dilation with a rotated version $\rho^0_\phi B$ of the structuring element. Equation (50) says that $Q_B(X)$ can also be found as a union of translates, that is, by a dilation

$$\delta_B^T(X) := \bigcup_{x \in X} \tau_x \check{B},$$

where

$$\check{B} := \mathbb{R}^+ \uplus \tau^{-1}(B).$$
Figure 6: The forbidden set $Q_B(X)$ for a robot $B$ with translational and rotational degrees of freedom. Heavy dots: the obstacle space $X$; $b$: base pointer. Arrows attached to heavy and open dots: the forbidden set.

Eq. (52) differs from a usual T-dilation through the fact that the structuring element (53) is not a planar, but a 3D subset of $C$. The construction of $\hat{B}$ is straightforward: (i) take the reflection $\hat{\tilde{B}}$ of $B$; (ii) lift $\tilde{B}$ to $C$ by applying $\tau$; (iii) construct rotated copies of $\tau(\hat{\tilde{B}})$ in $C$.

These results can be nicely presented geometrically using the representation by pointers, see Fig. 6. Alternatively, one may use a 3D representation in configuration space $C$.

5 Conclusions

In this paper we have presented an extension of mathematical morphology with a larger symmetry group than the group of Euclidean translations. It has been shown how to obtain generalizations of morphological transformations for arbitrary homogeneous spaces.

Using this framework, one of the limitations mathematical morphology is addressed here: how to take the projective geometry of the imaging process into account. In Sect. 3 we explained how to define morphological operations on images of a planar patch produced by perspective projection, where the allowed 3D motion is restricted to translation within the object plane. The main result (Theorem 7) justifies existing procedures of adapting the structuring element as a function of the position in the image plane w.r.t. the vanishing line.

Finally in Sect. 4 a problem from robot vision for path planning was approached by morphological methods. Here the problem is to find a path for an object, say a robot or a car, moving in a space with obstacles. It was shown, both when translations and rotations of the robot are allowed, how morphological methods may be used to solve the empty-space problem: find the allowed states of the robot.
6 References


