Threshold Superposition in Morphological Image Analysis Systems

PETROS MARAGOS AND ROBERT D. ZIFF

Abstract—**In this correspondence it is shown that many composite morphological systems, such as morphological edge detection, peak/valley extraction, skeletonization, and shape-size distributions obey a weak linear superposition, called *threshold-linear superposition*. Namely, the output image signal or measurement from each system is the sum of outputs due to input binary images that result from thresholding the input gray-level image at all levels. Then these results are generalized to a vector space formulation, e.g., to any finite linear combination of simple morphological systems. Thus many such systems processing gray-level images are reduced to corresponding binary image processing systems, which are easier to analyze and implement.**

Index Terms—**Image analysis, mathematical morphology, thresholding.**

I. Introduction

Morphological image analysis systems [16], [17], [26], [27], [11], [30], [6], [14], [15] are useful for feature extraction, shape analysis, and nonlinear filtering. A major limitation, however, in further advancing their practical design and theoretical understanding has been the lack of analytic tools (analogous in functionality with the tools of linear systems theory), due to the nonlinearity of the signal operations involved. Specifically, the morphological image operations do not obey the well-known *additive superposition* principle, which is obeyed by all linear systems. However, it can be shown that (see Sections II and III) a special class of simple morphological operations, in particular the gray-level erosions, dilations, openings, closings that use a binary (flat) structuring element, obey a weak additive superposition: namely, if the input gray-level image is expressed as the sum of all its binary threshold versions, then the output image from any of these filters is the sum of the filtered input threshold binary images. We call this system property *threshold-linear superposition*. Such ideas have been proven very useful in analyzing and implementing median and other rank-order filters [4], [32], [7], [24], [8]. A very similar threshold superposition was used to analyze the basic morphological [26], [14] and median-type filters [15].

In practice, the useful morphological image analysis systems do not consist of individual erosions, dilations, openings, and closings, but they include parallel and/or series interconnections of simple morphological operations. For example: 1) the morphological peak/valley extractor involves a difference between the image and its opening [18], [19], [26]; 2) a morphological edge gradient scheme involves the difference between the image and its erosion [18], [1], [5]; 3) the gray-level skeleton is the sum of components, each of which is the difference between erosions and openings [25], [22], [26], [13]; 4) differential size distributions involve areas of differences among openings or closings by structuring elements of varying size [16], [17], [26], [12], [3].

In this correspondence (Section III), we show that all four above composite morphological systems obey the threshold-linear superposition. That is, given any input gray-level image, their outputs are the sum of the individual system outputs corresponding to input binary images that resulted from exhaustive thresholding of the input image. The processing of these threshold binary images is much easier to analyze and implement. Thus our results offer new tools for the theoretical analysis of these nonlinear systems and suggest new parallel implementations since the processing of the threshold binary images can take place in parallel at all threshold levels simultaneously. Finally (in Section IV), we generalize the above results by showing that the four above systems together with any other system that obeys threshold-linear superposition form a vector space.

Manuscript received September 1, 1988; revised August 15, 1989. This work was supported by the National Science Foundation under Grant MIPS-86-58150 with matching funds from Bellcore, Xerox, Sun, and an IBM Departmental Grant, and in part by ARO under Grant DAAL03-86-K-0171.

The authors are with the Division of Applied Sciences, Harvard University, Cambridge, MA 02138.

IEEE Log Number 8933764.

II. Threshold Superposition

In this correspondence, by a *system* Ψ processing an input image f we mean either an *image transformation* where the system output $\Psi(f)$ is an image signal, or an *image measurement*. In the latter case $\Psi(f)$ is either a real number (e.g., the area of the image) or a real function of several parameters measuring some characteristics of the image.

Consider a digital gray-level image signal represented by a nonnegative 2-D sequence $f(m, n)$, which assumes $A + 1$ possible integer intensity values: $a = 0, 1, 2, \cdots, A$. For example, if we deal with 8 bit/pixel imagery, $A = 255$. By thresholding f at all possible amplitude levels $0 \le a \le A$ we obtain the *threshold binary images*

$$f_a(m, n) = \begin{cases} 1, & f(m, n) \ge a \\ 0, & f(m, n) < a. \end{cases} \tag{1}$$

If there is a risk of notational confusion, we will also denote the signal f_a by $t_a(f)$. It is simple to show that f can be reconstructed exactly from all its binary thresholded versions; i.e., $\forall (m, n)$

$$f(m, n) = \sum_{a=1}^{A} f_a(m, n) \tag{2}$$

$$= \max \{a: f_a(m, n) = 1\}. \tag{3}$$

We shall say that an image transformation Ψ *commutes with thresholding* provided that

$$\Psi[t_a(f)] = t_a[\Psi(f)] \tag{4}$$

for any input image f and any amplitude level a. Note that a necessary condition for Ψ to obey (4) is, whenever Ψ processes a binary image, to leave this signal characteristic unchanged. Thus if a system Ψ commutes with thresholding, processing by Ψ the threshold binary image f_a gives the same result with processing first by Ψ the gray-level image f and then thresholding $\Psi(f)$ at level a.

The four basic morphological transformations of *erosion* $f \ominus B$ of an image f by a 2-D structuring set (finite window) B, *dilation* $f \oplus B$, *opening* $f \circ B$, and *closing* $f \bullet B$, which are defined below, commute with thresholding [26], [23].

$$(f \ominus B)(m, n) = \min \{f(m + i, n + j): (i, j) \in B\} \tag{5}$$

$$(f \oplus B)(m, n) = \max \{f(m - i, n - j): (i, j) \in B\} \tag{6}$$

$$f \circ B = (f \ominus B) \oplus B \tag{7}$$

$$f \bullet B = (f \oplus B) \ominus B \tag{8}$$

Thus, it is simple to show that

$$t_a(f \ominus B) = f_a \ominus B \tag{9}$$

$$t_a(f \oplus B) = f_a \oplus B \tag{10}$$

where the notation $x = y$ for two signals means $x(m, n) = y(m, n)\ \forall\ (m, n)$. Since any finite cascade (series interconnection) of systems commuting with thresholding commutes with thresholding too [23], the opening and closing commute with thresholding.

We shall say that a system Ψ obeys the *threshold-linear superposition* provided that

$$\Psi(f) = \sum_{a=1}^{A} \Psi(f_a) \tag{11}$$

for any input image f. (Although f_a is binary, note that $\Psi(f_a)$, if it is an image signal, could be binary or multilevel.) Such a system Ψ can be realized by decomposing f into all its threshold binary images f_a, processing them by Ψ, and create the output $\Psi(f)$ by adding the processed f_a. A fundamental motivation for such a realization of Ψ is that, due to their binary range, the processing of the f_a's by Ψ is easier to analyze and implement than the processing of f.

The erosion, dilation, opening, and closing (5)–(8) obey a *threshold-max superposition* [26]:

$$[\Psi(f)](m, n) = \max\{a: [\Psi(f_a)](m, n) = 1\}. \tag{12}$$

This max-superposition is also obeyed by median and rank-order filters [4], [15]. Thus, these simple morphological and median-type systems obey both the threshold-sum superposition (11) and the threshold-max superposition (12). This happens because they commute with thresholding, which is a sufficient (but not necessary) condition for both types of threshold superposition.

From one viewpoint, the threshold-max superposition is more general than the threshold-sum superposition since the latter applies only to nonnegative input signals, while the former applies to any real-valued input signals. From a different viewpoint, the max-superposition restricts the class of systems since it requires that $\Psi(f_a)$ are signals and binary, an assumption not needed by the sum-superposition. In addition, the threshold-sum superposition ties well with linear systems, because it is just a weak form of linear superposition. This last viewpoint will be instrumental for our analysis throughout the rest of this paper. Therefore, we focus henceforth on systems obeying (11).

III. Special Cases

A. Morphological Edge Gradients

Given a gray-level image $f(m, n)$ and a small 2-D symmetric structuring set K containing the origin, the simple system [5]

$$EG(f) = f - (f \ominus K), \tag{13}$$

which we may call an *erosion gradient*, produces a gray-level image $EG(f)$ with enhanced edges, where $-$ denotes pointwise subtraction. Its value at (m, n) is equal to the maximum amplitude difference of f within the neighborhood K centered at (m, n). A binary edge map can be obtained by thresholding $EG(f)$, which is nonnegative everywhere.[1] Now, from (9) and (2) it follows that the erosion obeys threshold-linear superposition:

$$f \ominus K = \sum_a f_a \ominus K. \tag{14}$$

Hence, using the threshold representation (2) of f yields

$$\begin{aligned} EG(f) &= EG\left(\sum_{a=1}^{A} f_a\right) \\ &= \left(\sum_{a=1}^{A} f_a\right) - \left[\left(\sum_{a=1}^{A} f_a\right) \ominus K\right] \\ &= \sum_{a=1}^{A} f_a - \sum_{a=1}^{A} f_a \ominus K = \sum_{a=1}^{A} [f_a - (f_a \ominus K)] \\ &= \sum_{a=1}^{A} EG(f_a). \end{aligned} \tag{15}$$

[1]The same binary edge map can obtained by the system $\max_b \{t_{b+a}(f) - t_{b+1}(f \ominus K)\}$ proposed in [18] using set-theoretic terminology, where a is the variable edge height. In [26, pp. 435–437] this binary edge map was shown to be equal to $t_a[f - (f \ominus K)]$.

Thus the erosion gradient system (13) obeys the threshold-linear superposition. Examples are given in Figs. 1 and 2. Note that, since $f_a \ominus K$ and $f_a - (f_a \ominus K)$ are binary images for all a, the binary edge gradients $EG(f_a)$ can be implemented very simply by using only *pixel counting* or Boolean logic. Namely, if $|\cdot|$ denotes number of pixels, then

$$[EG(f_a)](m, n) = \begin{cases} 1, & \text{if } f_a(m, n) = 1 \text{ and } \sum_{(i,j)\in K} f_a(m + i, n + j) < |K| \\ 0, & \text{otherwise.} \end{cases}$$

Other edge gradients similar to (13) are the *dilation gradient* $DG(f) = (f \oplus K) - f$ and Beucher's *morphological gradient* [1], [26, p. 441]

$$MG(f) = (f \oplus K) - (f \ominus K). \tag{16}$$

These two systems also obey threshold-linear superposition; e.g., $MG(f) = \Sigma_a MG(f_a)$. This can be proven by proceeding as in (15) and observing that the dilation $f \oplus K = \Sigma_a f_a \oplus K$ obeys (11).

By combining the erosion and dilation gradients, various edge operators have been developed for robust edge detection schemes. Examples include: 1) the morphological edge-strength operators $\min[EG(f), DG(f)]$ and $\max[EG(f), DG(f)]$ in [10]; 2) the nonlinear Laplace operator $DG(f) - EG(f)$ in [31]. It is simple to show that these nonlinear edge operators obey the threshold-linear superposition.

B. Peak/Valley Extraction

A very useful morphological peak extractor, also called *top-hat transformation* [18], [19], [26, pp. 435–437], is the system

$$PE(f) = f - (f \circ B), \tag{17}$$

where B is any 2-D structuring set (the "base of the hat"). During this peak extraction, the peaks that cannot contain B remain, while the rest get eliminated. For any B, $f \geq f \circ B$ everywhere; hence, $PE(f)$ is a nonnegative image signal. Since the opening $f \circ B$ obeys (11),

$$\begin{aligned} PE(f) &= \left(\sum_{a=1}^{A} f_a\right) - \left[\left(\sum_{a=1}^{A} f_a\right) \circ B\right] \\ &= \sum_{a=1}^{A} f_a - \sum_{a=1}^{A} f_a \circ B = \sum_{a=1}^{A} [f_a - (f_a \circ B)] \\ &= \sum_{a=1}^{A} PE(f_a). \end{aligned} \tag{18}$$

As an example, consider the 1-D image $f(m)$

$$f = \cdots 0\,2\,1\,2\,3\,4\,0\,4\,4\,1\,2\,3\,2\,1\,0 \cdots$$

where $\cdots$ denotes an infinite sequence of trailing zero values. If we want to extract from f all peaks with a width less than 3 pixels, we select $B = \{0, 1, 2\}$. Then the gray-level opening is

$$f \circ B = \cdots 0\,1\,1\,2\,2\,2\,0\,1\,1\,1\,2\,2\,2\,1\,0 \cdots$$

and the gray-level peak extraction gives the peak image

$$PE(f) = \cdots 0\,1\,0\,0\,1\,2\,0\,3\,3\,0\,0\,1\,0\,0\,0 \cdots.$$

Now the threshold binary images of f are f_a, $1 \leq a \leq 4$:

$$f_4 = \cdots 0\,0\,0\,0\,0\,1\,0\,1\,1\,0\,0\,0\,0\,0\,0 \cdots$$

$$f_3 = \cdots 0\,0\,0\,0\,1\,1\,0\,1\,1\,0\,0\,1\,0\,0\,0 \cdots$$

$$f_2 = \cdots 0\,1\,0\,1\,1\,1\,0\,1\,1\,0\,1\,1\,1\,0\,0 \cdots$$

$$f_1 = \cdots 0\,1\,1\,1\,1\,1\,0\,1\,1\,1\,1\,1\,1\,1\,0 \cdots$$

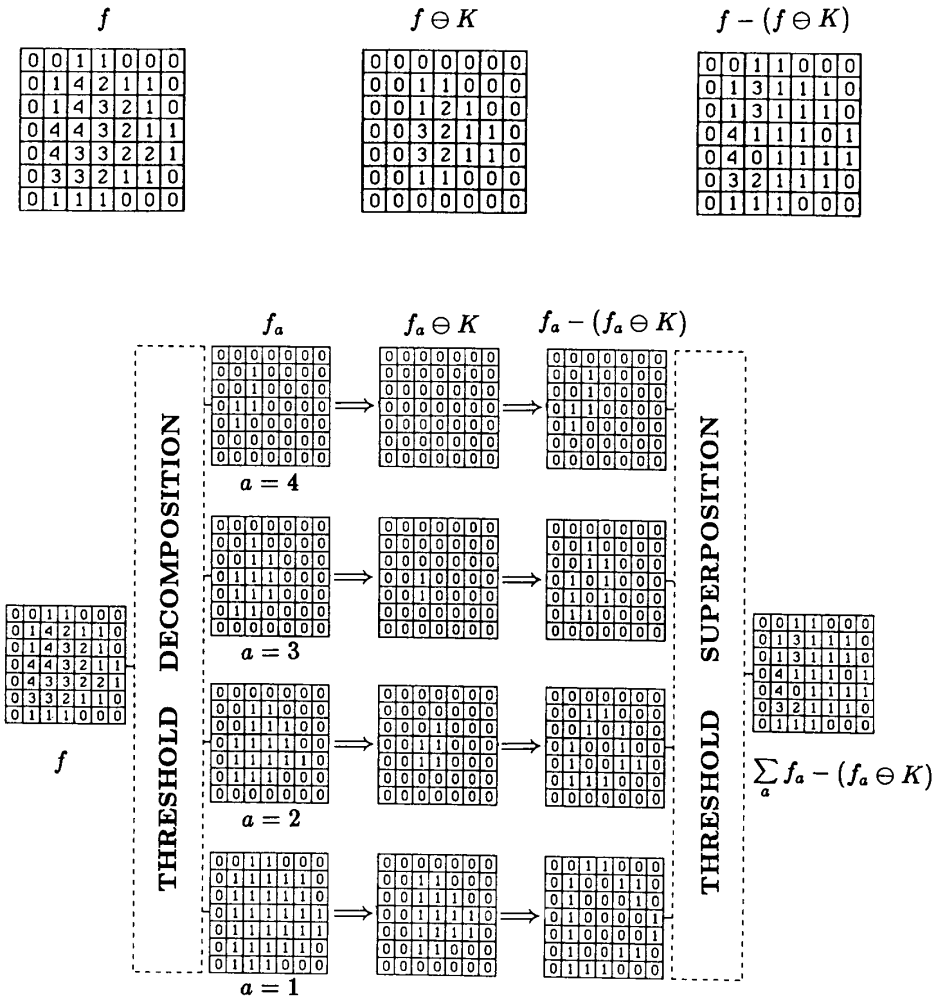

Fig. 1. Threshold superposition of erosion edge gradient. (Original image f has maximum intensity level $A = 4$. The structuring element K is the 5-pixel discrete disk of radius 1.)

and $f_0(m) = 1$ for all m. The binary openings $f_a \circ B$ are

$$f_4 \circ B = \cdots 0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0 \cdots$$
$$f_3 \circ B = \cdots 0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0 \cdots$$
$$f_2 \circ B = \cdots 0\,0\,0\,1\,1\,1\,0\,0\,0\,0\,1\,1\,1\,0\,0 \cdots$$
$$f_1 \circ B = \cdots 0\,1\,1\,1\,1\,1\,0\,1\,1\,1\,1\,1\,1\,1\,0 \cdots .$$

The binary peak extractions $PE(f_a) = f_a - (f_a \circ B)$ are

$$PE(f_4) = \cdots 0\,0\,0\,0\,0\,1\,0\,1\,1\,0\,0\,0\,0\,0\,0\,0 \cdots$$
$$PE(f_3) = \cdots 0\,0\,0\,0\,1\,1\,0\,1\,1\,0\,0\,1\,0\,0\,0 \cdots$$
$$PE(f_2) = \cdots 0\,1\,0\,0\,0\,0\,0\,1\,1\,0\,0\,0\,0\,0\,0 \cdots$$
$$PE(f_1) = \cdots 0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0 \cdots .$$

Thus summing the signals $PE(f_a)$ for all a gives us the original signal $PE(f)$. Clearly, the binary peak extractors are trivial to implement. $PE(f_a)$ simply consists of eliminating from the binary image f_a all connected components that contain any shifted version of B.

If we also consider the *valley extractor* system $VE(f) = (f \bullet B) - f$, by working as above, it can be shown that $VE(f) = \Sigma_a VE(f_a)$.

C. Skeletonization

Next, a morphological skeleton for a gray-level image f is defined by generalizing the algorithm in [25]. If B is a 2-D structuring set, let

$$nB = B \oplus B \oplus \cdots \oplus B \quad (n \text{ times}) \tag{19}$$

denote the n-fold dilation of B with itself, which creates a set of size $n = 0, 1, 2, \cdots$ times larger than B. (If $n = 0$, $nB = \{(0, 0)\}$.) The nth *skeleton component* of f is

$$SK_n(f) = (f \ominus nB) - [(f \ominus nB) \circ B], \quad 0 \le n \le N \tag{20}$$

where $N = \max\{n: f \ominus nB \neq 0\}$. (We assume here images f with a finite support.) These components $SK_n(f)$, indexed by the

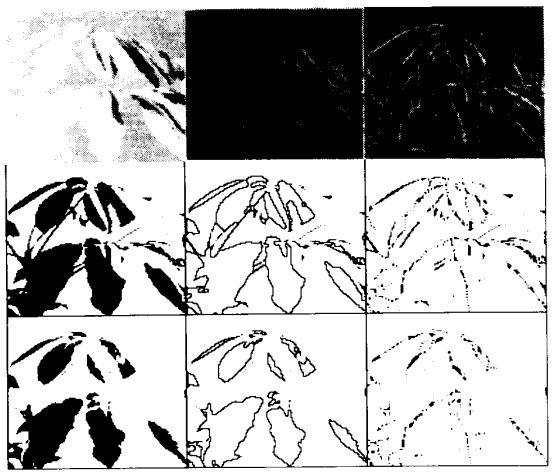

Fig. 2. The top row (from left to right) shows an original gray-level image f of 110 × 128 pixels with 8 bits/pixel, the gray-level erosion edge gradient $f - (f \ominus K)$, and the gray-level skeleton $SK(f)$ with respect to K, where K is a 3 × 3-pixel square structuring element. The other images show (from middle to bottom row): left—threshold binary images f_a for $a = 180$ and 210; middle—their binary edge images $f_a - (f_a \ominus K)$; right—their binary skeletons $SK(f_a)$. (In the top row the edge and skeleton image amplitude has been magnified; in the middle and bottom rows, the black (white) areas correspond to image foreground (background).)

discrete size parameter n, are nonnegative everywhere. Thus they are gray-level images, usually very sparse, and their ensemble can exactly reconstruct f. A skeleton, i.e., a thinned caricature, of f can be defined[2] as the gray-level image

$$SK(f) = \sum_{n=0}^{N} SK_n(f). \tag{21}$$

Since erosions and openings of the binary images f_a by sets B of dimensionality ≤ 2 yield binary outputs and since $f_a \ominus nB \geq (f_a \ominus nB) \circ B$, the skeleton component $SK_n(f_a)$ of f_a is also a binary image. The skeleton $SK(f_a)$ of f_a is defined [22], [21], [26], [13] as the union of all the binary skeleton components $SK_n(f_a)$, represented by 2-D sets. But, for each a, this union-definition of $SK(f_a)$ is equivalent to a sum-definition as in (21) because the binary images $SK_n(f_a)$ are disjoint [13]. Putting all these ideas together yields

$$\begin{aligned} SK_n(f) &= \left[\left(\sum_{a=1}^{A} f_a\right) \ominus nB\right] - \left[\left(\sum_{a=1}^{A} f_a\right) \ominus nB\right] \circ B \\ &= \left(\sum_{a=1}^{A} f_a \ominus nB\right) - \left[\sum_{a=1}^{A} (f_a \ominus nB) \circ B\right] \\ &= \sum_{a=1}^{A} \left[(f_a \ominus nB) - (f_a \ominus nB) \circ B\right] \\ &= \sum_{a=1}^{A} SK_n(f_a). \end{aligned} \tag{22}$$

[2]In [21, pp. 94–95] the idea appears to create a gray-level image gradient or skeleton by adding the binary contours or skeletons, respectively, of threshold binary images; but this was not related to any system producing gradients or skeletons by operating directly on gray-level images.

Thus, the nth skeleton component system obeys the threshold-linear superposition. Now,

$$\begin{aligned} SK(f) &= \sum_{n=0}^{N} SK_n\left(\sum_{a=1}^{A} f_a\right) \\ &= \sum_{n=0}^{N} \sum_{a=1}^{A} SK_n(f_a) = \sum_{a=1}^{A} \sum_{n=0}^{N} SK_n(f_a) \\ &= \sum_{a=1}^{A} SK(f_a). \end{aligned} \tag{23}$$

Hence, the morphological skeleton system also obeys threshold-linear superposition. An example is given in Fig. 2.

D. Shape-Size Distributions

Matheron [16], [17] introduced the *granulometries* as openings $X \circ rB$ of sets X in $\boldsymbol{R}^d$ by one-parameter (r) families of compact convex structuring elements $rB = \{rb: b \in B\}$. He also used Lebesgue measures of these openings to define probabilistic *size distributions* of random sets. As described in [26, ch. 10], Serra and his co-workers have used extensively these set size distributions in image analysis applications to petrography and biology. In [20] closings by hexagons of increasing size were used to analyze digital gray-level biomedical images. In [12] some generalizations and extensions of these size distributions were developed and applied to multiscale shape representation and description.

For a gray-level image $f(i, j)$, consider the (differential size distribution) function [12]

$$\begin{aligned} [SH(f)](+n, B) &= \sum_i \sum_j [f \circ nB - f \circ (n+1)B](i, j), \\ &\qquad n \geq 0 \\ [SH(f)](-n, B) &= \sum_i \sum_j [f \bullet nB - f \bullet (n-1)B](i, j), \\ &\qquad n > 0 \end{aligned} \tag{24}$$

where nB is given by (19); i.e., the integer n is a discrete size parameter and B is *any* 2-D structuring set whose shape can vary. The function $SH(f)$ can be viewed as a *shape-size histogram*. In [12] it was called "pattern spectrum" as a symbolic term to illustrate some of its conceptual similarities with classical Fourier spectrum as well as its close relationships with morphological skeleton transforms for shape representation. It measures the size (n) and shape (B) distribution of f, providing useful information about critical scales and the general shape-size content of f. In [3] the normalized pattern spectrum for positive n (called "pecstrum") was used for shape recognition.

We emphasize here that the function (24) is *nonnegative* for any f and *any* set B. To check this conjecture (proven in [12]), observe that for all $n \geq 0$

$$\begin{aligned} f \circ (n+1)B &= \big(((f \ominus nB) \ominus B) \oplus B\big) \oplus nB \\ &= \big((f \ominus nB) \circ B\big) \oplus nB \\ &\leq (f \ominus nB) \oplus nB = f \circ nB. \end{aligned}$$

Similarly we can show that $f \bullet (n+1)B \geq f \bullet nB$ for all $n \geq 0$.

Regarding threshold superposition, for $n \geq 0$,

$$\begin{aligned} &[SH(f)](n, B) \\ &= \sum_i \sum_j (f \circ nB)(i, j) - \sum_i \sum_j [f \circ (n+1)B](i, j) \\ &= \sum_i \sum_j \left[\left(\sum_{a=1}^{A} f_a\right) \circ nB\right](i, j) \\ &\quad - \sum_i \sum_j \left[\left(\sum_{a=1}^{A} f_a\right) \circ (n+1)B\right](i, j) \end{aligned}$$

$$= \sum_i \sum_j \sum_a [f_a \circ nB](i,j) - \sum_i \sum_j \sum_a [f_a \circ (n+1)B](i,j)$$

$$= \sum_a \left[\sum_i \sum_j [f_a \circ nB - f_a \circ (n+1)B](i,j) \right]$$

$$= \sum_{a=1}^{A} [SH(f_a)](n, B). \tag{25}$$

An identical result to (25) is easily obtained for $n < 0$ by replacing openings $f \circ nB$ with closings $f \bullet nB$. Thus the shape-size histogram obeys the threshold-linear superposition. To illustrate this, consider the example of the 1-D image f in Section III-B. Fixing $B = \{0, 1\}$ yields

n	-2	-1	0	1	2	3	4	5	6
$[SH(f)](n)$	2	6	3	8	6	0	5	0	7
$[SH(f_4)](n)$	0	1	1	2	0	0	0	0	0
$[SH(f_3)](n)$	2	1	1	4	0	0	0	0	0
$[SH(f_2)](n)$	0	3	1	2	6	0	0	0	0
$[SH(f_1)](n)$	0	1	0	0	0	0	5	0	7

Computing the size histograms of the binary images f_a is much easier than for f. For example, for 1-D images f and $B = \{0, 1\}$, the value of $SH(f_a)$ at $(n-1)$ is equal to n times the number of *runs* of n consecutive 1's if $n \geq 1$; likewise for runs of 0's and negative n.

Note that the shape-size histogram system performs an image measurement, while the three previous morphological systems examined in Sections III-A, B, and C perform an image transformation.

IV. General Result

The four morphological systems of Section III, which we showed that obeyed the threshold linear superposition, consisted of pointwise additions/subtractions of simple morphological operations. Next we show that these four examples are special cases of a more general result. Let F be the class of all d-dimensional ($d = 1, 2, \cdots$) *nonnegative* real-valued signals f: $E \to V$ (not necessarily images), where $E = \boldsymbol{R}^d$ or $E = \boldsymbol{Z}^d$. ($\boldsymbol{R}$ and $\boldsymbol{Z}$ denote, respectively, the set of reals and integers.) The input signals' range V could be either discrete ($V = \{n \in \boldsymbol{Z}: n \geq 0\}$) or continuous ($V = \{r \in \boldsymbol{R}: r \geq 0\}$). If V is discrete, f is reconstructed from its thresholds f_a by summation

$$f(x) = \sum_{a=1}^{\infty} f_a(x), \qquad \forall x \in E,$$

and a system Ψ with input f obeys the threshold-linear superposition if (11) is satisfied. In the case of continuous range $V = [0, +\infty)$, f can be reconstructed from the f_a's via integration

$$f(x) = \int_0^{\infty} f_a(x)\, da, \qquad \forall x \in E,$$

and we say that a system Ψ with input f obeys the threshold-linear superposition if

$$\Psi(f) = \int_0^{\infty} \Psi(f_a)\, da.$$

Let S be the class of all systems Ψ: $F \to G$ that obey the threshold-linear superposition, with the restriction that either all $\Psi \in S$ are signal transformations or all are signal measurements but *not* both. G is the class of system outputs, which are either real-valued signals like the signals in F (but not necessarily nonnegative) or real-valued measurements (constants or functions of several parameters). Let us view each system Ψ in S as a vector point. Then let us define a binary operation $\Psi_1 + \Psi_2$ called *system (vector) addition* between any $\Psi_1, \Psi_2 \in S$ and a unary operation $r \cdot \Psi$ called *scalar multiplication of a system* Ψ by any real number $r \in \boldsymbol{R}$ as follows:

$$[\Psi_1 + \Psi_2](f) \stackrel{\text{def}}{=} \Psi_1(f) + \Psi_2(f), \qquad f \in F \tag{26}$$

$$[r \cdot \Psi](f) \stackrel{\text{def}}{=} r \cdot \Psi(f), \qquad f \in F. \tag{27}$$

There is a different interpretation of the symbols + and · between the left and right parts of these definitions. In the right part of (26) "+" denotes pointwise addition of signals if S is a class of signal transformations or addition of real numbers if S is a class of signal measurements. In the right part of (27) "·" denotes multiplication of the signal or measurement $\Psi(f)$ by the scalar r. Thus $\Psi_1 + \Psi_2$ is a *parallel* interconnection of the systems Ψ_1 and Ψ_2, whereas $r \cdot \Psi$ just scales Ψ by r.

Theorem A: The class S of systems Ψ that obey the threshold-linear superposition forms a vector space over the field of real numbers under the vector addition (26) and scalar multiplication (27).

Proof: From [9], we must prove that, for all $\Psi, \Phi \in S$ and $r, q \in \boldsymbol{R}$,

A1: $(S, +)$ is an Abelian group.
A2: $r \cdot \Psi \in S$.
A3: $r \cdot (\Psi + \Phi) = r \cdot \Psi + r \cdot \Phi$.
A4: $(r + q) \cdot \Psi = r \cdot \Psi + q \cdot \Psi$.
A5: $r \cdot (q \cdot \Psi) = (rq) \cdot \Psi$.
A6: $1 \cdot \Psi = \Psi$.

Assume first a discrete range V for input signals. (A1): S is closed under system + because

$$[\Psi + \Phi](f) = \Psi(f) + \Phi(f) = \sum_a \Psi(f_a) + \sum_a \Phi(f_a)$$

$$= \sum_a [\Phi + \Psi](f_a). \tag{28}$$

Further, the system + is associative, commutative, and has a *zero* element (the system Ψ_0, where $\Psi_0(f)$ is the all-zero signal for all f or just zero in case of signal measurements). Finally each Ψ has its *inverse* $-\Psi$, defined as $[-\Psi](f) = -\Psi(f)$. Hence, $(S, +)$ is an Abelian group. (A2) is true because

$$[r \cdot \Psi](f) = r \cdot \Psi(f) = r \cdot \sum_a \Psi(f_a)$$

$$= \sum_a r \cdot \Psi(f_a) = \sum_a [r \cdot \Psi](f_a). \tag{29}$$

The proof of the rest of the axioms (A3)–(A6) is easy and hence omitted.

If V is continuous, replacing everywhere in the above proof summation $\Sigma_a(\cdot)$ with integration $\int(\cdot)\, da$ yields again that S is a vector space. Q.E.D.

The above result establishes that the principle of threshold-linear superposition is obeyed by any composite system formed as a linear combination of systems that obey it. As a special case consider any systems Ψ_k among the following classes used often in image analysis: 1) erosion, dilation, opening, and closing by binary (flat) structuring elements; 2) median, rank-order, and stack filters; and 3) any finite cascade or parallel (i.e., using pointwise max/min) combination of systems in 1) and/or 2). All such Ψ_k commute with thresholding and hence obey threshold superposition. Hence, Theorem A implies that any linear finite combination system $\Psi(f) = \Sigma_k r_k \Psi_k(f)$ will also obey (11). Such a useful linear combination is, for instance, the general order statistic filter [2] where the systems Ψ_k are all the distinct kth rank-order filters.

Next a few remarks follow about our general result:

1) The class of systems obeying threshold-linear superposition contains the class of all linear systems, because threshold-linear superposition is a weak form of linear superposition.

2) Theorem A is independent of:

- whether or not the systems are translation-invariant;
- the signals' dimensionality;
- whether the signals have continuous or discrete argument(s);
- whether the input or output signals have a discrete or continuous range.

3) Theorem A requires that the *input* signals are *nonnegative*; hence, it especially applies to image analysis systems.

4) The results for the four *discrete* morphological systems (edge gradients, peak detection, skeletonization, size distributions) of Section III follow now as simple corollaries of Theorem A. However, some *continuous* versions of these four systems may or may not fall in the category of systems to which Theorem A applies, as explained next.

We shall briefly address some issues on whether our result about the morphological edge gradient can be extended to the Euclidean case. (The initial question and the following (30)–(34) have been suggested by Serra [28].) Consider signals f: $\boldsymbol{R}^d \to [0, +\infty)$, their threshold binary signals f_a and the corresponding threshold sets $F_a = \{x \in \boldsymbol{R}^d: f_a(x) = 1\}$. The *boundary* of F_a is given by

$$\partial F_a = \bar{F}_a \backslash \mathring{F}_a = \bar{F}_a \cap (\mathring{F}_a)^c, \tag{30}$$

where $\bar{X}$, $\mathring{X}$, X^c denote topological closure, topological interior, and the complement of a set X. Now, if rB is the ball of radius r, then [16, p. 150]

$$\mathring{X} = \bigcup_{r>0} X \ominus rB \tag{31}$$

$$\bar{X} = \bigcap_{r>0} X \oplus rB. \tag{32}$$

Hence, it follows that

$$\partial F_a = \bigcap_{r>0} (F_a \oplus rB) \backslash (F_a \ominus rB). \tag{33}$$

Then we could define a morphological edge gradient for the binary signal f_a by

$$[MG(f_a)](x) = \inf_{r>0} [(f_a \oplus rB) - (f_a \ominus rB)](x), \tag{34}$$

which is equal to the indicator function of ∂F_a. Now, since $f_a \oplus rB \le f_a \oplus sB$ and $f_a \ominus rB \ge f_a \ominus sB$ for $r < s$,

$$[MG(f_a)](x) = \lim_{r \downarrow 0} [(f_a \oplus rB) - (f_a \ominus rB)](x). \tag{35}$$

Integrating (35) over a yields

$$\int_0^\infty [MG(f_a)](x)\, da = \lim_{r \downarrow 0} [(f \oplus rB) - (f \ominus rB)](x) \tag{36}$$

which does not behave like a gradient; it actually gives zero response when f is continuous at x. However, Beucher's morphological gradient in $\boldsymbol{R}^d$ is defined [1], [26, p. 441] as

$$[MG(f)](x) = \lim_{r \to 0} \frac{[(f \oplus rB) - (f \ominus rB)](x)}{2r} \tag{37}$$

which is not equal to the right side of (36). In addition, the definition (37) applied to the binary signals f_a may not give an edge function equal to (35).

Similarly, there may arise several problems if one attempts to extend our result on threshold superposition of the skeleton and the size distributions to their corresponding systems in the Euclidean space. (For properties of the skeleton in Euclidean spaces see Matheron [27, ch. 11].)

V. Concluding Remarks

An important factor on which our results in Section III depend is that the erosions, dilations, openings, and closings used by the four analyzed morphological systems involve flat (binary) structuring elements. That is, for a 2-D image signal, only 2-D or 1-D sets can be used as structuring elements; likewise, for a 1-D signal, the structuring element must be a 1-D set. For the more general erosions (min of differences), dilations (max of additions), and their combinations, which use a nonbinary structuring element [30], [26], [14], [6], our results in this correspondence do not apply. Some forms of threshold superposition for such general erosions/dilations have been investigated in [26, pp. 443–444] and in [29].

The key idea of our results is that a large class of morphological and other systems for gray-level image analysis reduces to corresponding systems for binary signals. But the latter are much easier to analyze. Hence our results provide a theoretical tool that facilitates the analysis of many morphological and related nonlinear systems. In addition, they suggest new implementations based on threshold superposition. Of course, *software* implementations of these ideas on current serial computer architectures are discouraging because of the large number of threshold binary images required. However, VLSI hardware implementations exploiting the threshold superposition of composite gray-level morphological systems (as already has been done for simple rank-order and morphological operations [7], [24], [8]) is very promising because binary morphological operations can be done using only pixel counting or Boolean logic. Further, the binary operations on each threshold binary image can be done in *parallel* for all threshold levels.

Acknowledgment

We wish to thank J. Serra from École des Mines de Paris, at Fontainebleau, for many helpful suggestions and detailed comments.

References

[1] S. Beucher and T. Hersant, "Detection du relief de fractures metalliques par correlation d'images," in *Proc. Int. Symp. Quantit. Metallography*, Florence, Italy, Nov. 1978.

[2] A. C. Bovik, "A generalization of median filtering using linear combinations of order statistics," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 1342–1349, Dec. 1983.

[3] J. F. Bronskill and A. N. Venetsanopoulos, "Multidimensional shape description and recognition using mathematical morphology," *J. Intell. Robotic Syst.*, vol. 1, pp. 117–143, 1988.

[4] J. P. Fitch, E. J. Coyle, and N. C. Gallagher, "Median filtering threshold decomposition," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-32, pp. 1183–1188, Dec. 1984.

[5] V. Goetcherian, "From binary to grey tone image processing using fuzzy logic concepts," *Pattern Recogn.*, vol. 12, pp. 7–15, 1980.

[6] R. M. Haralick, S. R. Sternberg, and X. Zhuang, "Image analysis using mathematical morphology," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-9, pp. 523–550, July 1987.

[7] R. G. Harber, S. C. Bass, and G. W. Neudeck, "VLSI implementation of a fast rank order filtering algorithm," in *Proc. IEEE ICASSP-85*, Tampa, FL, Mar. 1985.

[8] J. M. Hereford and W. T. Rhodes, "Nonlinear optical image filtering by time-sequential threshold decomposition," *Opt. Eng.*, vol. 27, pp. 274–279, Apr. 1988.

[9] I. N. Herstein, *Topics in Algebra*. New York: Wiley, 1975.

[10] J. S. J. Lee, R. M. Haralick, and L. G. Shapiro, "Morphologic edge detection," *IEEE Trans. Robotics Automat.*, vol. RA-3, pp. 142–156, Apr. 1987.

[11] R. M. Lougheed, D. L. McCubbrey, and S. R. Sternberg, "Cytocomputers: Architectures for parallel image processing," in *Proc. Workshop Picture Data Descr. Management*, Pacific Grove, CA, 1980.

[12] P. Maragos, "Pattern spectrum and multi-scale shape representation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 701–716, July 1989.

[13] P. Maragos and R. W. Schafer, "Morphological skeleton representation and coding of binary images," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 1228–1244, Oct. 1986.

[14] ——, "Morphological filters—Part I: Their set-theoretic analysis and relations to linear shift-invariant filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 1153–1169, Aug. 1987.
[15] ——, "Morphological filters—Part II: Their relations to median, order-statistic, and stack filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 1170–1184, Aug. 1987; corrections, vol. 37, p. 597, Apr. 1989.
[16] G. Matheron, *Éléments pour une Théorie des Milieux Poreux*. Paris: Masson, 1967.
[17] ——, *Random Sets and Integral Geometry*. New York: Wiley, 1975.
[18] F. Meyer, "Contrast feature extraction," in *Special Issues of Practical Metallography*. Stuttgart, West Germany: Riederer Verlag GmbH, 1978; see also *Proc. 2nd European Symp. Quantitative Analysis of Microstruct. in Materials Science, Biology, and Medicine*, France, Oct. 1977.
[19] ——, "Iterative image transformations for an automatic screening of cervical smears," *J. Histochem. Cytochem.*, vol. 27, pp. 128–135, 1979.
[20] ——, "Empiricism and idealism," in *Pattern Recognition in Practice*, E. S. Gelsema and L. N. Kanal, Eds. Amsterdam, The Netherlands: North-Holland, 1980.
[21] P. E. Miller, "An investigation of boolean image neighborhood transformations," Ph.D. dissertation, Ohio State Univ., 1978.
[22] J. C. Mott-Smith, "Medial axis transformations," in *Picture Processing and Psychopictorics*, B. S. Lipkin and A. Rosenfeld, Eds. New York: Academic, 1970.
[23] Y. Nakagawa and A. Rosenfeld, "A note on the use of local min and max operations in digital picture processing," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-8, Aug. 1978.
[24] E. Ochoa, J. P. Allebach, and D. W. Sweeney, "Optical median filtering by threshold decomposition," *Appl. Opt.*, vol. 26, pp. 252–260, Jan. 1987.
[25] S. Peleg and A. Rosenfeld, "A min-max medial axis transformation," *IEEE Trans. Pattern. Anal. Machine Intell.*, vol. PAMI-3, pp. 208–210, Mar. 1981.
[26] J. Serra, *Image Analysis and Mathematical Morphology*. New York: Academic, 1982.
[27] ——, Ed., *Image Analysis and Mathematical Morphology Vol. 2: Theoretical Advances*. New York: Academic, 1988.
[28] ——, personal communication, 1989.
[29] F. Y.-C. Shih and O. R. Mitchell, "Threshold decomposition of gray-scale morphology into binary morphology," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 31–42, Jan. 1989.
[30] S. R. Sternberg, "Grayscale morphology," *Comput. Vision, Graphics, Image Processing*, vol. 35, pp. 333–355, 1986.
[31] L. J. van Vliet, I. T. Young, and G. L. Beckers, "A nonlinear Laplace operator as edge operator in noisy images," *Comput. Vision, Graphics, Image Processing*, vol. 45, pp. 167–195, 1989.
[32] P. D. Wendt, E. J. Coyle, and N. C. Gallagher, "Stack filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 898–911, Aug. 1986.

Correspondenceless Stereo and Motion: Planar Surfaces

JOHN (YIANNIS) ALOIMONOS AND JEAN-YVES HERVÉ

***Abstract*—We show that a binocular observer can recover the depth and three-dimensional motion of a rigid planar patch, without using any correspondences between the left and right image frames (static) or between the successive dynamic frames (dynamic). We study uniqueness and robustness issues with respect to this problem and we provide experimental results from the application of our theory to real images.**

***Index Terms*—Correspondence problem, stereo, visual motion.**

Manuscript received May 4, 1988; revised December 13, 1989. Recommended for acceptance by W. B. Thompson.

The authors are with the Computer Vision Laboratory, Center for Automation Research, University of Maryland, College Park, MD 20742.

IEEE Log Number 9034376.

I. Introduction

An important problem in computer vision is to recover the depth and 3-D motion of a moving object from its successive images. Dynamic visual information can be produced by a sensor moving through the environment and/or by independently moving objects in the observer's visual field. The interpretation of such dynamic imagery information consists of dynamic segmentation, recovery of the 3-D motion (of the sensor and the objects in the environment), as well as determination of the structure of the environment. The results of such an interpetation can be used to control behavior, as for example in robotics, tracking, and autonomous navigation. Up to now there have been, basically, three approaches towards the solution of this problem: 1) the differential or continuous; 2) the discrete; and 3) the direct.

In the differential approach, the dynamic image is supposed to be a 3-D function of two "spatial" arguments and a "temporal" argument. Then, if this function is locally well-behaved and its spatio-temporal gradients are computable, the local image velocity or "optical flow" may be computed [7], [9], [10].

The second approach considers the case where the motion is large and the first technique is not applicable. In these instances, the measurement of the image motion relies upon isolating and tracking highlights and feature points (corners, edges, etc.) in the image through time. This entails solving the correspondence problem which has proved to be difficult in many situations [3], [6], [21], [23].

In both of the above approaches, after the image motion (optical flow or discrete displacements) is computed, constraints are developed between the retinal motion and the 3-D motion parameters; these constraints become the basis of a whole variety of algorithms for the recovery of the 3-D motion [1], [4], [5], [8], [18], [19], [24], [25], [27]–[38].

The third approach which is worth mentioning [2], [15], [22], [32], [33] attempts to directly compute 3-D motion from the spatiotemporal derivatives of the image intensity function. There is good research in this direction but a stability analysis of the methods is still needed [31].

As the problem has been formulated over the years, usually one camera is used and so the 3-D motion parameters that can be computed are five: two for the direction of translation and three for the rotation. In our approach, we assume a binocular observer and so we can recover six motion parameters: three for the translation and three for the rotation. In this correspondence we restrict our work to the motion of a rigid planar patch. In the traditional one-camera approach for the estimation of the 3-D motion parameters of a rigid planar patch, it was recently pointed out [26] that one should use image point correspondences for object points not on a single planar patch when estimating 3-D motions of rigid objects. But it was not known how many solutions there are, what the minimum number of points and views needed to assure uniqueness is, and how those solutions can be computed without using any iterative search (i.e., without having to solve nonlinear systems). It was proved [27], [28], [30] that there are exactly two solutions for the 3-D motion parameters and plane orientations, given at least four image point correspondences in two perspective views, with the exception of a few degenerate cases. However, the solutions are unique if three views of a planar patch are given or two views of at least two planar patches. In our approach, the duality problem does not exist in

0162-8828/90/0500-0504$01.00 © 1990 IEEE