Centrum voor Wiskunde en Informatica

REPORT*RAPPORT*

Morphological operators for image sequences

J. Goutsias, H.J.A.M. Heijmans, K. Sivakumar

Department of Operations Research, Statistics, and System Theory

Report BS-R9419 May 1994

CWI is the National Research Institute for Mathematics and Computer Science. CWI is part of the Stichting Mathematisch Centrum (SMC), the Dutch foundation for promotion of mathematics and computer science and their applications.
SMC is sponsored by the Netherlands Organization for Scientific Research (NWO). CWI is a member of ERCIM, the European Research Consortium for Informatics and Mathematics.

P.O. Box 94079, 1090 GB Amsterdam (NL)
Kruislaan 413, 1098 SJ Amsterdam (NL)
Telephone +31 20 592 9333
Telefax +31 20 592 4199

Morphological Operators for Image Sequences

John Goutsias

Department of Electrical and Computer Engineering
Image Analysis and Communications Laboratory
The Johns Hopkins University, Baltimore, MD 21218, USA

Henk J.A.M. Heijmans

CWI
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands

K. Sivakumar

Department of Electrical and Computer Engineering
Image Analysis and Communications Laboratory
The Johns Hopkins University, Baltimore, MD 21218, USA

Abstract

This paper presents a unifying approach to the problem of morphologically processing image sequences (or, equivalently, vector-valued images), by means of lattice theory, thus providing a mathematical foundation for vector morphology. Lattice theory is an abstract algebraic tool that has been extensively used as a theoretical framework for scalar morphology (i.e., mathematical morphology applied on single images). Two approaches to vector morphology are discussed. According to the first approach, vector morphology is viewed as a natural extension of the well known scalar morphology. This approach formalizes and generalizes Wilson's matrix morphology and shows that the latter is a direct consequence of marginal vector ordering. The derivation of the second approach is more delicate and requires careful treatment. This approach is a direct consequence of a vector transformation followed by marginal ordering. When the vector transformation is the identity transformation, the two approaches are equivalent. A number of examples demonstrate the applicability of the proposed theory in a number of image processing and analysis problems.

AMS Subject Classification (1991): 68U10

Keywords & Phrases: image sequences, mathematical morphology, vector morphology, matrix morphology, lattice theory, dilation, erosion, adjunction, h-adjunction, h-dilation, h-erosion,h-filter, vector ordering, maximum noise fraction (MNF) transform, Mahalanobis distance, vector distance transform

1. Introduction

In recent years, the wide applicability of fast and affordable computer hardware, and the increasing sophistication in image data gathering, have resulted in problems which require processing

Report BS-R9419
ISSN 0924-0659
CWI
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands

and analysis of *image sequences.* One good example is in remote sensing, where a sequence $\{I_k \mid k = 1, 2, \ldots, K\}$ of K multi-spectral images is usually available, with K typically being between 2 and 12 [16, 28]. Color image processing is another area of increasing interest, due, primarily to the recent development of high resolution video systems (e.g., HDTV systems). In this case, three images are usually available, each containing information about the red, green, and blue color image components [16]. Another interesting example is in magnetic resonance imaging, where a number of images (e.g., proton density (PD) weighted, T1-weighted, and T2-weighted [39]) are recorded, each image containing a certain type of tissue information, which corresponds to the same anatomic site of interest.

A straightforward approach to processing these images is by means of a component-wise transformation (i.e., a transformation which is independently applied on each image component). However, this approach is bound to fail, since, in most cases, image components are highly correlated. Two simple examples, demonstrating the pitfalls of carelessly applying component-wise processing, can be found in [1]. Therefore, it is strongly recommended that the design of a multi-frame image processing and analysis module should take into consideration the existence of correlations among images in the same sequence [1, 11, 16, 25, 28, 35]. In this paper, we shall consider the problem of processing image sequences by means of *mathematical morphology* [10, 13, 33, 34]. So far, mathematical morphology has been primarily developed for the case of single (scalar) images. If we choose to apply this technique on image sequences, we are immediately limited to component-wise processing. Therefore, the need for developing a theory of *vector morphology* is imminent.

Recently, a number of morphological techniques have been proposed for dealing with the problem of processing vector images; see [35], [42, 43, 44] and [8]. One of these techniques is Wilson's *matrix morphology* [44]; see also [42, 43]. This technique is based on a direct generalization of the standard *scalar morphology* concepts, achieved by imposing a vector space structure on the image sequences under consideration, and by employing standard translation invariant erosions and dilations. A number of interesting examples, which demonstrate the effectiveness of matrix morphology in a variety of image processing and analysis applications, can be found in [44] and [45]. However, matrix morphology is somehow limited by, mainly, two factors.

It has been pointed out in [13, 14, 34] that there exist applications which require development of more general types of erosions and dilations. Furthermore, as we shall demonstrate later in this paper, Wilson's matrix morphology is a direct consequence of a marginal vector ordering principle, which may not be desirable in certain applications. Therefore, it would be natural to investigate whether matrix morphology is the only type of vector morphology. As we shall demonstrate in this paper, matrix morphology is a special case of a more general approach. However, and in order to develop such an approach, we shall need to employ *lattice theory* [3].

Lattice theory has proven itself to be fundamental in mathematical morphology, since it provides a powerful tool for understanding and abstracting a number of important morphological concepts [13, 14, 29, 30, 34]. This theory has been recently employed in [35] and [8] for the development of a general approach to vector morphology. The majority of the work in [35] is oriented towards constructing vector morphological transformations by means of set primitives and flat operators. On the other hand, the work in [8], based on generalizing fundamental lattice theoretic concepts to the case of image sequences, is directly related to the theory of matrix morphology, and it complements the work in [35].

The present paper is an extended version of [8], and is structured as follows: Section 2 provides a brief review of the theory of complete lattices and adjunctions, and introduces the new notion of h-adjunction. Furthermore, it contains a general discussion on two complete lattices for image sequences. In Section 3, a more specialized discussion is provided, regarding the two complete lattices derived in Section 2, which leads to two distinct, in general, approaches to vector morphology. Finally, in Section 4, a number of image processing and analysis examples demonstrate the utility of the proposed approaches in practical applications. Additionally, it is shown that Wilson's matrix morphology is a special case of the first vector morphology approach and that the widely used Euclidean distance transform can be obtained by means of the second approach.

2. Complete lattices and image sequences

In this section, we shall review the fundamentals of the theory of complete lattices and introduce two important classes of complete lattices of image sequences. For a detailed exposition on complete lattice theory refer to [3] and [20]. We shall first repeat some results from this theory for notational purposes and completeness. Then we introduce h-adjunctions, a new concept which forms the basis of our theory. The connection of complete lattice theory with mathematical morphology is thoroughly discussed in [34], [14, 30], and [13].

2.1. Complete lattice theory fundamentals

A set $\mathcal{L}$ with a partial ordering $\leq$ is called a *complete lattice* if every subset $\mathcal{H}$ of $\mathcal{L}$ has a least upper bound (supremum) $\bigvee \mathcal{H}$ and a greatest lower bound (infimum) $\bigwedge \mathcal{H}$.

Let $\mathcal{L}, \mathcal{M}$ be two complete lattices. The set of all operators mapping $\mathcal{L}$ into $\mathcal{M}$ forms a complete lattice under the partial ordering given by

$$\phi \leq \psi \iff \phi(X) \leq \psi(X), \quad \text{for all } X \in \mathcal{L}.$$

The identity operator on $\mathcal{L}$, which maps every element onto itself, is denoted by $\mathrm{id}_{\mathcal{L}}$, or, when no confusion is possible, by id. An operator $\psi : \mathcal{L} \to \mathcal{M}$ is *increasing* if $X \leq Y$ implies that $\psi(X) \leq \psi(Y)$. It is called a *dilation* if $\psi(\bigvee_{i\in I} X_i) = \bigvee_{i\in I} \psi(X_i)$, for every collection $\{X_i \mid i \in I\}$. Dually, it is called an *erosion* if $\psi(\bigwedge_{i\in I} X_i) = \bigwedge_{i\in I} \psi(X_i)$. Let $\varepsilon : \mathcal{L} \to \mathcal{M}$ and $\delta : \mathcal{M} \to \mathcal{L}$ be two operators; we say that the pair (ε, δ) is an *adjunction* between $\mathcal{L}$ and $\mathcal{M}$ if

$$\delta(Y) \leq X \iff Y \leq \varepsilon(X), \quad \text{for all } X \in \mathcal{L}, Y \in \mathcal{M}.$$

If (ε, δ) is an adjunction, then ε is an erosion and δ a dilation. With every erosion $\varepsilon : \mathcal{L} \to \mathcal{M}$ there corresponds a unique dilation $\delta : \mathcal{M} \to \mathcal{L}$ such that (ε, δ) is an adjunction. Similarly, with every dilation δ one can associate a unique erosion ε such that this holds. We say that ε and δ are adjoint to each other.

Assume that (ε, δ) is an adjunction between $\mathcal{L}$ and $\mathcal{M}$. Then

$$\varepsilon\delta\varepsilon = \varepsilon \quad \text{and} \quad \delta\varepsilon\delta = \delta.$$

Furthermore,

$$\delta\varepsilon \leq \mathrm{id}_{\mathcal{L}} \quad \text{and} \quad \varepsilon\delta \geq \mathrm{id}_{\mathcal{M}},$$

where $\mathrm{id}_{\mathcal{L}}, \mathrm{id}_{\mathcal{M}}$ are the identity operators on $\mathcal{L}$ and $\mathcal{M}$, respectively. The operators $\delta\varepsilon$ and $\varepsilon\delta$ form an *opening* on $\mathcal{L}$ and a *closing* on $\mathcal{M}$, respectively.

An operator $\psi : \mathcal{L} \to \mathcal{L}$ is called a *lattice isomorphism* if ψ is bijective and $X \leq Y$ iff $\psi(X) \leq \psi(Y)$, for all $X, Y \in \mathcal{L}$. In this case, the pairs (ψ^{-1}, ψ) and (ψ, ψ^{-1}) form adjunctions on $\mathcal{L}$. Further results about adjunctions can be found in [14] and [13].

We conclude this subsection with a more detailed examination of adjunctions on the complete lattice $\overline{\mathbb{R}}$, as these play an important role in the sequel of this paper.

2.1. Example. In this paper we use the following convention: in general, dilations and erosions are denoted by δ and ε, respectively. However, if the underlying lattice has the interpretation of a set of grey-values (one- or multi-dimensional), then we use d and e instead.

First we note (see [13, Lemma 11.22]) that d is a dilation on $\overline{\mathbb{R}}$ if d is increasing, continuous from the left, and satisfies $d(-\infty) = -\infty$. Dually, e is an erosion on $\overline{\mathbb{R}}$ if e is increasing, continuous from the right, and satisfies $e(+\infty) = +\infty$. In order that (e, d) constitutes an adjunction we must have

$$e(t) = \bigvee\{s \in \overline{\mathbb{R}} \mid d(s) \le t\} \text{ and } d(t) = \bigwedge\{s \in \overline{\mathbb{R}} \mid t \le e(s)\}.$$

The pair $d(t) = at + b$, $e(t) = (t - b)/a$, where $a > 0$ and $b \in \overline{\mathbb{R}}$, forms an adjunction on $\overline{\mathbb{R}}$. Here we define $s + t = -\infty$ if $s = -\infty$ or $t = -\infty$, and $s - t = +\infty$ if $s = +\infty$ or $t = -\infty$.

To conclude this section, we consider power lattices as these play an important role in this paper. Let $\mathcal{L}$ be a complete lattice and P an arbitrary nonempty set. We denote the elements of the power set $\mathcal{L}^P$ by $\mathbf{X}$. For a given $p \in P$, the value of $\mathbf{X}$ at p is denoted by X_p. The set $\mathcal{L}^P$ is a complete lattice, known as a *complete power lattice*, with the pointwise ordering

$$\mathbf{X} \le \mathbf{Y} \text{ iff } X_p \le Y_p,\ p \in P,$$

whenever $\mathbf{X}, \mathbf{Y} \in \mathcal{L}^P$; see [13, Ex. 2.10].

Adjunctions between a complete power lattice $\mathcal{L}^P$ and another complete power lattice $\mathcal{M}^K$ may be written in terms of adjuctions between $\mathcal{L}$ and $\mathcal{M}$; refer to [13, Prop. 5.3].

2.2. Proposition. *The pair (ε, δ) is an adjunction between the complete power lattices $\mathcal{L}^P$ and $\mathcal{M}^K$ if and only if there exist adjunctions $(\varepsilon_{p,k}, \delta_{k,p})$ between $\mathcal{L}$ and $\mathcal{M}$, for all $p \in P$, $k \in K$, such that*

$$(\varepsilon(\mathbf{X}))_k = \bigwedge_{p\in P} \varepsilon_{p,k}(X_p),$$

$$(\delta(\mathbf{Y}))_p = \bigvee_{k\in K} \delta_{k,p}(Y_k),$$

for $\mathbf{X} \in \mathcal{L}^P, \mathbf{Y} \in \mathcal{M}^K$.

2.2. h-adjunctions

Assume that $\mathcal{R}$ is a complete lattice and that $\mathcal{T}$ is a nonempty set. Furthermore, let $h : \mathcal{T} \to \mathcal{R}$ be a surjective mapping. Define an equivalence relation $=_h$ on $\mathcal{T}$ as follows:

$$t =_h t' \iff h(t) = h(t'), \quad t, t' \in \mathcal{T}.$$

We define another relation $\le_h$ on $\mathcal{T}$ in the following way:

$$t \le_h t' \iff h(t) \le h(t'), \quad t, t' \in \mathcal{T}.$$

It is evident that this relation is reflexive ($t \le_h t$) and transitive ($t_1 \le_h t_2$ and $t_2 \le_h t_3$ implies that $t_1 \le_h t_3$). However, $\le_h$ is not a partial ordering because $t \le_h t'$ and $t' \le_h t$ implies only that $t =_h t'$ but not $t = t'$. We refer to $\le_h$ as the *h-ordering*.

We need some further definitions.

2.3. Definition. An operator $\psi : \mathcal{T} \to \mathcal{T}$ is *h-increasing* if $t \le_h t'$ implies that $\psi(t) \le_h \psi(t')$.

It is easy to see that the composition of two h-increasing operators is again h-increasing. Let ψ_1, ψ_2 be two operators on $\mathcal{T}$; we write $\psi_1 \leq_h \psi_2$ if $\psi_1(t) \leq_h \psi_2(t)$ for every $t \in \mathcal{T}$.

Let $r \in \mathcal{R}$; since h is surjective, there exists a $t_r \in \mathcal{T}$ such that $h(t_r) = r$. Define the equivalence class $\mathcal{T}[r] = \{t \in \mathcal{T} \mid h(t) = r\}$. The Axiom of Choice [7] implies that there exist mappings $h^{\leftarrow} : \mathcal{R} \to \mathcal{T}$ such that

$$hh^{\leftarrow}(r) = r, \quad \text{for } r \in \mathcal{R}.$$

Unless h is injective, there exist more than one (possibly an infinite number of) such "inverse" mappings. Note that $h^{\leftarrow}h$ is not the identity mapping in general (but $h^{\leftarrow}h =_h \mathsf{id}$); we call $h^{\leftarrow}$ the *semi-inverse* of h. For every semi-inverse $h^{\leftarrow}$ and for every $r \in \mathcal{R}$ we have $h^{\leftarrow}(r) \in \mathcal{T}[r]$.

2.4. Lemma. *Let $\psi : \mathcal{T} \to \mathcal{T}$ be h-increasing, and let $h^{\leftarrow}$ be a semi-inverse of h. Then*

$$h\psi h^{\leftarrow} h = h\psi. \tag{2.1}$$

PROOF. Since $h^{\leftarrow}h =_h \mathsf{id}$ we find that $\psi h^{\leftarrow}h =_h \psi$. This yields the result. ∎

2.5. Proposition. *A mapping $\psi : \mathcal{T} \to \mathcal{T}$ is h-increasing iff there exists an increasing mapping $\psi^{\circ} : \mathcal{R} \to \mathcal{R}$ such that*

$$\psi^{\circ} h = h\psi. \tag{2.2}$$

The mapping ψ° is uniquely determined by ψ, and can be computed from

$$\psi^{\circ} = h\psi h^{\leftarrow}, \tag{2.3}$$

where $h^{\leftarrow}$ is an arbitrary semi-inverse of h. The expression in (2.3) is independent of $h^{\leftarrow}$.

PROOF. 'if': assume that there exists an increasing mapping $\psi^{\circ} : \mathcal{R} \to \mathcal{R}$ such that $\psi^{\circ}h = h\psi$. Let $t, t' \in \mathcal{T}$ such that $t \leq_h t'$, that is, $h(t) \leq h(t')$. Then $h\psi(t) = \psi^{\circ}h(t) \leq \psi^{\circ}h(t') = h\psi(t')$, i.e., $\psi(t) \leq_h \psi(t')$.

'only if': assume that ψ is h-increasing. Choose an arbitrary semi-inverse $h^{\leftarrow}$ and define $\psi^{\circ} = h\psi h^{\leftarrow}$. Then $\psi^{\circ}h = h\psi h^{\leftarrow}h = h\psi$ by Lemma 2.4. We must show that ψ° is increasing. Let $r, r' \in \mathcal{R}$ such that $r \leq r'$, and define $t = h^{\leftarrow}(r)$, $t' = h^{\leftarrow}(r')$. Clearly, $t \leq_h t'$, hence $\psi(t) \leq_h \psi(t')$. This means that

$$\psi^{\circ}(r) = \psi^{\circ}hh^{\leftarrow}(r) = \psi^{\circ}h(t) = h\psi(t) \leq h\psi(t') = \psi^{\circ}h(t') = \psi^{\circ}hh^{\leftarrow}(r') = \psi^{\circ}(r'),$$

which shows that ψ° is increasing.

It remains to be shown that ψ° given by (2.3) does not depend on the choice of the semi-inverse $h^{\leftarrow}$. If $h_1^{\leftarrow}, h_2^{\leftarrow}$ are two different semi-inverses, then $h_1^{\leftarrow} =_h h_2^{\leftarrow}$, hence $\psi h_1^{\leftarrow} =_h \psi h_2^{\leftarrow}$, yielding that $h\psi h_1^{\leftarrow} = h\psi h_2^{\leftarrow}$. This concludes the proof. ∎

We write $\psi \overset{h}{\mapsto} \psi^{\circ}$ if the assumptions in Proposition 2.5 are satisfied. It is obvious that

$$\psi \overset{h}{\mapsto} \psi^{\circ} \iff h^{\leftarrow}h\psi \overset{h}{\mapsto} \psi^{\circ},$$

for every semi-inverse $h^{\leftarrow}$.

2.6. Definition. Let $\varepsilon, \delta : \mathcal{T} \to \mathcal{T}$ be two mappings with the property that for $s, t \in \mathcal{T}$,

$$\delta(s) \leq_h t \iff s \leq_h \varepsilon(t);$$

then the pair (ε, δ) is called an *h-adjunction*.

h-adjunctions inherit a large number of properties from ordinary adjunctions between complete lattices.

2.7. Proposition. *If (ε, δ) is an h-adjunction on $\mathcal{T}$, then both ε and δ are h-increasing.*

PROOF. We show here that ε is h-increasing; the proof for δ is similar. Let $t \leq_h t'$. We get

$$\begin{aligned} \varepsilon(t) \leq_h \varepsilon(t) &\Rightarrow \delta\varepsilon(t) \leq_h t \\ &\Rightarrow \delta\varepsilon(t) \leq_h t' \\ &\Rightarrow \varepsilon(t) \leq_h \varepsilon(t'). \end{aligned}$$

This proves the result. ∎

2.8. Proposition. *Let (ε, δ) be h-increasing mappings on $\mathcal{T}$, and let $\varepsilon \overset{h}{\mapsto} \varepsilon^\circ$, $\delta \overset{h}{\mapsto} \delta^\circ$. Then (ε, δ) is an h-adjunction on $\mathcal{T}$ iff $(\varepsilon^\circ, \delta^\circ)$ is an adjunction on $\mathcal{R}$.*

PROOF. We prove the 'only if'-statement. The proof of the 'if'-statement is left to the reader. Assume that (ε, δ) is an h-adjunction. Let $r, s \in \mathcal{R}$ and $\delta^\circ(r) \leq s$. We must show that $r \leq \varepsilon^\circ(s)$. Let $t_r, t_s \in \mathcal{T}$ be such that $h(t_r) = r$ and $h(t_s) = s$. Thus $\delta^\circ h(t_r) \leq h(t_s)$, i.e., $h\delta(t_r) \leq h(t_s)$. In other words, $\delta(t_r) \leq_h t_s$, which yields that $t_r \leq_h \varepsilon(t_s)$, i.e., $h(t_r) \leq h\varepsilon(t_s) = \varepsilon^\circ h(t_s)$. This means that $r \leq \varepsilon^\circ(s)$, which had to be shown. Analogously, one shows that $r \leq \varepsilon^\circ(s)$ implies $\delta^\circ(r) \leq s$. ∎

If (ε, δ) is an h-adjunction on $\mathcal{T}$ and if $\varepsilon' = h_1^{\leftarrow} h\varepsilon$, $\delta' = h_2^{\leftarrow} h\delta$, for two arbitrary semi-inverses $h_1^{\leftarrow}, h_2^{\leftarrow}$ of h, then (ε', δ') is an h-adjunction as well.

We have seen that for every lattice isomorphism ψ on a complete lattice $\mathcal{L}$ with inverse ψ^{-1}, the pairs (ψ, ψ^{-1}) and (ψ^{-1}, ψ) are adjunctions on $\mathcal{L}$. A similar result holds for h-adjunctions.

A mapping $\psi : \mathcal{T} \to \mathcal{T}$ which is h-increasing, bijective, and has an h-increasing inverse ψ^{-1} is called an *h-isomorphism* on $\mathcal{T}$.

2.9. Proposition. *Assume that ψ is an h-isomorphism on $\mathcal{T}$, then both (ψ, ψ^{-1}) and (ψ^{-1}, ψ) are h-adjunctions on $\mathcal{T}$.*

The proof is straightforward. Note that $\psi \overset{h}{\mapsto} \psi^\circ$ iff $\psi^{-1} \overset{h}{\mapsto} (\psi^\circ)^{-1}$. Indeed, if $\psi^{-1} \overset{h}{\mapsto} \phi^\circ$, then $\psi^\circ\phi^\circ = h\psi h^{\leftarrow} h\psi^{-1} h^{\leftarrow}$. However, since $h^{\leftarrow} h =_h \mathrm{id}$, then $\psi h^{\leftarrow} h =_h \psi$, or $h\psi h^{\leftarrow} h = h\psi$, and, therefore, $\psi^\circ\phi^\circ = h\psi\psi^{-1}h^{\leftarrow} = hh^{\leftarrow} = \mathrm{id}$. Similarly, $\phi^\circ\psi^\circ = \mathrm{id}$. Therefore, $\phi^\circ = (\psi^\circ)^{-1}$. A similar argument leads to the 'if' part.

We now present some simple examples.

2.10. Examples.
(a) Let $\mathcal{T} = \overline{\mathbb{R}}, \mathcal{R} = \overline{\mathbb{R}}_+$, and let $h : \mathcal{T} \to \mathcal{R}$ be given by $h(t) = |t|$. The pair (e, d) of mappings on $\overline{\mathbb{R}}$ given by $e(t) = t - b$, $d(t) = t + b$, where $b \in \overline{\mathbb{R}}$, is an adjunction on $\overline{\mathbb{R}}$; see Example 2.1. However, it is not an h-adjunction unless $b = 0$. Namely, in order that this pair is an h-adjunction the following condition must hold:

$$|s + b| \leq |t| \iff |s| \leq |t - b|,$$

for $s, t \in \overline{\mathbb{R}}$. It is easy to check that this holds if and only if $b = 0$. Note that $h^{\leftarrow} : \overline{\mathbb{R}}_+ \to \overline{\mathbb{R}}$ is a semi-inverse iff $h^{\leftarrow}(r) \in \{-r, r\}$ for $r \in \overline{\mathbb{R}}_+$.
(b) Let $\mathcal{T} = \overline{\mathbb{R}}^2, \mathcal{R} = \overline{\mathbb{R}}$, and $h(t_1, t_2) = t_1 + t_2$. In this case, the pair (ε, δ) given by

$$e(t_1, t_2) = (t_1 - a, t_2 - b) \text{ and } d(t_1, t_2) = (t_1 + a, t_2 + b),$$

with $a, b \in \overline{\mathbb{R}}$, defines an h-adjunction on $\overline{\mathbb{R}}^2$. Furthermore, $e \overset{h}{\mapsto} e^\circ, d \overset{h}{\mapsto} d^\circ$, where (e°, d°) is the adjunction on $\overline{\mathbb{R}}$ given by $e^\circ(r) = r - a - b$, $d^\circ(r) = r + a + b$.

Let $d' : \overline{\mathbb{R}}^2 \to \overline{\mathbb{R}}^2$ be given by $d'(t_1, t_2) = (t_1 + b, t_2 + a)$, then (e, d') is an h-adjunction, too.

2.11. Proposition. *Assume that (ε, δ) is an h-adjunction, then*

$$\delta\varepsilon \leq_h \mathsf{id} \leq_h \varepsilon\delta,$$

and also

$$\varepsilon\delta\varepsilon =_h \varepsilon \quad \textit{and} \quad \delta\varepsilon\delta =_h \delta.$$

The proof is straightforward. A mapping ψ on $\mathcal{T}$ with $\psi \overset{h}{\mapsto} \psi^{\circ}$ is called *h-dilation* (resp. *h-erosion*) if ψ° is a dilation (erosion) on $\mathcal{R}$. Similarly, an operator α on $\mathcal{T}$ is called an *h-opening* if the operator α° on $\mathcal{R}$ determined by $\alpha \overset{h}{\mapsto} \alpha^{\circ}$ is an opening. It is easy to check that any α is an h-opening iff it is h-increasing, and satisfies $\alpha^2 =_h \alpha$ (*h-idempotence*) and $\alpha \leq_h \mathsf{id}$ (*h-anti-extensivity*). The previous result shows that $\delta\varepsilon$ is an h-opening if (ε, δ) is an h-adjunction. An *h-closing* is similarly defined.

If he or she wishes, the reader can readily verify that many results from the theory of morphological filters have h-analogues. We illustrate this by means of the *alternating filters*, which are compositions of openings and closings [36]. An operator ψ on $\mathcal{T}$ is called an *h-filter* if it is h-increasing and satisfies $\psi^2 =_h \psi$.

2.12. Proposition. *Let α be an h-opening and β an h-closing on $\mathcal{T}$. Then all compositions of these operators are h-filters and*

$$\alpha \leq_h \alpha\beta\alpha \leq_h \left\{ \begin{matrix} \beta\alpha \\ \alpha\beta \end{matrix} \right\} \leq_h \beta\alpha\beta \leq_h \beta.$$

We leave the proof as an exercise to the reader. (Hint: use (2.1)-(2.2).)

2.3. Image sequences

Let $E, \mathcal{T}$ be nonempty sets. We denote by $\mathrm{Fun}(E, \mathcal{T})$ the power set $\mathcal{T}^E$, i.e., the functions from E into $\mathcal{T}$. If $\mathcal{T}$ is a complete lattice, then $\mathrm{Fun}(E, \mathcal{T})$ is a complete lattice too. Given an index set P, we denote by $\mathrm{Fun}(E, \mathcal{T})^P$ all image sequences indexed by P. It is obvious that $\mathrm{Fun}(E, \mathcal{T})^P$ is isomorphic with $\mathrm{Fun}(E, \mathcal{T}^P)$; both expressions yield two alternative representations of the same family. In the following, we shall examine their general structure. More details will be discussed in Section 3.

We denote the elements of $\mathrm{Fun}(E, \mathcal{T})^P$ by $\boldsymbol{F}$, and the value of $\boldsymbol{F}$ at index $p \in P$, which is an element of $\mathrm{Fun}(E, \mathcal{T})$, by F_p. Using the representation $\mathrm{Fun}(E, \mathcal{T}^P)$ the value of $\boldsymbol{F}$ at a point $x \in E$, which lies in $\mathcal{T}^P$, is denoted by $\boldsymbol{F}(x)$. Note that

$$\boldsymbol{F} \leq \boldsymbol{G} \iff F_p \leq G_p,\ p \in P \iff F_p(x) \leq G_p(x),\ p \in P, x \in E. \tag{2.4}$$

The inequality at the right hand-side has to be interpreted in the complete lattice $\mathcal{T}$.

Assume that $\mathcal{T}$ is a nonempty set, $\mathcal{R}$ is a complete lattice, and P, K are nonempty index sets. Furthermore, let $h : \mathcal{T}^P \to \mathcal{R}^K$ be a surjective mapping. We can extend h as a mapping $h : \mathrm{Fun}(E, \mathcal{T}^P) \to \mathrm{Fun}(E, \mathcal{R}^K)$ by putting

$$h(\boldsymbol{F})(x) = h(\boldsymbol{F}(x)),\ x \in E.$$

With this definition the relations $\leq_h$ and $=_h$ can also be extended to $\mathrm{Fun}(E, \mathcal{T}^P)$. Let, for example, $\boldsymbol{F}, \boldsymbol{G} \in \mathrm{Fun}(E, \mathcal{T}^P)$; we put

$$\boldsymbol{F} \leq_h \boldsymbol{G} \iff \boldsymbol{F}(x) \leq_h \boldsymbol{G}(x), \text{ for all } x \in E.$$

The properties of these relations remain also valid for their extensions.

To conclude this section we consider some examples for the case that $\mathcal{T} = \mathcal{R} = \overline{\mathbb{R}}$. First assume that $P = K$ and that h is the identity mapping $h(\boldsymbol{t}) = \boldsymbol{t}$. Then the partial ordering induced by h is the usual pointwise ordering (see (2.4)), also known as *marginal ordering* or *M-ordering* [2]; i.e., ordering of elements $\boldsymbol{t}$ in $\overline{\mathbb{R}}^P$ takes place within the marginal components of $\boldsymbol{t}$.

When K contains one element, h is a mapping from $\overline{\mathbb{R}}^P$ into $\overline{\mathbb{R}}$, and in this case the h-ordering is called the *reduced* or *R-ordering* [2]. In this case, each element $\boldsymbol{t}$ in $\overline{\mathbb{R}}^P$ is reduced to a single value $h(\boldsymbol{t})$ by means of some combination of its components. In contrast to M-ordering, the aim of R-ordering is to produce a *total ordering*, i.e., for every pair $(\boldsymbol{s}, \boldsymbol{t})$ in $\overline{\mathbb{R}}^P \times \overline{\mathbb{R}}^P$ we have $\boldsymbol{s} \leq_h \boldsymbol{t}$ or $\boldsymbol{t} \leq_h \boldsymbol{s}$ [3].

Two more interesting types of ordering in $\overline{\mathbb{R}}^P$ have been proposed in [2]. The first type is called *partial ordering* or *P-ordering*, whereas the second type is called *conditional (sequential) ordering* or *C-ordering*. In the case of C-ordering, the elements $\boldsymbol{t}$ of $\overline{\mathbb{R}}^P$ are ordered by means of some marginal components of $\boldsymbol{t}$ conditioned on the selection, or ordering, of other marginal components of $\boldsymbol{t}$. We can usually formulate a desired type of C-ordering analytically, but the resulting formulation is quite complicated and depends on the particular type of C-ordering under consideration. We will not expand on this subject any further. P-ordering is rather geometrical in nature and is of no special interest here. For additional discussion on these vector ordering principles see also [11].

As noted in [2], M-ordering may serve as a prelude to some further partial ordering principle. Additionally, M-ordering may be applied to transformations of the available data set, instead. For example, in general $h(\boldsymbol{t}) = \{h_k(\boldsymbol{t}) \mid k \in K\}$, in which case the partial ordering $\leq_h$ on $\overline{\mathbb{R}}^P$ amounts to a data transformation from $\overline{\mathbb{R}}^P$ onto $\overline{\mathbb{R}}^K$ via h, followed by an M-ordering in $\overline{\mathbb{R}}^K$. Indeed, $\boldsymbol{s} \leq_h \boldsymbol{t}$ is equivalent to $h(\boldsymbol{s}) \leq h(\boldsymbol{t})$, which in turn is equivalent to $h_k(\boldsymbol{s}) \leq h_k(\boldsymbol{t})$, for every $k \in K$. In the rest of the paper we shall illustrate the potential of these concepts in terms of some interesting image processing and analysis examples.

3. Morphological operators for image sequences

In this section we give, in some detail, general expressions for adjunctions and h-adjunctions on $\mathrm{Fun}(E, \mathcal{T})^P$ and $\mathrm{Fun}(E, \mathcal{T}^P)$, respectively.

3.1. Adjunctions on $\mathrm{Fun}(E, \mathcal{T})^P$

In this subsection we assume that $\mathcal{T}$ is a complete lattice. We present a general expression for adjunctions on $\mathrm{Fun}(E, \mathcal{T})^P$ in terms of adjunctions on $\mathcal{T}$. Thereto we apply Proposition 2.2 twice. Using (2.4), we get that $(\mathcal{E}, \Delta)$ is an adjunction on $\mathrm{Fun}(E, \mathcal{T})^P$ iff there exist adjunctions $(\mathcal{E}_{p,q}, \Delta_{q,p})$ on $\mathrm{Fun}(E, \mathcal{T})$, for all $p, q \in P$, such that

$$(\mathcal{E}(\boldsymbol{F}))_p = \bigwedge_{q \in P} \mathcal{E}_{q,p}(F_q),$$

$$(\Delta(\boldsymbol{F}))_p = \bigvee_{q \in P} \Delta_{q,p}(F_q),$$

for $\boldsymbol{F} \in \mathrm{Fun}(E, \mathcal{T})^P$.

Now we use that $\text{Fun}(E,\mathcal{T}) = \mathcal{T}^E$. Applying Proposition 2.2 once more we find that $(\mathcal{E}_{p,q},\Delta_{q,p})$ is an adjunction on $\text{Fun}(E,\mathcal{T})$ iff there exist adjunctions $(\tilde{e}_{p,q,x,y},\tilde{d}_{q,p,y,x})$ on $\mathcal{T}$, for all $x,y \in E$, such that

$$\mathcal{E}_{p,q}(F)(x) = \bigwedge_{y\in E} \tilde{e}_{p,q,y,x}(F(y)),$$

$$\Delta_{q,p}(F)(x) = \bigvee_{y\in E} \tilde{d}_{q,p,y,x}(F(y)),$$

for $F \in \text{Fun}(E,\mathcal{T})$. Combining these two facts we arrive at the following result.

3.1. Proposition. *$(\mathcal{E},\Delta)$ is an adjunction on* $\text{Fun}(E,\mathcal{T})^P$ *if and only if there exist adjunctions $(\tilde{e}_{p,q,x,y},\tilde{d}_{q,p,y,x})$ on $\mathcal{T}$, for all $x,y \in E, p,q \in P$, such that*

$$(\mathcal{E}(\boldsymbol{F}))_p(x) = \bigwedge_{q\in P}\bigwedge_{y\in E} \tilde{e}_{q,p,y,x}(F_q(y)), \tag{3.1}$$

$$(\Delta(\boldsymbol{F}))_p(x) = \bigvee_{q\in P}\bigvee_{y\in E} \tilde{d}_{q,p,y,x}(F_q(y)), \tag{3.2}$$

for $\boldsymbol{F} \in \text{Fun}(E,\mathcal{T})^P$.

From Proposition 3.1 we see that, in the case of transformations on $\text{Fun}(E,\mathcal{T})^P$, the erosion of an image sequence $\boldsymbol{F}$ can be obtained by first eroding each single frame F_p of $\boldsymbol{F}$. The resulting erosion at frame q is the infimum of erosions on F_p, $p \in P$. This is illustrated in Figure 1 for the special case where $P = \{1,2\}$. For dilations we have found an analogous result.

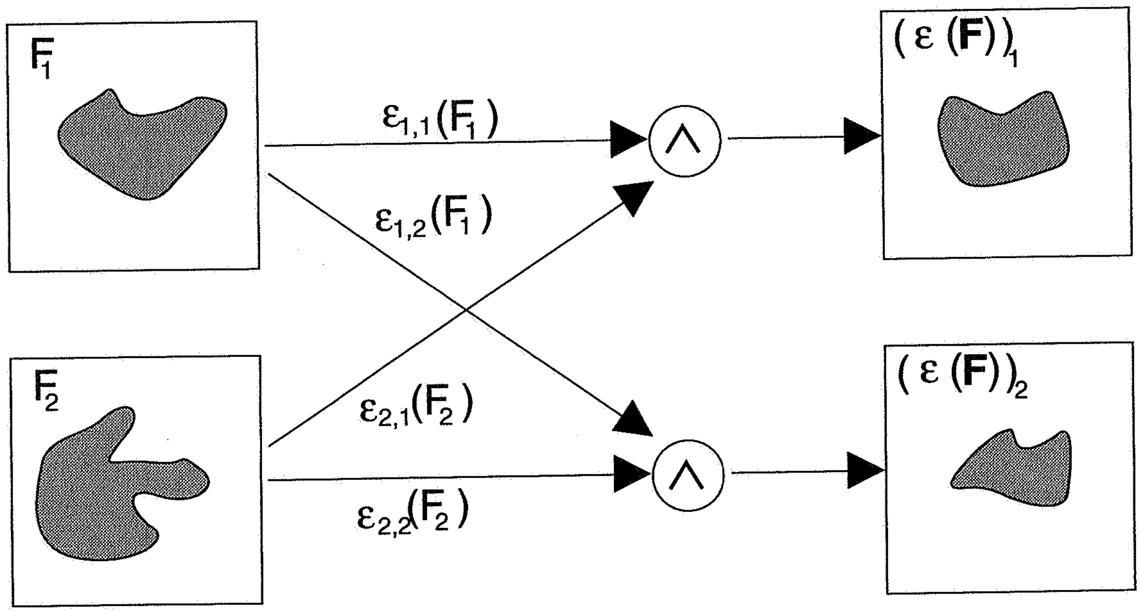

Fig. 1. Erosion of an image sequence $\boldsymbol{F}$ can be obtained by first eroding each single frame F_p of $\boldsymbol{F}$, and by then combining the results by means of infimum; $P = \{1,2\}$ in this example

In gray-scale morphology one can distinguish two important classes of morphological operators, the *H-operators* and the *T-operators* [12, 13]. We will extend these definitions to the framework of image sequences considered here. To introduce H-operators we have to assume that E has an abelian group structure, which we denote by '+'. We define the translation $\boldsymbol{F}_z$ of a function $\boldsymbol{F} \in \mathrm{Fun}(E, \mathcal{T})^P$ by an element $z \in E$ by $\boldsymbol{F}_z(x) = \boldsymbol{F}(x - z)$. An operator Ψ is called an H-operator if

$$\Psi(\boldsymbol{F}_z) = [\Psi(\boldsymbol{F})]_z, \quad \boldsymbol{F} \in \mathrm{Fun}(E, \mathcal{T})^P, \ z \in E.$$

To define a T-operator, we assume in addition that $\mathcal{T} = \overline{\mathbb{R}}$ (though the same definitions apply to $\mathcal{T} = \overline{\mathbb{Z}}$). If $\boldsymbol{F} \in \mathrm{Fun}(E, \overline{\mathbb{R}})^P$ and $\boldsymbol{t} \in \mathbb{R}^P$ we define

$$(\boldsymbol{F} + \boldsymbol{t})(x) = \boldsymbol{F}(x) + \boldsymbol{t}, \ x \in E.$$

We use the convention that $+\infty + t = +\infty$ and $-\infty + t = -\infty$, for $t \in \mathbb{R}$. An operator $\Psi : \mathrm{Fun}(E, \overline{\mathbb{R}})^P \to \mathrm{Fun}(E, \overline{\mathbb{R}})^P$ is called a T-operator if it is an H-operator, and if

$$\Psi(\boldsymbol{F} + \boldsymbol{t}) = \Psi(\boldsymbol{F}) + \boldsymbol{t},$$

for $\boldsymbol{F} \in \mathrm{Fun}(E, \overline{\mathbb{R}})^P$ and $\boldsymbol{t} \in \mathbb{R}^P$.

In the same way as in [13] one can show that the operators $\mathcal{E}, \Delta$ given by (3.1)–(3.2) are H-operators if and only if

$$\tilde{e}_{p,q,x,y} = \tilde{e}_{p,q,x+z,y+z}, \ \tilde{d}_{q,p,y,x} = \tilde{d}_{q,p,y+z,x+z},$$

for every $z \in E$. Writing $\tilde{e}_{p,q,z,0} = e_{p,q,z}$ and $\tilde{d}_{q,p,-z,0} = d_{q,p,z}$ the expressions in (3.1)–(3.2) reduce to

$$(\mathcal{E}(\boldsymbol{F}))_p(x) = \bigwedge_{q \in P} \bigwedge_{z \in E} e_{q,p,z}(F_q(x + z)),$$

$$(\Delta(\boldsymbol{F}))_p(x) = \bigvee_{q \in P} \bigvee_{z \in E} d_{q,p,z}(F_q(x - z)).$$

Every H-adjunction on $\mathrm{Fun}(E, \mathcal{T})^P$ is of this form.

Assume in addition that $\mathcal{T} = \overline{\mathbb{R}}$. In order that the adjunction given previously is a T-adjunction we must have that the pair $(e_{p,q,z}, d_{q,p,z})$ is T-invariant, i.e.,

$$e_{p,q,z}(s + t) = e_{p,q,z}(s) + t \ \text{ and } \ d_{q,p,z}(s + t) = d_{q,p,z}(s) + t,$$

for $s \in \overline{\mathbb{R}}, t \in \mathbb{R}$. Putting

$$B_{q,p}(z) = d_{q,p,z}(0),$$

we find

$$e_{p,q,z}(t) = t - B_{q,p}(z) \ \text{ and } \ d_{q,p,z}(t) = t + B_{q,p}(z). \tag{3.3}$$

This shows that every T-adjunction on $\mathrm{Fun}(E, \overline{\mathbb{R}})^P$ is given by

$$(\mathcal{E}(\boldsymbol{F}))_p(x) = \bigwedge_{q \in P} \bigwedge_{z \in E} [F_q(x + z) - B_{p,q}(z)], \tag{3.4}$$

$$(\Delta(\boldsymbol{F}))_p(x) = \bigvee_{q \in P} \bigvee_{z \in E} [F_q(x - z) + B_{q,p}(z)]. \tag{3.5}$$

Finally, it is worthwile noticing that the previous discussion carries over almost wordly to the more general case of adjunctions between $\mathrm{Fun}(E, \mathcal{T})^P$ and $\mathrm{Fun}(E, \mathcal{T})^K$, where K is a nonempty index set.

3.2. h-adjunctions on $\mathrm{Fun}(E, \mathcal{T}^P)$

The results stated in Proposition 3.1 are possible because $\mathcal{T}$ is assumed to be a complete lattice and there exists a factorization of the complete lattice $\mathrm{Fun}(E, \mathcal{T})^P$. In certain applications such a factorization may be undesirable; see Section 4. In this section we do not assume that $\mathcal{T}$ (hence $\mathcal{T}^P$) is a complete lattice, but rather that there exists a surjective mapping $h : \mathcal{T}^P \to \mathcal{R}^K$, where K is a non-empty set, and $\mathcal{R}$ (hence $\mathcal{R}^K$) is a complete lattice. The least and greatest elements in $\mathcal{R}^K$ are denoted by $-\infty$ and $+\infty$, respectively. In Subsection 2.3 we have explained that h induces a surjection from $\mathrm{Fun}(E, \mathcal{T}^P)$ to $\mathrm{Fun}(E, \mathcal{R}^K)$. This means in particular that all results about h-operators derived in Subsection 2.2 apply in the present situation. The next result gives a complete characterization of h-adjunctions on $\mathrm{Fun}(E, \mathcal{T}^P)$ in terms of h-adjunctions on $\mathcal{T}^P$ and in terms of adjunctions on $\mathcal{R}^K$.

3.2. Proposition.

(a) *The pair $(\mathcal{E}, \Delta)$ is an h-adjunction on $\mathrm{Fun}(E, \mathcal{T}^P)$ if and only if for every $x, y \in E$ there exists an h-adjunction $(e_{x,y}, d_{y,x})$ on $\mathcal{T}^P$ such that*

$$h(\mathcal{E}(\boldsymbol{F})(x)) = \bigwedge_{y \in E} h(e_{y,x}(\boldsymbol{F}(y))), \tag{3.6}$$

$$h(\Delta(\boldsymbol{F})(x)) = \bigvee_{y \in E} h(d_{y,x}(\boldsymbol{F}(y))), \tag{3.7}$$

for $\boldsymbol{F} \in \mathrm{Fun}(E, \mathcal{T}^P)$.

(b) *Let $e_{x,y} \overset{h}{\mapsto} e^{\circ}_{x,y}$ and $d_{y,x} \overset{h}{\mapsto} d^{\circ}_{y,x}$, then $(e^{\circ}_{x,y}, d^{\circ}_{y,x})$ is an adjunction on $\mathcal{R}^K$. Let $(\mathcal{E}^{\circ}, \Delta^{\circ})$ be the adjunction on $\mathrm{Fun}(E, \mathcal{R}^K)$ given by*

$$\mathcal{E}^{\circ}(\boldsymbol{F})(x) = \bigwedge_{y \in E} e^{\circ}_{y,x}(\boldsymbol{F}(y)),$$

$$\Delta^{\circ}(\boldsymbol{F})(x) = \bigvee_{y \in E} d^{\circ}_{y,x}(\boldsymbol{F}(y)),$$

for $\boldsymbol{F} \in \mathrm{Fun}(E, \mathcal{R}^K)$. Instead of (3.6)–(3.7) we can write

$$h(\mathcal{E}(\boldsymbol{F})(x)) = \mathcal{E}^{\circ}(h(\boldsymbol{F}))(x) = \bigwedge_{y \in E} e^{\circ}_{y,x}(h(\boldsymbol{F}(y))), \tag{3.8}$$

$$h(\Delta(\boldsymbol{F})(x)) = \Delta^{\circ}(h(\boldsymbol{F}))(x) = \bigvee_{y \in E} d^{\circ}_{y,x}(h(\boldsymbol{F}(y))), \tag{3.9}$$

for $\boldsymbol{F} \in \mathrm{Fun}(E, \mathcal{T}^P)$. In other words, $\mathcal{E} \overset{h}{\mapsto} \mathcal{E}^{\circ}$ and $\Delta \overset{h}{\mapsto} \Delta^{\circ}$.

The proof of this result follows by a straightforward combination of Proposition 2.2 and Proposition 2.8.

3.3. Remarks.

(a) Note that, in general, $\mathcal{E}(\boldsymbol{F})$ and $\Delta(\boldsymbol{F})$ are not uniquely determined by Proposition 3.2. For example, if the function $\boldsymbol{G}$ satisfies the equation in (3.6), i.e., if $h(\boldsymbol{G}(x)) = \bigwedge_{y \in E} h(e_{y,x}(\boldsymbol{F}(y)))$, then the same holds for $h^{\leftarrow} h(\boldsymbol{G})$, for every semi-inverse $h^{\leftarrow}$.

(b) The adjunctions $(e^{\circ}_{x,y}, d^{\circ}_{y,x})$ on $\mathcal{R}^K$ can be decomposed in terms of adjunctions on $\mathcal{R}$; see Proposition 2.2.

From now on we will restrict ourselves to the H-invariant case. We get

$$h(\mathcal{E}(\boldsymbol{F})(x)) = \bigwedge_{z\in E} h(e_z(\boldsymbol{F}(x+z))) = \bigwedge_{z\in E} e_z^\circ(h(\boldsymbol{F}(x+z))), \tag{3.10}$$

$$h(\Delta(\boldsymbol{F})(x)) = \bigvee_{z\in E} h(d_z(\boldsymbol{F}(x-z))) = \bigvee_{z\in E} d_z^\circ(h(\boldsymbol{F}(x-z))), \tag{3.11}$$

where $e_z = e_{z,0}, d_z = d_{-z,0}, e_z \overset{h}{\mapsto} e_z^\circ$, and $d_z \overset{h}{\mapsto} d_z^\circ$. Notice that (e_z, d_z) and (e_z°, d_z°) are h-adjunctions on $\mathcal{T}^P$ and adjunctions on $\mathcal{R}^K$, respectively. Assume further that $\mathcal{R} = \overline{\mathbb{R}}$. The corresponding adjunctions on $\mathrm{Fun}(E, \overline{\mathbb{R}}^K)$ are given by

$$\mathcal{E}^\circ(\boldsymbol{F})(x) = \bigwedge_{z\in E} e_z^\circ(\boldsymbol{F}(x+z)),$$

$$\Delta^\circ(\boldsymbol{F})(x) = \bigvee_{z\in E} d_z^\circ(\boldsymbol{F}(x-z)).$$

Let $e_{k,l,z}^\circ, d_{l,k,z}^\circ$, where $k, l \in K$, denote the decompositions of e_z°, d_z° in terms of mappings on $\overline{\mathbb{R}}$. We may write

$$(\mathcal{E}^\circ(\boldsymbol{F}))_k(x) = \bigwedge_{l\in K} \bigwedge_{z\in E} e_{l,k,z}^\circ(F_l(x+z)),$$

$$(\Delta^\circ(\boldsymbol{F}))_k(x) = \bigvee_{l\in K} \bigvee_{z\in E} d_{l,k,z}^\circ(F_l(x-z)).$$

In the T-invariant case we have

$$e_{k,l,z}^\circ(r) = r - A_{l,k}^\circ(z) \quad \text{and} \quad d_{l,k,z}^\circ(r) = r + A_{l,k}^\circ(z),$$

(cf. (3.3)) where $A_{l,k}^\circ(\cdot) : E \to \overline{\mathbb{R}}$ is given by $A_{l,k}^\circ(z) = d_{l,k,z}^\circ(0)$. Note that, when $A_{l,k}^\circ(z) = -\infty$, then $(e_{k,l,z}^\circ, d_{l,k,z}^\circ)$ is the trivial adjunction which is identically $(+\infty, -\infty)$. We consider the special case where

$$A_{l,k}^\circ(z) = -\infty \ \text{ if } \ l \neq k, \ z \in E.$$

Define $\boldsymbol{B}^\circ : E \to \overline{\mathbb{R}}^K$ by

$$B_k^\circ(z) := A_{k,k}^\circ(z), \ k \in K, \ z \in E.$$

The corresponding H-invariant h-adjunction is now given by

$$h(\mathcal{E}(\boldsymbol{F})(x)) = \bigwedge_{z\in E} \Big[h(\boldsymbol{F}(x+z)) - \boldsymbol{B}^\circ(z)\Big], \tag{3.12}$$

$$h(\Delta(\boldsymbol{F})(x)) = \bigvee_{z\in E} \Big[h(\boldsymbol{F}(x-z)) + \boldsymbol{B}^\circ(z)\Big]. \tag{3.13}$$

In practical situations, $\mathcal{T}$ will often be equal to $\overline{\mathbb{R}}$ (or a subset of it). In that case we can, for a given $\boldsymbol{B} : E \to \overline{\mathbb{R}}^P$, modify (3.12)-(3.13) in the following way:

$$h(\mathcal{E}(\boldsymbol{F})(x)) = \bigwedge_{z\in E} h(\boldsymbol{F}(x+z) - \boldsymbol{B}(z)),$$

$$h(\Delta(\boldsymbol{F})(x)) = \bigvee_{z\in E} h(\boldsymbol{F}(x-z) + \boldsymbol{B}(z)).$$

Under what conditions does $(\mathcal{E}, \Delta)$, which satisfies these expressions, define an h-adjunction? To answer this question we use the expressions in (3.10)–(3.11). Here (e_z, d_z) must be an h-adjunction on $\overline{\mathbb{R}}^P$. To arrive at the previous expressions we must take $e_z(\boldsymbol{t}) = \boldsymbol{t} - \boldsymbol{B}(z)$, $d_z(\boldsymbol{t}) = \boldsymbol{t} + \boldsymbol{B}(z)$. This defines an h-adjunction on $\overline{\mathbb{R}}^P$ if h satisfies the condition stated in the following proposition.

3.4. Proposition. *The pair* (e, d) *on* $\overline{\mathbb{R}}^P$ *given by* $e(\mathbf{t}) = \mathbf{t} - \mathbf{b}$, $d(\mathbf{t}) = \mathbf{t} + \mathbf{b}$, *where* $\mathbf{b} \in \overline{\mathbb{R}}^P$, *defines an h-adjunction if and only if*

$$h(\mathbf{s}) \leq h(\mathbf{t}) \iff h(\mathbf{s} + \mathbf{b}) \leq h(\mathbf{t} + \mathbf{b}), \tag{3.14}$$

for $\mathbf{s}, \mathbf{t} \in \overline{\mathbb{R}}^P$.

Refer to Example 2.10(b) for an example.

3.5. Proposition. *Let* $\mathbf{B} : E \to \overline{\mathbb{R}}^P$ *be a function such that* h *satisfies condition* (3.14) *for every* $\mathbf{b} \in \{\mathbf{B}(z) \mid z \in E\}$. *Then the pair* $(\mathcal{E}, \Delta)$ *of operators on* $\mathrm{Fun}(E, \overline{\mathbb{R}}^P)$ *determined by the expressions*

$$h(\mathcal{E}(\mathbf{F})(x)) = \bigwedge_{z \in E} h(\mathbf{F}(x + z) - \mathbf{B}(z)),$$

$$h(\Delta(\mathbf{F})(x)) = \bigvee_{z \in E} h(\mathbf{F}(x - z) + \mathbf{B}(z)),$$

defines an H-invariant h-adjunction on $\mathrm{Fun}(E, \overline{\mathbb{R}}^P)$.

We emphasize that, in general, h cannot be omitted in both expressions, and therefore $\mathcal{E}, \Delta$ are not T-invariant in general. Furthermore, condition (3.14) will only be satisfied for $\mathbf{b}$ in a subset of $\overline{\mathbb{R}}^P$. We present some examples.

3.6. Examples. We assume that P is a finite set. With an abuse of notation, we replace P by $\{1, 2, \ldots, P\}$, where $P \geq 1$.
(a) Let $h : \overline{\mathbb{R}}^P \to \overline{\mathbb{R}}$ be one of the mappings given by

$$h_{\mathrm{inf}}(t_1, t_2, \ldots, t_P) = \bigwedge_{p=1}^{P} t_p\,, \quad h_{\mathrm{sup}}(t_1, t_2, \ldots, t_P) = \bigvee_{p=1}^{P} t_p.$$

Then (3.14) is satisfied for $\mathbf{b} \in \overline{\mathbb{R}}^P$ with $b_p = b$ for all $p = 1, 2, \ldots, P$, where $b \in \overline{\mathbb{R}}$.
(b) We can extend the previous example to arbitrary order statistics. Let $(t^{(1)}, t^{(2)}, \ldots, t^{(P)})$ be the ordered values of $(t_1, t_2, \ldots, t_P)$ such that $t^{(1)} \geq t^{(2)} \geq \ldots \geq t^{(P)}$. Let $h : \overline{\mathbb{R}}^P \to \overline{\mathbb{R}}$ be given by $h(t_1, t_2, \ldots, t_P) = t^{(n)}$, where $1 \leq n \leq P$. Then (3.14) is satisfied for $\mathbf{b} \in \overline{\mathbb{R}}^P$ with $b_p = b$ for all $p = 1, 2, \ldots, P$, where $b \in \overline{\mathbb{R}}$.

We may also take $h(t_1, t_2, \ldots, t_P) = \sum_{p=1}^{P} w_p t^{(p)}$, where w_p are given weights. Note that $h = h_{\mathrm{sup}}$ if $w_1 = 1$ and $w_p = 0$ for $p > 1$. Dually, $h = h_{\mathrm{inf}}$ if $w_P = 1$ and $w_p = 0$ for $p < P$.
(c) If we take for $h : \overline{\mathbb{R}}^P \to \overline{\mathbb{R}}$ the linear transformation

$$h(t_1, t_2, \ldots, t_P) = \sum_{p=1}^{P} w_p t_p,$$

where w_p are given weights, then (3.14) is satisfied for any $\mathbf{b} \in \overline{\mathbb{R}}^P$. We point out that in this example one has to be careful with expressions which contain $+\infty$ or $-\infty$. It is evident how one must generalize the present example to linear transformations $h : \overline{\mathbb{R}}^P \to \overline{\mathbb{R}}^K$.

A class of H-invariant h-adjunctions which is of special interest are the *flat h-adjunctions.* This class is obtained by making the following special choices for (e_z°, d_z°) in (3.10)–(3.11). Let $A \subseteq E$ be a given structuring element, and define $e_z^\circ(\mathbf{t}) = d_z^\circ(\mathbf{t}) = \mathbf{t}$ for $\mathbf{t} \in \overline{\mathbb{R}}^K$ and $z \in A$. For

z outside A we define $e_z^\circ(\boldsymbol{t}) = +\infty$ and $d_z^\circ(\boldsymbol{t}) = -\infty$ for all $\boldsymbol{t} \in \overline{\mathbb{R}}^K$. The corresponding h-adjunction $(\mathcal{E}, \Delta)$ is given by

$$h(\mathcal{E}(\boldsymbol{F})(x)) = \bigwedge_{z \in A} h(\boldsymbol{F}(x+z)), \tag{3.15}$$

$$h(\Delta(\boldsymbol{F})(x)) = \bigvee_{z \in A} h(\boldsymbol{F}(x-z)). \tag{3.16}$$

3.7. Algorithm. In general, $\mathcal{E}(\boldsymbol{F})(x)$ is not uniquely determined by (3.6). Dually, $\Delta(\boldsymbol{F})(x)$ is not uniquely determined by (3.7). We indicate a special, but practically useful, way to choose $\mathcal{E}(\boldsymbol{F})(x)$ and $\Delta(\boldsymbol{F})(x)$ from the various possibilities. Instead of (3.6)-(3.7) we consider (3.15)-(3.16), and we assume that $\overline{\mathbb{R}}^K$ is totally ordered (which holds automatically if $K = 1$) and that the structuring element A is finite (which is true in almost all cases of interest). Consider (3.15) for a given $\boldsymbol{F}$. Let, for $x \in E$, $z_e(x) \in A$ be such that $h(\boldsymbol{F}(x+z))$ assumes its minimum at $z = z_e(x)$; then

$$h(\mathcal{E}(\boldsymbol{F})(x)) = h\big(\boldsymbol{F}(x + z_e(x))\big),$$

and we may choose

$$\mathcal{E}(\boldsymbol{F})(x) = \boldsymbol{F}(x + z_e(x)).$$

For the dilation in (3.16), let $z = z_d(x)$ be such that

$$h(\Delta(\boldsymbol{F})(x)) = h\big(\boldsymbol{F}(x - z_d(x))\big),$$

and choose

$$\Delta(\boldsymbol{F})(x) = \boldsymbol{F}(x - z_d(x)).$$

In general, there will not be a unique choice for z_e and z_d.

It is obvious how to extend this algorithm for non-flat structuring functions.

3.3. h-increasing operators on $\mathrm{Fun}(E, \mathcal{T}^P)$

So far, we have mainly concentrated on (H-invariant) h-adjunctions on $\mathrm{Fun}(E, \mathcal{T}^P)$. We are now in a position to study more complicated transformations, like openings and closings. In [13, Thm. 11.23], Matheron's theorem has been extended to H-operators for gray-scale functions. The proof given there carries over almost literally to the present framework. Let $\boldsymbol{O}, \boldsymbol{I}$ denote the functions which are identically $-\infty$ and $+\infty$, respectively.

3.8. Proposition.

(a) *Every increasing H-operator Ψ on $\mathrm{Fun}(E, \mathcal{R}^K)$ with $\Psi(\boldsymbol{O}) = \boldsymbol{O}$ can be represented as an infimum of H-dilations on $\mathrm{Fun}(E, \mathcal{R}^K)$.*

(b) *Every increasing H-operator Ψ on $\mathrm{Fun}(E, \mathcal{R}^K)$ with $\Psi(\boldsymbol{I}) = \boldsymbol{I}$ can be represented as a supremum of H-erosions on $\mathrm{Fun}(E, \mathcal{R}^K)$.*

Recall from Subsection 2.2 that $\mathcal{T}^P[\boldsymbol{r}] = \{\boldsymbol{t} \in \mathcal{T}^P \mid h(\boldsymbol{t}) = \boldsymbol{r}\}$, for a given $\boldsymbol{r} \in \mathcal{R}^K$. For h-increasing operators, the condition $\Psi(\boldsymbol{O}) = \boldsymbol{O}$ translates into "the class of functions in $\mathrm{Fun}(E, \mathcal{T}^P)$ with values in $\mathcal{T}^P[-\infty]$ is invariant under Ψ". Dually, the condition $\Psi(\boldsymbol{I}) = \boldsymbol{I}$ translates into "the class of functions in $\mathrm{Fun}(E, \mathcal{T}^P)$ with values in $\mathcal{T}^P[+\infty]$ is invariant under Ψ"

3.9. Theorem.

(a) *Let Ψ be an h-increasing H-operator on* $\mathrm{Fun}(E,\mathcal{T}^P)$ *such that the class of functions in* $\mathrm{Fun}(E,\mathcal{T}^P)$ *with values in* $\mathcal{T}^P[-\infty]$ *is invariant under* Ψ*; then there exists a collection* Δ_i, $i \in I$, *of H-invariant h-dilations on* $\mathrm{Fun}(E,\mathcal{T}^P)$ *such that*

$$h\Psi = \bigwedge_{i\in I} h\Delta_i.$$

(b) *Let Ψ be an h-increasing H-operator on* $\mathrm{Fun}(E,\mathcal{T}^P)$ *such that the class of functions in* $\mathrm{Fun}(E,\mathcal{T}^P)$ *with values in* $\mathcal{T}^P[+\infty]$ *is invariant under* Ψ*; then there exists a collection* $\mathcal{E}_i$, $i \in I$, *of H-invariant h-erosions on* $\mathrm{Fun}(E,\mathcal{T}^P)$ *such that*

$$h\Psi = \bigvee_{i\in I} h\mathcal{E}_i.$$

PROOF. We prove (a); then (b) follows by duality. Assume that Ψ satisfies the assumptions in (a), and let $\Psi \overset{h}{\mapsto} \Psi^\circ$. Then Ψ° satisfies the assumptions of Proposition 3.8(a), and we have a decomposition $\Psi^\circ = \bigwedge_{i\in I}\Delta_i^\circ$, for some collection of H-invariant dilations Δ_i°, $i \in I$, on $\mathrm{Fun}(E,\mathcal{R}^K)$. This yields that $\Psi^\circ h = \bigwedge_{i\in I}\Delta_i^\circ h$. Let $i \in I$ be fixed. For every semi-inverse $h^{\leftarrow}$, the operator $\Delta_i = h^{\leftarrow}\Delta_i^\circ h$ is an H-invariant h-dilation on $\mathrm{Fun}(E,\mathcal{T}^P)$ with $\Delta_i^\circ h = h\Delta_i$. We get that $h\Psi = \Psi^\circ h = \bigwedge_{i\in I} h\Delta_i$. This proves the result. ∎

Recall that an operator α on $\mathrm{Fun}(E,\mathcal{T}^P)$ is an h-opening if
- α is h-increasing;
- $\alpha^2 =_h \alpha$ (h-idempotence);
- $\alpha \le_h \mathrm{id}$ (h-anti-extensivity).

h-closings are defined similarly.

From Proposition 2.11 we know that $\Delta\mathcal{E}$ and $\mathcal{E}\Delta$ are an h-opening and h-closing, respectively, when $(\mathcal{E},\Delta)$ is an h-adjunction. As stated in Proposition 2.12, composition of an h-opening α and an h-closing β yields h-filters, the so-called *alternating h-filters.*

4. Examples and applications

In this section, we shall discuss a number of interesting image processing and analysis examples which utilize the theory developed in the previous sections. As a direct result of our discussions, we shall immediately see that the proposed theory is general enough to be applied in a number of problems, including noise smoothing, color image processing, biomedical image processing, analysis of multi-spectral images, computation of the Euclidean distance transform, etc. Like in Example 3.6 we replace the index set P by the finite set $\{1,2,\ldots,P\}$.

4.1. Matrix morphology

The theory developed in Subsection 3.1 is directly related to Wilson's theory of *matrix morphology* [42, 43, 44, 45]; see also [13, §5.5.D]. In fact, our theory extends Wilson's theory to the case of general erosions, dilations, openings, and closings, and shows that Wilson's approach to matrix morphology is a direct consequence of vector M-ordering. Indeed, let us consider a $1 \times P$ row vector image

$$\boldsymbol{F} = [\,F_1 \quad F_2 \quad \cdots \quad F_P\,],$$

and a $P \times P$ matrix structuring function

$$\boldsymbol{B} = \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1P} \\ B_{21} & B_{22} & \cdots & B_{2P} \\ \vdots & \vdots & \ddots & \vdots \\ B_{P1} & B_{P2} & \cdots & B_{PP} \end{bmatrix}.$$

The matrix erosion of $\boldsymbol{F}$ by $\boldsymbol{B}$ is defined by

$$\boldsymbol{F} \boxminus \boldsymbol{B}^T = [\, F_1 \quad F_2 \quad \cdots \quad F_P \,] \boxminus \begin{bmatrix} B_{11} & B_{21} & \cdots & B_{P1} \\ B_{12} & B_{22} & \cdots & B_{P2} \\ \vdots & \vdots & \ddots & \vdots \\ B_{1P} & B_{2P} & \cdots & B_{PP} \end{bmatrix} = [\, E_1 \quad E_2 \quad \cdots \quad E_P \,],$$

where

$$E_p = \bigwedge_{q=1}^{P} F_q \ominus B_{pq}.$$

Similarly, the matrix dilation is defined by

$$\boldsymbol{F} \boxplus \boldsymbol{B} = [\, F_1 \quad F_2 \quad \cdots \quad F_P \,] \boxplus \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1P} \\ B_{21} & B_{22} & \cdots & B_{2P} \\ \vdots & \vdots & \ddots & \vdots \\ B_{P1} & B_{P2} & \cdots & B_{PP} \end{bmatrix} = [\, D_1 \quad D_2 \quad \cdots \quad D_P \,],$$

where

$$D_p = \bigvee_{q=1}^{P} F_q \oplus B_{qp}.$$

These expressions coincide with those given in (3.4)–(3.5). These results have been generalized in [44] by considering an $N \times P$ matrix image

$$\boldsymbol{F} = \begin{bmatrix} F_{11} & F_{12} & \cdots & F_{1P} \\ F_{21} & F_{22} & \cdots & F_{2P} \\ \vdots & \vdots & \ddots & \vdots \\ F_{N1} & F_{N2} & \cdots & F_{NP} \end{bmatrix}.$$

The matrix erosion of $\boldsymbol{F}$ by $\boldsymbol{B}$ is now given by

$$\boldsymbol{F} \boxminus \boldsymbol{B}^T = \begin{bmatrix} F_{11} & F_{12} & \cdots & F_{1P} \\ F_{21} & F_{22} & \cdots & F_{2P} \\ \vdots & \vdots & \ddots & \vdots \\ F_{N1} & F_{N2} & \cdots & F_{NP} \end{bmatrix} \boxminus \begin{bmatrix} B_{11} & B_{21} & \cdots & B_{P1} \\ B_{12} & B_{22} & \cdots & B_{P2} \\ \vdots & \vdots & \ddots & \vdots \\ B_{1P} & B_{2P} & \cdots & B_{PP} \end{bmatrix} = \begin{bmatrix} E_{11} & E_{12} & \cdots & E_{1P} \\ E_{21} & E_{22} & \cdots & E_{2P} \\ \vdots & \vdots & \ddots & \vdots \\ E_{N1} & E_{N2} & \cdots & E_{NP} \end{bmatrix},$$

where

$$E_{np} = \bigwedge_{q=1}^{P} F_{nq} \ominus B_{pq}, \; n = 1, \ldots, N.$$

The theory presented in Subsection 3.1 trivially extends to this case, by considering N different image sequences (each with P components), which are processed by means of the same $P \times P$ matrix structuring function $\boldsymbol{B}$.

Since a number of interesting examples, illustrating this theory, have been already presented in [44] and [45], we shall not provide any additional examples here. Rather, we shall focus our attention on discussing a number of applications associated with h-operators on $\mathrm{Fun}(E, \mathcal{T}^P)$.

4.2. The case that h is invertible

As we have previously discussed, we may decide upon the functional form of h and transform our image sequence from $\mathcal{T}^P$ into $\mathcal{R}^K$. We can now decide upon the proper set of adjunctions $(e^{\circ}_{x,y}, d^{\circ}_{y,x})$ on $\mathcal{R}^K$, and use (3.8)-(3.9), in order to define erosions and dilations on $\text{Fun}(E, \mathcal{T}^P)$, provided that h is invertible. The simplest invertible transformation that we may think of is the linear transformation:

$$h(\boldsymbol{t}) = \boldsymbol{A}\boldsymbol{t}, \tag{4.1}$$

where $\boldsymbol{A}$ is an invertible $P \times P$ matrix. When processing multi-spectral images, $\boldsymbol{A}$ might be the *Karhunen-Loeve transform matrix* [16, 28], the *Kauth-Thomas tasseled cap transform matrix* [28], or the *maximum noise fraction* (MNF) *transform* matrix [9, 17].

In the case of the MNF transform, a sequence of (possibly) mutually correlated and noisy multi-spectral images $\boldsymbol{F} = \boldsymbol{F}_0 + \boldsymbol{N} = (F_1, F_2, \ldots, F_P)$, consisting of P bands is transformed, by means of a linear invertible transformation, to a new image sequence $\boldsymbol{G} = (G_1, G_2, \ldots, G_P)$. If the corrupting additive noise sequence $\boldsymbol{N} = (N_1, N_2, \ldots, N_P)$ is uncorrelated from $\boldsymbol{F}_0$, and if $\boldsymbol{C}_{\boldsymbol{F}_0} = \text{Cov}(\boldsymbol{F}_0(x))$ and $\boldsymbol{C}_{\boldsymbol{N}} = \text{Cov}(\boldsymbol{N}(x))$, for every $x \in E$, are the (stationary) covariance matrices of $\boldsymbol{F}_0(x)$ and $\boldsymbol{N}(x)$, respectively, then

$$\boldsymbol{G}(x) = \boldsymbol{A}_{\text{mnf}}\boldsymbol{F}(x), \quad \text{for every} \quad x \in E,$$

with

$$\boldsymbol{A}^T_{\text{mnf}} = [\, \boldsymbol{e}_1 \quad \boldsymbol{e}_2 \quad \cdots \quad \boldsymbol{e}_P \,],$$

where $\boldsymbol{e}_p$, $p = 1, 2, \ldots, P$, are the left eigenvectors of matrix $\boldsymbol{C}_{\boldsymbol{N}}\boldsymbol{C}^{-1}_{\boldsymbol{F}_0}$, with corresponding eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_P$. It can be easily shown [9, 17] that $G_1(x), G_2(x), \ldots, G_P(x)$ are mutually uncorrelated and that the *signal-to-noise ratio* (SNR) in G_{p+1} will be greater than the SNR in G_p, for every $p = 1, 2, \ldots, P-1$. Since the components of $\boldsymbol{G}$ are now mutually uncorrelated, they can be processed independently (i.e., component-wise processing of $\boldsymbol{G}$ will be natural here).

If noise smoothing is desirable, we may decide to apply morphological filtering (e.g., by means of *alternating filtering* [34, 36]) on components $G_1, G_2, \ldots, G_{p'}$, for some $p' \leq P$, leaving components $G_{p'+1}, G_{p'+2}, \ldots, G_P$ unprocessed, since these components will be characterized by high SNR's. It is shown in [9] that the SNR in component G_p is proportional to $1/\lambda_p$; therefore, p' may be chosen such that $1/\lambda_p \geq S$, for every $p > p'$, whereas, $1/\lambda_p < S$, for every $p \leq p'$, where S is a predefined SNR threshold. This type of filtering will produce an image $\boldsymbol{G}'$ which, in turn, will result in a filtered version $\boldsymbol{F}'$ of the image $\boldsymbol{F}$, by means of $\boldsymbol{F}'(x) = \boldsymbol{A}^{-1}_{\text{mnf}}\boldsymbol{G}'(x)$, for every $x \in E$. It is now not difficult to show that the overall transformation $\boldsymbol{F} \mapsto \boldsymbol{F}'$ will be a morphological h-filter on $\text{Fun}(E, \mathcal{T}^P)$, where h is given by (4.1).

To illustrate this process, let us consider the three-band, 256×256 pixel, 256 gray-level image $\boldsymbol{F}_0$ depicted in the first row of Figure 2. This image is a proton density weighted, T1-weighted, and T2-weighted magnetic resonance vector image of the same anatomical site of a human brain. The second row of Figure 2 depicts the bands of image $\boldsymbol{F}$ obtained by corrupting the original bands with additive i.i.d. vector salt and pepper noise $\boldsymbol{N}$, with zero mean and covariance matrix given by

$$\boldsymbol{C}_{\boldsymbol{N}} = \begin{bmatrix} 1.0 & 0.8 & 0.1 \\ 0.8 & 1.0 & -0.5 \\ 0.1 & -0.5 & 1.0 \end{bmatrix}.$$

The image covariance matrix has been estimated to be

$$\boldsymbol{C}_{\boldsymbol{F}_0} = \begin{bmatrix} 0.3405 & 0.2510 & 0.5189 \\ 0.2510 & 0.4159 & 0.5814 \\ 0.5189 & 0.5814 & 1.0000 \end{bmatrix}.$$

In this case the MNF transform matrix is given by

$$\boldsymbol{A}_{\text{mnf}} = \begin{bmatrix} 0.5563 & 0.5551 & -0.6184 \\ 0.8366 & -0.5399 & 0.0931 \\ -0.5896 & 0.6910 & 0.4182 \end{bmatrix},$$

whereas, the eigenvalues of matrix $\boldsymbol{C_N C_{F_0}^{-1}}$ are given by $\lambda_1 = 1.1961 \times 10^2$, $\lambda_2 = 2.0946$, and $\lambda_3 = 2.6802 \times 10^{-2}$. The uncorrelated image $\boldsymbol{G}$ is depicted in the third row of Figure 2.

Observe that most of the noise energy is concentrated in the first band of $\boldsymbol{G}$, whereas the third band is virtually noise free. Image $\boldsymbol{G'}$, whose bands are depicted in the fourth row of Figure 2, is obtained by independently applying alternating filtering on the first and second bands of $\boldsymbol{G}$. The filtering is based on a flat structuring function B, with domain $\{(0,0),(0,1),(1,0),(1,1)\}$. Observe that no filtering is applied on G_3. The fifth row of Figure 2 depicts the resulting filtered image $\boldsymbol{F'}$. The quality of $\boldsymbol{F'}$ is remarkable. Most of the noise has been eliminated from all three bands, whereas, morphological filtering has been successful in preserving most image features of interest.

Another important application of using invertible transformations h is color image processing. Color images contain three primary colors: red (R), green (G), and blue (B). If morphological operators are directly applied on each of these three primary color images, the resulting image will not preserve its original color; see [24, 35]. One way to preserve color is to transform the RGB image sequence $\boldsymbol{F} = (F_r, F_g, F_b)$ into an HLS (Hue-Lightness-Saturation) image sequence $\boldsymbol{G} = (G_h, G_l, G_s)$, by means of a non-linear invertible transformation h, such that $\boldsymbol{G}(x) = h(\boldsymbol{F}(x))$, for every $x \in E$ [16, 24, 35]. We can, then, independently process all, or some, of the components in $\boldsymbol{G}$, in order to produce a morphologically processed image sequence $\boldsymbol{G'}$. For example, in [24], only the L-component G_l of $\boldsymbol{G}$ is processed, whereas, in [35], morphological processing of both G_h and G_s is suggested. Image sequence $\boldsymbol{G'}$ is now transformed back to the RGB space, in order to produce an image sequence $\boldsymbol{F'}$, such that $\boldsymbol{F'}(x) = h^{-1}(\boldsymbol{G'}(x))$, for every $x \in E$. Alternatively, we may transform the RGB image sequence $\boldsymbol{F}$ into a chrominance-luminance sequence $\boldsymbol{G}$, by means of an invertible linear transformation, morphologically process the luminance components in $\boldsymbol{G}$, independently from its chrominance components, in order to obtain $\boldsymbol{G'}$, and transform back to the RGB space, in order to obtain $\boldsymbol{F'}$. This, for example, may be achieved by means of the RGB to NTSC transmission system transformation, given by [16]

$$\begin{bmatrix} G_y \\ G_i \\ G_q \end{bmatrix} = \begin{bmatrix} 0.177 & 0.814 & 0.011 \\ 0.534 & -0.247 & -0.179 \\ 0.247 & -0.679 & 0.405 \end{bmatrix} \begin{bmatrix} F_r \\ F_g \\ F_b \end{bmatrix},$$

where G_y is the luminance component, and G_i and G_q are two chrominance components in $\boldsymbol{G}$. Again, the overall transformation $\boldsymbol{F} \mapsto \boldsymbol{F'}$ will be a morphological operator on $\text{Fun}(E, \mathcal{T}^P)$.

Finally, it has been suggested by Serra [35] that an image sequence should be morphologically processed by introducing priorities in processing its individual components. This idea leads to the so-called *conditional lattices*, which involve functional relationships between their components. When linear relationships are employed, it has been suggested in [35] that the partial order $\preceq$ relationship associated with a conditional lattice should be given by (for the case when $P = 3$)

$$\boldsymbol{s} \preceq \boldsymbol{t} \Longleftrightarrow \begin{cases} s_1 \leq t_1 \\ s_2 \leq t_2 + a_1(t_1 - s_1) \\ s_3 \leq t_3 + a_3(t_1 - s_1) + a_2(t_2 - s_2), \end{cases}$$

where a_1, a_2, and a_3 are constants. Instead of $\boldsymbol{s} \preceq \boldsymbol{t}$ we can write $\boldsymbol{s} \leq_h \boldsymbol{t}$, where h is given by (4.1), with

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 \\ a_1 & 1 & 0 \\ a_3 & a_2 & 1 \end{bmatrix}.$$

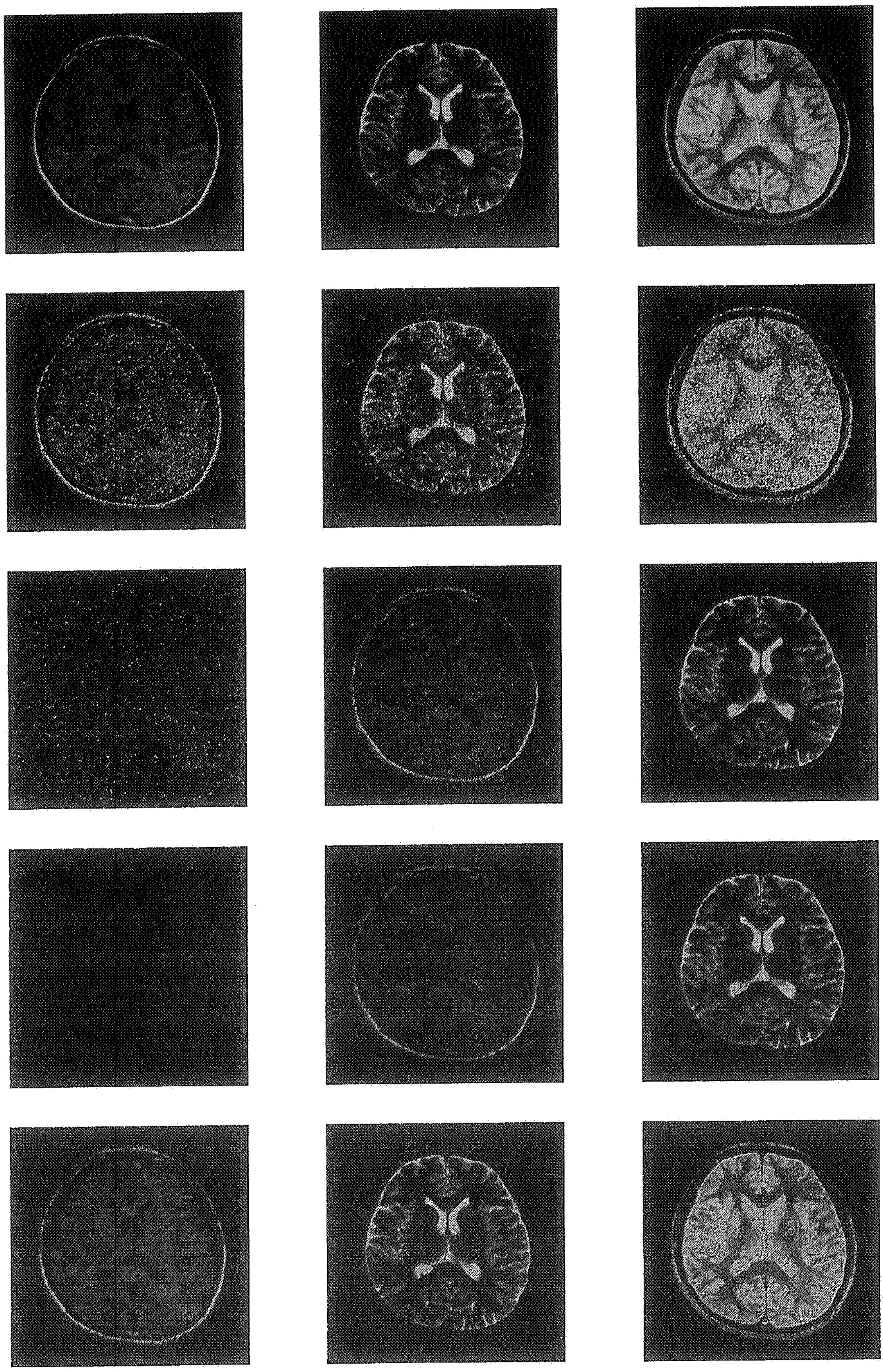

Fig. 2. An example of vector morphological filtering by means of an invertible linear transformation.

4.3. Operators based on Mahalanobis distance

We shall now discuss a simple example where h is a scalar transformation from $\overline{\mathbb{R}}^P$ into $\overline{\mathbb{R}}$. A more complicated example follows. In both cases, we shall consider erosions and dilations obtained from Algorithm 3.7.

Let us first consider the image sequence $\boldsymbol{F}$, employed in the MNF example of the previous sub-section. It has been suggested in [11] and [25] that if filtering of $\boldsymbol{F}$ is of interest, then the vector data under consideration may be ordered by means of the *Mahalanobis distance*

$$h(\boldsymbol{t}) = (\boldsymbol{t} - \boldsymbol{\tau}_0)^T C_{\boldsymbol{F}_0}^{-1} (\boldsymbol{t} - \boldsymbol{\tau}_0), \tag{4.2}$$

where $\boldsymbol{\tau}_0 = E(\boldsymbol{F}_0(x))$, for every $x \in E$, is the (stationary) mean vector of $\boldsymbol{F}_0(x)$. The Mahalanobis distance has been employed in [11] and [25] for the development of rank-order filters applied on image sequences. This vector-to-scalar transformation can be also used in the morphological framework of this paper, provided that only flat structuring elements are used (since, in this case, (3.14) is satisfied). Figures 3 and 4 depict an example of applying Algorithm 3.7, where h is given by (4.2), on $\boldsymbol{F}$.

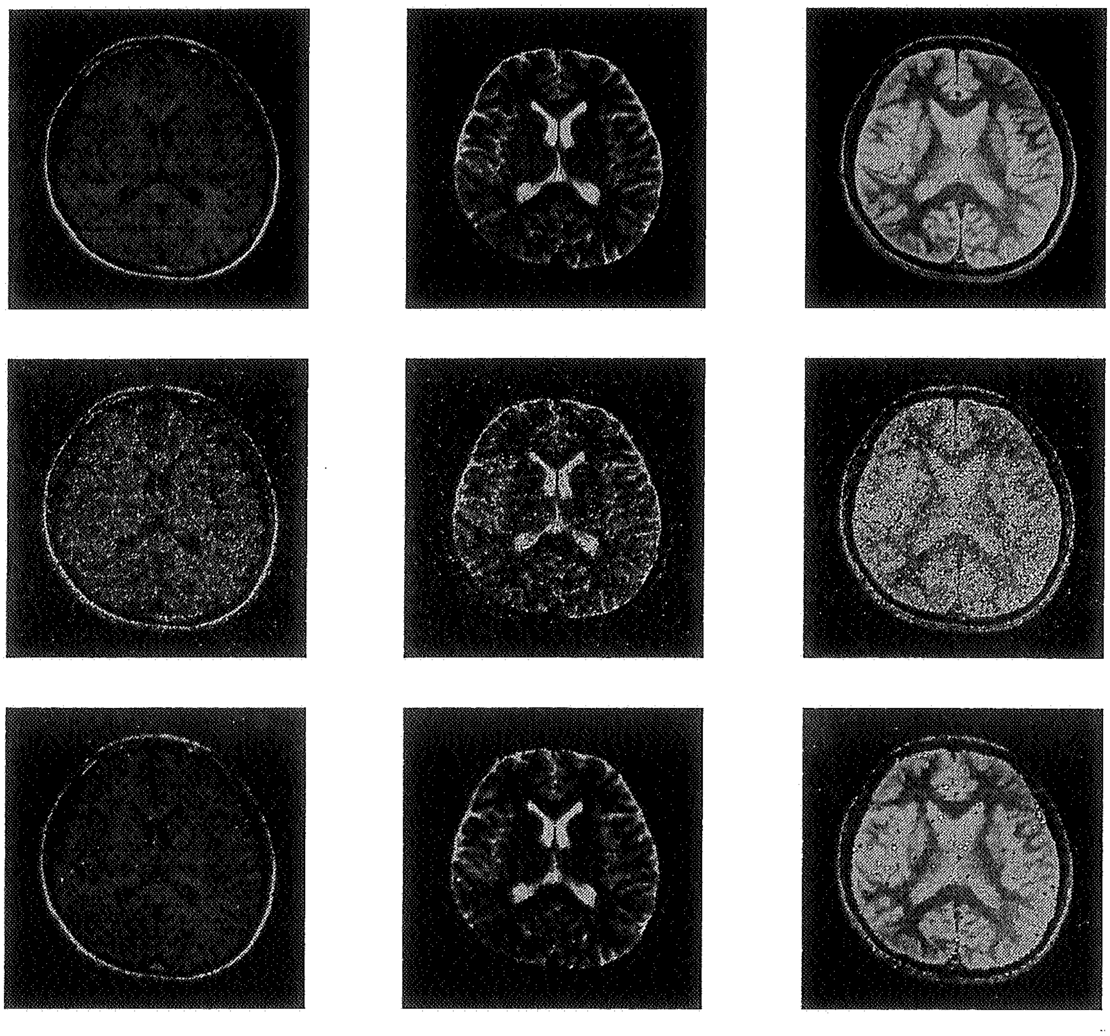

Fig. 3. An example of vector morphological filtering by means of a scalar Mahalanobis distance transformation.

The mean vector τ_0 has been estimated to be

$$\tau_0 = \begin{bmatrix} 0.5823 \\ 0.5745 \\ 0.9913 \end{bmatrix}.$$

The first row of Figure 3 depicts the three bands of the noise-free image $\boldsymbol{F}_0$, whereas, the second row depicts the three bands of the noisy image $\boldsymbol{F}$. The filtered image $\boldsymbol{F}'$, obtained by applying alternating filtering on $\boldsymbol{F}$ is depicted in the third row of Figure 3. The same flat structuring function, as the one used in Figure 2, has been employed here. The resulting images are very similar to the ones depicted in the fifth row of Figure 2. However, the MNF transform based procedure has resulted in slightly better noise smoothing performance. Finally, Figure 4 depicts the Mahalanobis distance images, obtained by applying transformation (4.2) on $\boldsymbol{F}$, $\boldsymbol{F}_0$, and $\boldsymbol{F}'$.

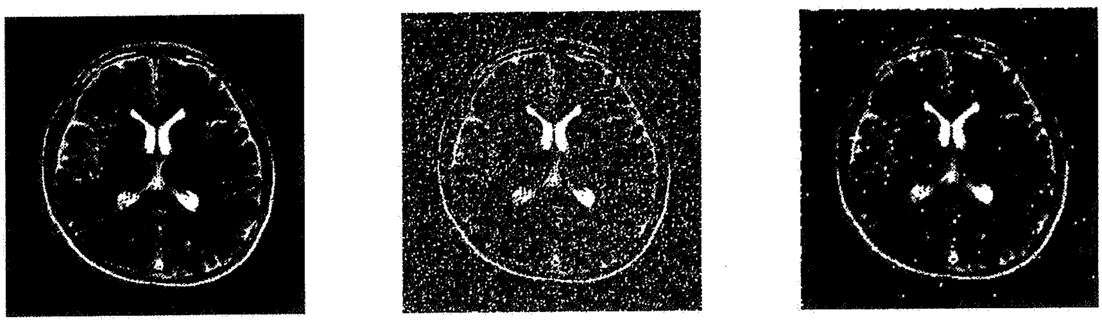

Fig. 4. The Mahalanobis distance images of the (a) noise-free image $\boldsymbol{F}_0$, (b) noisy image $\boldsymbol{F}$, and (c) filtered image $\boldsymbol{F}'$ depicted in Figure 3.

These images verify the fact that a vector morphological filtering approach, based on (4.2), is able to remove vector image outliers, due to corrupting noise, by first removing outliers in the distance domain and by then mapping the results back to the vector domain.

4.4. Vector distance transform

Finally, we shall apply our theory to the problem of calculating distance transforms. In many image processing and analysis applications, it is necessary to devise a procedure which calculates the *distance* between a point $\boldsymbol{u}$ and a set X in the P-dimensional Euclidean vector space $\mathbb{R}^P$, given by

$$\Delta_X(\boldsymbol{u}) = \begin{cases} \bigwedge_{\mathbf{x}\in X} \|\mathbf{x} - \mathbf{u}\|, & \boldsymbol{u} \notin X, \\ 0, & \boldsymbol{u} \in X, \end{cases} \tag{4.3}$$

where $\|\cdot\|$ is a norm on $\mathbb{R}^P$. The function $\Delta_X(\cdot)$ is known as the *distance transform* of set X, and has been extensively used for skeletonization, segmentation, smoothing, matching, etc., of objects in $\mathbb{R}^P$ [18, 21, 22, 23, 26, 27, 31, 33, 41].

Quite frequently, computation of $\Delta_X(\boldsymbol{u})$ may not be sufficient; rather, knowledge of a vector $\boldsymbol{d}_X(\boldsymbol{u})$, which satisfies

$$\boldsymbol{d}_X(\boldsymbol{u}) = \begin{cases} \boldsymbol{p}_X(\boldsymbol{u}) - \boldsymbol{u}, & \boldsymbol{u} \notin X \\ \boldsymbol{0}, & \boldsymbol{u} \in X, \end{cases} \tag{4.4}$$

where $\boldsymbol{p}_X(\boldsymbol{u})$ is a vector in X such that

$$\|\boldsymbol{p}_X(\boldsymbol{u}) - \mathbf{u}\| = \bigwedge_{\mathbf{x}\in X} \|\mathbf{x} - \mathbf{u}\|, \ \boldsymbol{u} \notin X, \tag{4.5}$$

may be desirable. This is the case with many signal and image processing techniques and algorithms, such as template matching [16], logarithmic, sequential, and hierarchical search techniques for scene matching and detection [16], unsupervised learning [16, 32], predictive coding with motion compensation [16], and learning vector quantization [19], to mention a few (see also [27]). We call $\boldsymbol{p}_X(\boldsymbol{u})$ the *projection* of $\boldsymbol{u}$ on X and $\boldsymbol{d}_X(\boldsymbol{u})$ the *vector distance transform* of the set X.

Observe that

$$\Delta_X(\boldsymbol{u}) = \|\boldsymbol{d}_X(\boldsymbol{u})\|, \;\; \boldsymbol{u} \in I\!R^P.$$

However, $\boldsymbol{d}_X(\boldsymbol{u})$ is not uniquely determined by (4.4)–(4.5). If there is no danger of confusion we omit the subindex "X" and write, e.g., $\boldsymbol{d}(\boldsymbol{u})$ instead of $\boldsymbol{d}_X(\boldsymbol{u})$.

In practice, space $I\!R^P$ is discretized. Let $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_P$ be the usual P orthonormal vectors in $I\!R^P$. We shall denote by E the set of all vectors $\boldsymbol{u}$ in $I\!R^P$ such that $\boldsymbol{u} = u_1\mathbf{e}_1 + u_2\mathbf{e}_2 + \cdots + u_P\mathbf{e}_P$, for some $u_1, u_2, \ldots, u_P \in \mathbb{Z}$. In this case, the distance transform should be either re-defined [4, 5, 6, 21, 31], or computed via (4.3), by replacing X with $X \cap E$ and by considering $\boldsymbol{u} \in E$. The first approach is limited by the types of norms $\|\cdot\|$ allowed. The second approach is quite general and permits *exact* distance computations for any choice of $\|\cdot\|$. We shall adopt this approach here, and shall focus on the problem of computing the vector distance transform $\boldsymbol{d}_X(\boldsymbol{u}) \in E$, by means of the morphological operators discussed in Subsection 3.2.

Consider the space of all image sequences $\text{Fun}(E, \overline{\mathbb{Z}}^P)$ and let $X \subseteq E$. Let $\boldsymbol{X}$ be the element in $\text{Fun}(E, \overline{\mathbb{Z}}^P)$ with

$$X_p(\boldsymbol{u}) = \begin{cases} +\infty, & \boldsymbol{u} \notin X \\ 0, & \boldsymbol{u} \in X, \end{cases} \tag{4.6}$$

for every $p = 1, 2, \ldots, P$. Put

$$\boldsymbol{C}(\boldsymbol{z}) = -\boldsymbol{z}, \;\; \boldsymbol{z} \in E.$$

From (4.4),(4.5), and (4.6) we get that

$$\boldsymbol{d}(\boldsymbol{u}) = \boldsymbol{X}(\boldsymbol{u} + \boldsymbol{z}_e(\boldsymbol{u})) - \boldsymbol{C}(\boldsymbol{z}_e(\boldsymbol{u})), \;\; \boldsymbol{u} \in E \tag{4.7}$$

where $\boldsymbol{z}_e = \boldsymbol{z}_e(\boldsymbol{u})$ is such that

$$\|\boldsymbol{X}(\boldsymbol{u} + \boldsymbol{z}_e) - \boldsymbol{C}(\boldsymbol{z}_e)\| = \bigwedge_{\boldsymbol{z} \in E} \|\boldsymbol{X}(\boldsymbol{u} + \boldsymbol{z}) - \boldsymbol{C}(\boldsymbol{z})\|. \tag{4.8}$$

If the norm $\|\cdot\|$ would satisfy (3.14), i.e.,

$$\|\boldsymbol{u}\| \leq \|\mathbf{v}\| \iff \|\boldsymbol{u} + \boldsymbol{z}\| \leq \|\mathbf{v} + \boldsymbol{z}\|, \tag{4.9}$$

for $\boldsymbol{u}, \mathbf{v}, \boldsymbol{z} \in E$, then $\boldsymbol{d}(\boldsymbol{u})$ in (4.7) would be given by the h-erosion (with $h(\boldsymbol{u}) = \|\boldsymbol{u}\|$) of $\boldsymbol{X}$ by $\boldsymbol{C}$ at the point $\boldsymbol{u}$; see Proposition 3.5 and Algorithm 3.7. Computation of the vector distance transform by means of an erosion is desirable for a variety of theoretical and computational reasons. For example, we can always define the adjoint dilation, which, in turn, may allow the construction of useful morphological algorithms. Additionally, morphological computer hardware may be directly used here.

Unfortunately, the norm $\|\cdot\|$ does not satisfy (4.9) in general. We make the following assumption on the norm $\|\cdot\|$.

4.1. Assumption. There exists a mapping $\boldsymbol{r} : E \to \overline{\mathbb{Z}}^P$ and a mapping $h : \overline{\mathbb{Z}}^P \to \overline{\mathbb{R}}_+$ such that

$$\|\boldsymbol{u}\| = h(\boldsymbol{r}(\boldsymbol{u})),$$

$$\boldsymbol{r}(+\infty) = +\infty,$$

$$h(\boldsymbol{i}) \leq h(\boldsymbol{j}) \iff h(\boldsymbol{i}+\boldsymbol{n}) \leq h(\boldsymbol{j}+\boldsymbol{n}),$$

for all $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{n} \in \overline{\mathbb{Z}}^P$.

It is easy to verify that

$$\boldsymbol{r}(\boldsymbol{X}(\boldsymbol{u}+\boldsymbol{z}) - \boldsymbol{C}(\boldsymbol{z})) = \boldsymbol{X}(\boldsymbol{u}+\boldsymbol{z}) - \boldsymbol{B}(\boldsymbol{z}),$$

where

$$\boldsymbol{B}(\boldsymbol{z}) = -\boldsymbol{r}(-\boldsymbol{C}(\boldsymbol{z})) = -\boldsymbol{r}(\boldsymbol{z}), \ \ \boldsymbol{z} \in E.$$

Thus we get that

$$\|\boldsymbol{X}(\boldsymbol{u}+\boldsymbol{z}) - \boldsymbol{C}(\boldsymbol{z})\| = h(\boldsymbol{X}(\boldsymbol{u}+\boldsymbol{z}) - \boldsymbol{B}(\boldsymbol{z})).$$

From (4.7) and (4.8) we get that $\boldsymbol{r}(\boldsymbol{d}(\boldsymbol{u}))$ is the h-erosion of $\boldsymbol{X}$ by $\boldsymbol{B}$, i.e.,

$$\boldsymbol{r}(\boldsymbol{d}(\boldsymbol{u})) = \boldsymbol{X}(\boldsymbol{u}+\boldsymbol{z}_e) - \boldsymbol{B}(\boldsymbol{z}_e), \tag{4.10}$$

where $\boldsymbol{z}_e = \boldsymbol{z}_e(\boldsymbol{u})$ is such that

$$h(\boldsymbol{X}(\boldsymbol{u}+\boldsymbol{z}_e) - \boldsymbol{B}(\boldsymbol{z}_e)) = \bigwedge_{\boldsymbol{z}\in E} h(\boldsymbol{X}(\boldsymbol{u}+\boldsymbol{z}) - \boldsymbol{B}(\boldsymbol{z})). \tag{4.11}$$

Calculation of the vector distance transform by means of these formulas is not efficient, since it requires computation of the infimum $\bigwedge_{\boldsymbol{z}\in E} h(\boldsymbol{X}(\boldsymbol{u}+\boldsymbol{z}) - \boldsymbol{B}(\boldsymbol{z}))$, for every $\boldsymbol{u} \in E$. These calculations may be performed recursively, under additional constraints on $\boldsymbol{r}$ and h. Indeed, let us define subsets E_k of E by

$$E_k = \{\boldsymbol{u} \in E \mid \boldsymbol{u} = (u_1, \ldots, u_P), \ 0 \leq |u_p| \leq k, \ p = 1, 2, \ldots, P\}$$

for $k = 1, 2, \ldots$. Clearly, $E_k \subseteq E_{k+1} \subseteq E$ for all k. Let $D_k = E_k \setminus E_{k-1}$ and $E_0 = D_0 = \{\boldsymbol{0}\}$. In this case

$$D_k \cap D_l = \varnothing, \quad k \neq l,$$

$$\bigcup_{k=0}^{n} D_k = E_n, \quad n \geq 1,$$

and

$$\bigcup_{k=0}^{\infty} D_k = E.$$

Define a sequence $\{\boldsymbol{B}_k \mid k = 0, 1, \ldots\}$ of structuring functions by

$$\boldsymbol{B}_k(\boldsymbol{u}) = \begin{cases} -\boldsymbol{r}(\boldsymbol{u}), & \boldsymbol{u} \in E_k \\ -\infty, & \text{otherwise.} \end{cases} \tag{4.12}$$

Observe that

$$\boldsymbol{B}_k(\boldsymbol{u}) \leq \boldsymbol{B}_{k+1}(\boldsymbol{u}), \ \ \boldsymbol{u} \in E, \quad \text{and} \quad \boldsymbol{B}_k(\boldsymbol{u}) = \boldsymbol{B}_{k+1}(\boldsymbol{u}), \ \ \boldsymbol{u} \in E_k.$$

Consider another sequence $\{A_k \mid k = 0,1,\ldots\}$ of structuring functions defined by

$$A_k(u) = \begin{cases} r(ku) - r((k+1)u), & u \in E_1 \\ -\infty, & \text{otherwise.} \end{cases} \tag{4.13}$$

Clearly,

$$A_k(0) = 0, \quad k = 1,2,\ldots,$$

and, presumed that $r(0) = 0$,

$$A_0(u) = B_1(u), \quad u \in E.$$

For $F, A \in \mathrm{Fun}(E, \overline{\mathbb{Z}}^P)$, let us define

$$\{F\} \ominus_h A = \{G \in \mathrm{Fun}(E, \overline{\mathbb{Z}}^P) \mid G(u) = F(u+z_e) - A(z_e), \text{ for some } z_e \in C(F,A,u)\}, \tag{4.14}$$

where

$$C(F,A,u) = \{z_e \in E \mid h(F(u+z_e) - A(z_e)) = \bigwedge_{z \in E} h(F(u+z) - A(z))\}. \tag{4.15}$$

For $\mathcal{F} \subseteq \mathrm{Fun}(E, \overline{\mathbb{Z}}^P)$, we also define

$$\mathcal{F} \ominus_h A = \bigcup_{F \in \mathcal{F}} \{F\} \ominus_h A$$

and write

$$(\mathcal{F} \ominus_h A)(u) = \{G(u) \mid G \in \mathcal{F} \ominus_h A\}, \ u \in E. \tag{4.16}$$

Observe that, for $G_1, G_2 \in \{F\} \ominus_h A$, $h(G_1(u)) = h(G_2(u))$, for every $u \in E$. We shall denote by $h((\{F\} \ominus_h A)(u))$ the value $h(G(u))$, for $G \in \{F\} \ominus_h A$.

In the Appendix we shall prove the following result.

4.2. Proposition. *Let $r : E \to \overline{\mathbb{Z}}^P$ be a mapping that satisfies*

$$r(u) = r(u_1,\ldots,u_P) = (r_1(|u_1|),\ldots,r_P(|u_P|)), \tag{4.17}$$

where $r_p(0) = 0$, $r_p(+\infty) = +\infty$, $r_p(k) < \infty$ if $k < \infty$, and

$$r_p(k+1) - r_p(k) \geq r_p(k) - r_p(k-1) > 0, \tag{4.18}$$

for $k = 1,2,\ldots$ and $p = 1,2,\ldots,P$. Furthermore, let $h : \overline{\mathbb{Z}}^P \to \overline{\mathbb{R}}_+$ be a mapping such that $h = fg$, where $g(i) = \sum_{p=1}^P i_p$ and f is a lattice isomorphism on $\overline{\mathbb{R}}_+$. Then

$$(\{X\} \ominus_h B_{k-1}) \ominus_h A_{k-1} \subseteq \{X\} \ominus_h B_k.$$

In practice, we are interested in computing $d(u)$, at every point $u \in E$ such that $\Delta_X(u) < +\infty$. From (4.10), (4.11), (4.14), and (4.15), it is easy to show that,

$$i \in (\{X\} \ominus_h B_\infty)(u) = (\{X\} \ominus_h B_K)(u) \iff r(d(u)) = i, \tag{4.19}$$

for every u such that $\Delta_X(u) < +\infty$, where K is a large enough finite integer which depends on u. From Proposition 4.2 and (4.19), it is now easy to show by induction that

$$(\{X\} \ominus_h B_K)(u) \supseteq ((\cdots((\{X\} \ominus_h A_0) \ominus_h A_1) \ominus_h \cdots) \ominus_h A_{K-1})(u), \tag{4.20}$$

for every $u \in E$ such that $|d_p(u)| \leq K$, $p = 1, 2, \ldots, P$, and, therefore, any element in the right-hand side of (4.20) will be the vector distance transform, assigned at $\mathbf{u}$. Equations (4.10), (4.11), and (4.20) show that the vector distance transform (4.4) and (4.5) can be recursively computed by means of H-invariant h-erosions with varying vector structuring functions of size 3×3. From the assumed form (4.17) of $\mathbf{r}$, we see that only $|\mathbf{d}(\mathbf{u})|$ can be recovered from $\mathbf{r}(\mathbf{d}(\mathbf{u}))$, provided that $r_p(\cdot)$ is invertible, for every $p = 1, 2, \ldots, P$. Recovery of the correct signs in $\mathbf{d}(\mathbf{u})$ can be accomplished by considering all possible sign combinations and by checking whether the resulting vectors $\mathbf{u} + \mathbf{d}(\mathbf{u})$ are in X.

Let us now illustrate the previous discussion with an example. Consider the problem of calculating the *Hölder vector distance transform*, for which [40]

$$\|\mathbf{u}\| = \left(\sum_{p=1}^{P} |u_p|^m \right)^{1/m}, \ 1 \leq m < +\infty.$$

If we set

$$r_p(\mathbf{u}) = r_p(|u_p|) = |u_p|^m, \text{ for } p = 1, 2, \ldots, P, \tag{4.21}$$

$$g(\mathbf{i}) = \sum_{p=1}^{P} i_p,$$

and

$$f(s) = (s)^{1/m},$$

then $\|\mathbf{u}\| = f(g(\mathbf{r}(\mathbf{u})))$, where $\mathbf{r}$ is given by (4.17) and (4.21), and mappings r_p, g, and f, satisfy all requirements in Proposition 4.2. Furthermore, $r_p(\cdot)$ is invertible, and

$$h(\mathbf{i}) = \left(\sum_{p=1}^{P} i_p \right)^{1/m}.$$

The Hölder vector distance transform can be now computed by means of (4.20) , where (see (4.13), (4.17), and (4.21))

$$A_{k,p}(\mathbf{u}) = \begin{cases} (k^m - (k+1)^m)|u_p|, & \mathbf{u} \in E_1 \\ -\infty, & \text{otherwise} \end{cases}, \text{ for } k = 0, 1, \ldots.$$

When $m = 1$ (which is the case of the *city-block distance* [31, 41]) we have that

$$A_{k,p}(\mathbf{u}) = \begin{cases} -|u_p|, & \mathbf{u} \in E_1 \\ -\infty, & \text{otherwise} \end{cases}, \text{ for } k = 0, 1, \ldots,$$

whereas, when $m = 2$ (which is the case of the *Euclidean distance* [6, 31]), we have that

$$A_{k,p}(\mathbf{u}) = \begin{cases} -2(k+1)|u_p|, & \mathbf{u} \in E_1 \\ -\infty, & \text{otherwise} \end{cases}, \text{ for } k = 0, 1, \ldots.$$

In the first case, the vector structuring function is independent of the iteration number k, whereas, in the second case, the vector structuring function does depend on k.

To conclude this section, we should point-out that mathematical morphology has been also proposed in [15, 37, 38] (see also [46]) as a tool for calculating the two-dimensional Euclidean distance transform. However, our approach here is more general, and demonstrates the fact that, such computations can be performed by means of more general morphological transformations, which directly deal with image sequences.

5. Discussion

If one tries to extend mathematical morphology to image sequences (i.e., multi-valued functions) one is faced with a very serious problem: there does not exist a canonical total ordering for vectors like for scalars. One may attempt to circumvent this problem by endowing the space of multi-valued functions with the marginal ordering, which defines a partial ordering. In this way, the space becomes a complete lattice, and one may use the abstract theory of morphological operators on complete lattices. Such an approach is very useful indeed in many situations; among others, it leads (as was also explained in [13]) to Wilson's theory of matrix morphology.

However, this approach, which leads to component-wise transformations, is bound to fail in situations where image components are correlated. In this paper we therefore suggest an alternative approach, also strongly based upon the complete lattice framework. The basic idea is to transform the image data by means of a surjective transformation h. The underlying assumption is that the transformed image data is better suited for the "classical" morphological approach. There may be several reasons for this to be true. Either h can map onto a totally ordered set (e.g., a subset of the extended reals), in which case classical gray-scale morphological operators can be used. We have discussed an example based on the Mahalanobis distance. Alternatively, h can be chosen in such a way that the data becomes uncorrelated. We have discussed an application based on the maximum noise fraction transform.

The major drawback of applying morphological operators on the transformed images instead on the original ones is that it is not a priori clear how to interpret the outcome in terms of the original data: in general h will not be invertible. To deal with such problems we introduce some new concepts: h-dilations and h-erosions, and mappings derived from them. To illustrate the great generality of our approach we have shown that vector distance transforms can be computed by means of h-erosions.

Acknowledgment.

The work of J. Goutsias and K. Sivakumar has been supported by the Office of Naval Research, Mathematical Sciences Division, under ONR Grant N00014-90-J-1345.

References.

[1] Astola, J., Haavisto, P., and Neuvo, Y. Vector median filters. *Proceedings of the IEEE 78* (1990), 678–689.

[2] Barnett, V. The ordering of multivariate data, (with discussion). *Journal of the Royal Statistical Society, Series A 139* (1976), 318–354.

[3] Birkhoff, G. *Lattice Theory*, 3rd ed., vol. 25 of *American Mathematical Society Colloquium Publications.* American Mathematical Society, Providence, RI, 1967.

[4] Borgefors, G. Distance transformations in arbitrary dimensions. *Computer Vision, Graphics and Image Processing 27* (1984), 321–345.

[5] Borgefors, G. Distance transformations in digital images. *Computer Vision, Graphics and Image Processing 34* (1986), 344–371.

[6] Danielsson, P. E. Euclidean distance mapping. *Computer Graphics and Image Processing 14* (1980), 227–248.

[7] Dugundji, J. *Topology.* Allyn and Bacon, Boston, 1966.

[8] GOUTSIAS, J. Morphological transformations of image sequences: A lattice theory approach. In *Image Algebra and Morphological Image Processing III* (San Diego, 1992), vol. 1769 of *SPIE Proceedings*, pp. 306–317.

[9] GREEN, A. A., BERMAN, M., SWITZER, P., AND CRAIG, M. D. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Transactions on Geoscience and Remote Sensing 26* (1988), 65–74.

[10] HARALICK, R. M., STERNBERG, S. R., AND ZHUANG, X. Image analysis using mathematical morphology. *IEEE Transactions on Pattern Analysis and Machine Intelligence 9* (1987), 532–550.

[11] HARDIE, R. C., AND ARCE, G. R. Ranking in R^P and its use in multivariate image estimation. *IEEE Transactions on Circuits and Systems for Video Technology 1* (1991), 197–209.

[12] HEIJMANS, H. J. A. M. Theoretical aspects of gray-level morphology. *IEEE Transactions on Pattern Analysis and Machine Intelligence 13* (1991), 568–582.

[13] HEIJMANS, H. J. A. M. *Morphological Image Operators.* Academic Press, Boston, 1994.

[14] HEIJMANS, H. J. A. M., AND RONSE, C. The algebraic basis of mathematical morphology – part I: Dilations and erosions. *Computer Vision, Graphics and Image Processing 50* (1990), 245–295.

[15] HUANG, C.-T., AND MITCHELL, O. R. Rapid Euclidean distance transform using grayscale morphology decomposition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Maui, Hawaii* (1991), pp. 695–697.

[16] JAIN, A. K. *Fundamentals of Digital Image Processing.* Prentice Hall, Englewood Cliffs, New Jersey, 1989.

[17] LEE, J. B., WOODYATT, A. S., AND BERMAN, M. Enhancement of high spectral resolution remote-sensing data by a noise-adjusted principal components transform. *IEEE Transactions on Geoscience and Remote Sensing 28* (1990), 295–304.

[18] LEYMARIE, F., AND LEVINE, M. D. Simulating the grassfire transform using an active contour model. *IEEE Transactions on Pattern Analysis and Machine Intelligence 14* (1992), 56–75.

[19] MAKHOUL, J., ROUCOS, S., AND GISH, H. Vector quantization in speech coding. *Proceedings of the IEEE 73* (1985), 1551–1558.

[20] MEYER-NIEBERG, P. *Banach Lattices.* Springer-Verlag, Berlin, 1991.

[21] MONTANARI, U. A method for obtaining skeletons using a quasi-Euclidean distance. *Journal of the Association for Computing Machinery 15* (1968), 600–624.

[22] NIBLACK, C. W., GIBBONS, P. B., AND CAPSON, D. W. Generating skeletons and centerlines from the distance transform. *Computer Vision, Graphics and Image Processing: Graphical Models and Image Processing 54* (1992), 420–437.

[23] PAGLIERONI, D. W. Distance transforms: Properties and machine vision applications. *Computer Vision, Graphics and Image Processing: Graphical Models and Image Processing 54* (1992), 56–74.

[24] PEI, S.-C., AND CHEN, F.-C. Subband decomposition of monochrome and color images by mathematical morphology. *Optical Engineering 30* (1991), 921–933.

[25] PITAS, I., AND TSAKALIDES, P. Multivariate ordering in color image filtering. *IEEE Transactions on Circuits and Systems for Video Technology 1* (1991), 247–259.

[26] PRETEAUX, F. Watershed and skeleton by influence zones: A distance-based approach. *Journal of Mathematical Imaging and Vision 1* (1992), 239–255.

[27] RAGNEMALM, I. Generation of Euclidean distance maps. Thesis 206, Linköping Studies in Science and Technology, Department of Electrical Engineering, Linköping University, 1990.

[28] RICHARDS, J. A. *Remote Sensing Digital Image Analysis.* Springer-Verlag, Berlin, 1986.

[29] RONSE, C. Why mathematical morphology needs complete lattices. *Signal Processing 21* (1990), 129–154.

[30] RONSE, C., AND HEIJMANS, H. J. A. M. The algebraic basis of mathematical morphology – part II: Openings and closings. *Computer Vision, Graphics and Image Processing: Image Understanding 54* (1991), 74–97.

[31] ROSENFELD, A., AND PFALTZ, J. L. Distance functions on digital pictures. *Pattern Recognition 1* (1968), 33–61.

[32] SCHALKOFF, R. *Pattern Recognition: Statistical, Structural, and Neural Approaches.* John Wiley and Sons, New York, 1992.

[33] SERRA, J. *Image Analysis and Mathematical Morphology.* Academic Press, London, 1982.

[34] SERRA, J., Ed. *Image Analysis and Mathematical Morphology. II: Theoretical Advances.* Academic Press, London, 1988.

[35] SERRA, J. Anamorphoses and function lattices (multivalued morphology). In *Mathematical Morphology in Image Processing*, E. R. Dougherty, Ed. Marcel Dekker, New York, 1993, ch. 13, pp. 483–523.

[36] SERRA, J., AND VINCENT, L. An overview of morphological filtering. *Circuits, Systems and Signal Processing 11* (1992), 47–108.

[37] SHIH, F. Y., AND MITCHELL, O. R. A mathematical morphology approach to Euclidean distance transformation. *IEEE Transactions on Image Processing 1* (1992), 197–204.

[38] SHIH, F. Y., AND WU, H. Optimization on Euclidean distance transform using grayscale morphology. *Journal of Visual Communication and Image Representation 3* (1992), 104–114.

[39] STARK, D. D., AND BRADLEY JR., W. G. *Magnetic Resonance Imaging.* C. V. Mosby Company, St. Louis, Missouri, 1988.

[40] STEWART, G. W. *Introduction to Matrix Computations.* Academic Press, New York, 1973.

[41] TORIWAKI, J.-I., AND YOKOI, S. Distance transformations and skeletons of digitized pictures with applications. In *Progress in Pattern Recognition*, L. N. Kanal and A. Rosenfeld, Eds. North-Holland, New York, 1981, pp. 187–264.

[42] WILSON, S. S. Morphological networks. In *Visual Communications and Image Processing '89* (Philadelphia, 1989), vol. 1199 of *SPIE Proceedings*, pp. 483–493.

[43] WILSON, S. S. Vector morphology and iconic neural networks. *IEEE Transactions on Systems, Man, and Cybernetics 19* (1989), 1636–1644.

[44] WILSON, S. S. Theory of matrix morphology. *IEEE Transactions on Pattern Analysis and Machine Intelligence 14* (1992), 636–652.

[45] WILSON, S. S. Training structuring elements in morphological networks. In *Mathematical Morphology in Image Processing*, E. R. Dougherty, Ed. Marcel Dekker, New York, 1993, ch. 1, pp. 1–41.

[46] YAMADA, H. Complete Euclidean distance transform by parallel operation. In *Proceedings of the 7th IEEE International Conference on Pattern Recognition* (Montreal, Canada, 1984), pp. 69–71.

Appendix.

In this appendix we present a proof of Proposition 4.2. We start with the following lemma.

A.1. Lemma *Let* $\mathbf{r} : E \to \overline{\mathbb{Z}}^P$ *be a mapping that satisfies the requirements of Proposition 4.2. Then*

$$\mathbf{r}(\mathbf{u}) - \mathbf{r}(k\mathbf{v}) + \mathbf{r}((k+1)\mathbf{v}) \geq \mathbf{r}(|\mathbf{u}| + |\mathbf{v}|) \geq \mathbf{r}(\mathbf{u}+\mathbf{v}), \ \mathbf{u} \in E_k, \ \mathbf{v} \in E_1, \tag{A.1}$$

for $k = 1, 2, \ldots$, *where both inequalities and* $|\cdot|$ *are applied component-wise.*

PROOF. From the assumed form of mapping $\mathbf{r}$ we have

$$(\mathbf{r}(\mathbf{u}) - \mathbf{r}(k\mathbf{v}) + \mathbf{r}((k+1)\mathbf{v}))_p = r_p(|u_p|) - r_p(k|v_p|) + r_p((k+1)|v_p|).$$

Since $\mathbf{v} \in E_1$, $|v_p| \leq 1$, for $p = 1, 2, \ldots, P$. If $v_p = 0$ for some $p = 1, 2, \ldots, P$, then, since $r_p(0) = 0$,

$$r_p(|u_p|) - r_p(k|v_p|) + r_p((k+1)|v_p|) = r_p(|u_p|) = r_p(|u_p| + |v_p|).$$

Since $\mathbf{u} \in E_k$, $|u_p| \leq k$, for $p = 1, 2, \ldots, P$. If $|v_p| = 1$ for some $p = 1, 2, \ldots, P$, then (see also (4.18))

$$\begin{aligned} r_p(|u_p|) - r_p(k|v_p|) + r_p((k+1)|v_p|) &= r_p(|u_p| + 1) - (r_p(|u_p| + 1) - r_p(|u_p|)) \\ &\quad + (r_p(k+1) - r_p(k)) \\ &\geq r_p(|u_p| + 1) = r_p(|u_p| + |v_p|). \end{aligned}$$

This shows the first inequality in (A.1). The second inequality is obvious from (4.17) and (4.18), since $|u_p| + |v_p| \geq |u_p + v_p|$. ∎

PROOF OF PROPOSITION 4.2 It suffices to show that (see also (4.16)),

$$((\{\boldsymbol{X}\} \ominus_h \boldsymbol{B}_{k-1}) \ominus_h \boldsymbol{A}_{k-1})(\mathbf{u}) \subseteq (\{\boldsymbol{X}\} \ominus_h \boldsymbol{B}_k)(\mathbf{u}), \ \mathbf{u} \in E. \tag{A.2}$$

If $\mathbf{u} + \mathbf{z}_e \notin X$ or $\mathbf{z}_e \notin E_k$, for some, and hence for all, $\mathbf{z}_e \in C(\boldsymbol{X}, \boldsymbol{B}_k, \mathbf{u})$, then the right hand-side of (A.2) contains only $+\infty$. On the left hand-side, $\{\boldsymbol{X}\} \ominus_h \boldsymbol{B}_{k-1}$ is a collection of functions, all of which take the value $+\infty$ at the points $\mathbf{u} + \mathbf{v}$, $\mathbf{v} \in E_1$. Hence the set $((\{\boldsymbol{X}\} \ominus_h \boldsymbol{B}_{k-1}) \ominus_h \boldsymbol{A}_{k-1})(\mathbf{u})$ on the left hand-side contains only the vector $+\infty$. Thus (A.2) holds in this case.

We assume that $\mathbf{u} + \mathbf{z}_e \in X$ and $\mathbf{z}_e \in E_k$, for some, and hence for all, $\mathbf{z}_e \in C(\boldsymbol{X}, \boldsymbol{B}_k, \mathbf{u})$. We shall first show the following three assertions:

(1) For every $\mathbf{v} \in E_1$

$$h(\boldsymbol{G}(\mathbf{u}+\mathbf{v}) - \boldsymbol{A}_{k-1}(\mathbf{v})) \geq h((\{\boldsymbol{X}\} \ominus_h \boldsymbol{B}_k)(\mathbf{u})), \ \boldsymbol{G} \in \{\boldsymbol{X}\} \ominus_h \boldsymbol{B}_{k-1}. \tag{A.3}$$

(2) There exists a $\mathbf{v} \in E_1$ for which equality holds in (A.3). It is clear from the form of h that, for a given $\mathbf{v} \in E_1$, if equality holds in (A.3) for some $\boldsymbol{G} \in \{\boldsymbol{X}\} \ominus_h \boldsymbol{B}_{k-1}$, then equality holds for all $\boldsymbol{G} \in \{\boldsymbol{X}\} \ominus_h \boldsymbol{B}_{k-1}$. In the sequel, we shall denote by $E_1'(\mathbf{u}) \subseteq E_1$, the set of vectors $\mathbf{v}$ for which equality holds in (A.3).

(3) For every $\mathbf{v} \in E_1'(\mathbf{u})$, $\mathbf{v} + \mathbf{z}_e' \in C(\boldsymbol{X}, \boldsymbol{B}_k, \mathbf{u})$, for every $\mathbf{z}_e' \in C(\boldsymbol{X}, \boldsymbol{B}_{k-1}, \mathbf{u}+\mathbf{v})$.

When $\mathbf{z}_e \in C(\boldsymbol{X}, \boldsymbol{B}_k, \mathbf{u})$, we have (see also (4.14) and (4.15))

$$h((\{\boldsymbol{X}\} \ominus_h \boldsymbol{B}_k)(\mathbf{u})) = h(\boldsymbol{X}(\mathbf{u}+\mathbf{z}_e) - \boldsymbol{B}_k(\mathbf{z}_e)) = h(\mathbf{r}(\mathbf{z}_e)) \leq h(\mathbf{r}(\mathbf{z})) < +\infty, \tag{A.4}$$

for every $\mathbf{z} \in E_k$ such that $\mathbf{u} + \mathbf{z} \in X$. When $\boldsymbol{G} \in \{\boldsymbol{X}\} \ominus_h \boldsymbol{B}_{k-1}$, we have

$$\boldsymbol{G}(\mathbf{u}+\mathbf{v}) = \boldsymbol{X}(\mathbf{u}+\mathbf{v}+\mathbf{z}_e') - \boldsymbol{B}_{k-1}(\mathbf{z}_e'), \ \mathbf{v} \in E_1, \tag{A.5}$$

for some $z'_e \in C(X, B_{k-1}, u+v)$, and

$$h((\{X\} \ominus_h B_{k-1})(u+v)) = h(G(u+v)) = h(X(u+v+z'_e) - B_{k-1}(z'_e)), \tag{A.6}$$

for every $z'_e \in C(X, B_{k-1}, u+v)$. When $v \in E_1$, we have (see also (4.6) and (4.12))

$$X(u+v+z'_e) - B_{k-1}(z'_e) \geq r(z'_e),\ z'_e \in C(X, B_{k-1}, u+v), \tag{A.7}$$

where the inequality is a component-wise vector inequality. If $u+v+z'_e \notin X$ or $z'_e \notin E_{k-1}$, for some, and hence for all, $z'_e \in C(X, B_{k-1}, u+v)$, then clearly (see also (A.4)),

$$h(X(u+v+z'_e) - B_{k-1}(z'_e) - A_{k-1}(v)) = +\infty > h(r(z_e)), \tag{A.8}$$

which together with (A.4), (A.6) and the assumed form of h, shows (A.3). If $u+v+z'_e \in X$ and $z'_e \in E_{k-1}$, for some, and hence for all, $z'_e \in C(X, B_{k-1}, u+v)$, then (see also (4.13), (A.1), and (A.7))

$$X(u+v+z'_e) - B_{k-1}(z'_e) - A_{k-1}(v) \geq r(z'_e) - r((k-1)v) + r(kv) \geq r(v+z'_e), \tag{A.9}$$

where the inequalities are component-wise vector inequalities. From (A.4), (A.9), and the assumed form of h, we obtain

$$h(X(u+v+z'_e) - B_{k-1}(z'_e) - A_{k-1}(v)) \geq h(r(v+z'_e)) \geq h(r(z_e)), \tag{A.10}$$

where we have used the fact that $v+z'_e \in E_k$. From (A.4), (A.6), and (A.10), we obtain (A.3). This shows (1) above.

Next, we need to show that equality holds in (A.3), for some $v \in E_1$. Let $z_e = (a_1, \ldots, a_P) \in C(X, B_k, u)$ be fixed but arbitrary, so that $|a_p| \leq k$, for $p = 1, 2, \ldots, P$. Define $v \in E_1$ by

$$v_p = \begin{cases} 0, & \text{if } |a_p| < k \\ 1, & \text{if } a_p = k \\ -1, & \text{if } a_p = -k. \end{cases} \tag{A.11}$$

It follows that $z_e - v \in E_{k-1}$ and $|z_e - v| = |z_e| - |v|$. When $z'_e \in C(X, B_{k-1}, u+v)$ (see also (4.15)), we have that

$$h(r(z'_e)) = h(X(u+v+z'_e) - B_{k-1}(z'_e)) \leq h(X(u+v+z) - B_{k-1}(z)), \tag{A.12}$$

for all $z \in E_{k-1}$, and in particular,

$$h(r(z'_e)) \leq h(r(z_e - v)), \tag{A.13}$$

since $z_e - v \in E_{k-1}$ and $u+v+(z_e - v) = u + z_e \in X$. Let $v + z'_e = (m_1, \ldots, m_P)$, for some (for the moment fixed) $z'_e \in C(X, B_{k-1}, u+v)$, so that $|m_p| \leq k$, for $p = 1, 2, \ldots, P$. In this case, (A.10) can be written as

$$h(r(v+z'_e)) = f(\sum_{p=1}^{P} r_p(|m_p|)) \geq f(\sum_{p=1}^{P} r_p(|a_p|)) = h(r(z_e)), \tag{A.14}$$

and (A.13) can be written as

$$h(r(z'_e)) = f(\sum_{p=1}^{P} r_p(|m_p - v_p|)) \leq f(\sum_{p=1}^{P} r_p(|a_p - v_p|)) = h(r(z_e - v)). \tag{A.15}$$

Since $f(\cdot)$ is a lattice isomorphism, (A.14) and (A.15) become (see also (A.11))

$$\sum_{p=1}^{P} r_p(|m_p|) \geq \sum_{p=1}^{P} r_p(|a_p|), \tag{A.16}$$

and

$$\begin{aligned}
&\sum_{p=1}^{P} r_p(|m_p|) + \sum_{\{p|v_p=-1\}} [r_p(|m_p+1|) - r_p(|m_p|)] - \sum_{\{p|v_p=1\}} [r_p(|m_p|) - r_p(|m_p-1|)] \\
&\quad \leq \sum_{p=1}^{P} r_p(|a_p|) - \sum_{\{p|v_p=-1\}} [r_p(|a_p|) - r_p(|a_p+1|)] - \sum_{\{p|v_p=1\}} [r_p(|a_p|) - r_p(|a_p-1|)] \\
&\quad = \sum_{p=1}^{P} r_p(|a_p|) - \sum_{\{p|v_p=-1\}} [r_p(k) - r_p(k-1)] - \sum_{\{p|v_p=1\}} [r_p(k) - r_p(k-1)],
\end{aligned} \tag{A.17}$$

respectively. Combining (A.16) and (A.17) we get

$$\begin{aligned}
0 &\geq \sum_{p=1}^{P} r_p(|a_p|) - \sum_{p=1}^{P} r_p(|m_p|) \\
&\geq \sum_{\{p|v_p=-1\}} [r_p(|m_p+1|) - r_p(|m_p|)] - \sum_{\{p|v_p=1\}} [r_p(|m_p|) - r_p(|m_p-1|)] \\
&\quad + \sum_{\{p|v_p=-1\}} [r_p(k) - r_p(k-1)] + \sum_{\{p|v_p=1\}} [r_p(k) - r_p(k-1)] \\
&= \sum_{\{p|v_p=1,\, m_p>0\}} [(r_p(k) - r_p(k-1)) - (r_p(m_p) - r_p(m_p-1))] \\
&\quad + \sum_{\{p|v_p=1,\, m_p\leq 0\}} [(r_p(k) - r_p(k-1)) + (r_p(|m_p-1|) - r_p(|m_p|))] \\
&\quad + \sum_{\{p|v_p=-1,\, m_p\geq 0\}} [(r_p(k) - r_p(k-1)) + (r_p(m_p+1) - r_p(m_p))] \\
&\quad + \sum_{\{p|v_p=-1,\, m_p<0\}} [(r_p(k) - r_p(k-1)) - (r_p(|m_p|) - r_p(|m_p+1|))] \geq 0,
\end{aligned}$$

where the last inequality follows from (4.18) and the fact that $|m_p| \leq k$, for $p = 1, 2, \ldots, P$. Hence, we must have equality throughout, which, in particular, implies that equality holds in (A.14), for our choice of $\mathbf{v} \in E_1$. But since $\mathbf{z}'_e \in C(\mathbf{X}, \mathbf{B}_{k-1}, \mathbf{u}+\mathbf{v})$ was arbitrary, we have equality in (A.3). This shows (2) above.

Finally, if $\mathbf{v} \in E'_1(\mathbf{u})$, then we must have equality throughout in (A.9) and (A.10). Hence, $\mathbf{u}+\mathbf{v}+\mathbf{z}'_e \in X$ and $\mathbf{z}'_e \in E_{k-1}$, for some, and hence for all, $\mathbf{z}'_e \in C(\mathbf{X}, \mathbf{B}_{k-1}, \mathbf{u}+\mathbf{v})$, and

$$h(\mathbf{X}(\mathbf{u}+\mathbf{v}+\mathbf{z}'_e) - \mathbf{B}_{k-1}(\mathbf{z}'_e) - \mathbf{A}_{k-1}(\mathbf{v})) = h(\mathbf{r}(\mathbf{v}+\mathbf{z}'_e)) = h(\mathbf{r}(\mathbf{z}_e)) = h(\mathbf{X}(\mathbf{u}+\mathbf{z}_e) - \mathbf{B}_k(\mathbf{z}_e)),$$

which together with (4.15) and (A.4) gives $\mathbf{v}+\mathbf{z}'_e \in C(\mathbf{X}, \mathbf{B}_k, \mathbf{u})$, for every $\mathbf{z}'_e \in C(\mathbf{X}, \mathbf{B}_{k-1}, \mathbf{u}+\mathbf{v})$. This shows (3) above.

We can now complete the proof as follows: from (1) and (2), it follows that $E'_1(\mathbf{u})$ is non-empty and $C(\mathbf{G}, \mathbf{A}_{k-1}, \mathbf{u}) = E'_1(\mathbf{u})$, for every $\mathbf{G} \in \{\mathbf{X}\}\ominus_h \mathbf{B}_{k-1}$ (see also (4.15) and the assumed form of h). We then have (see also (4.12), (4.13), and (A.5))

$$
\begin{aligned}
&((\{\boldsymbol{X}\} \ominus_h \boldsymbol{B}_{k-1}) \ominus_h \boldsymbol{A}_{k-1})(\boldsymbol{u}) \\
&\quad = \{\boldsymbol{G}(\boldsymbol{u}+\mathbf{v}_e) - \boldsymbol{A}_{k-1}(\mathbf{v}_e) \mid \mathbf{v}_e \in C(\boldsymbol{G}, \boldsymbol{A}_{k-1}, \boldsymbol{u}),\ \boldsymbol{G} \in \{\boldsymbol{X}\} \ominus_h \boldsymbol{B}_{k-1}\} \\
&\quad = \{\mathbf{X}(\boldsymbol{u}+\mathbf{v}_e+\mathbf{z}'_e) - \boldsymbol{B}_{k-1}(\mathbf{z}'_e) - \boldsymbol{A}_{k-1}(\mathbf{v}_e) \mid \mathbf{v}_e \in E'_1(\boldsymbol{u}),\ \mathbf{z}'_e \in C(\mathbf{X}, \boldsymbol{B}_{k-1}, \boldsymbol{u}+\mathbf{v}_e)\} \\
&\quad = \{\mathbf{X}(\boldsymbol{u}+\mathbf{v}_e+\mathbf{z}'_e) + \mathbf{r}(\mathbf{z}'_e) - \mathbf{r}((k-1)\mathbf{v}_e) + \mathbf{r}(k\mathbf{v}_e) \mid \mathbf{v}_e \in E'_1(\boldsymbol{u}),\ \mathbf{z}'_e \in C(\mathbf{X}, \boldsymbol{B}_{k-1}, \boldsymbol{u}+\mathbf{v}_e)\}
\end{aligned}
$$

Also, for $\mathbf{v}_e \in E'_1(\boldsymbol{u})$ we have equality throughout in (A.9) and (A.10). Hence, we obtain

$$
\begin{aligned}
&((\{\mathbf{X}\} \ominus_h \boldsymbol{B}_{k-1}) \ominus_h \boldsymbol{A}_{k-1})(\boldsymbol{u}) \\
&\quad = \{\mathbf{X}(\boldsymbol{u}+\mathbf{v}_e+\mathbf{z}'_e) + \mathbf{r}(\mathbf{v}_e+\mathbf{z}'_e) \mid \mathbf{v}_e \in E'_1(\boldsymbol{u}),\ \mathbf{z}'_e \in C(\mathbf{X}, \boldsymbol{B}_{k-1}, \boldsymbol{u}+\mathbf{v}_e)\} \\
&\quad = \{\mathbf{X}(\boldsymbol{u}+\mathbf{v}_e+\mathbf{z}'_e) - \boldsymbol{B}_k(\mathbf{z}'_e+\mathbf{v}_e) \mid \mathbf{v}_e \in E'_1(\boldsymbol{u}),\ \mathbf{z}'_e \in C(\mathbf{X}, \boldsymbol{B}_{k-1}, \boldsymbol{u}+\mathbf{v}_e)\} \\
&\quad \subseteq \{\mathbf{X}(\boldsymbol{u}+\mathbf{z}_e) - \boldsymbol{B}_k(\mathbf{z}_e) \mid \mathbf{z}_e \in C(\mathbf{X}, \boldsymbol{B}_k, \boldsymbol{u})\} \\
&\quad = (\{\mathbf{X}\} \ominus_h \boldsymbol{B}_k)(\boldsymbol{u}),
\end{aligned}
$$

where we have used (3) in the inequality above. This shows (A.2) which completes the proof. ∎