THE HOLE SPECTRUM -- MODEL-BASED OPTIMIZATION
OF MORPHOLOGICAL FILTERS

Edward R. Dougherty
Center for Imaging Science
Rochester Institute of Technology

Robert M. Haralick
Department of Electrical Engineering
University of Washington

ABSTRACT

An image representation is developed that is useful for designing optimal morphological filters to restore images suffering from degradation due to subtractive noise. The image and noise models are predicated on the existence of some class of shape primitives into which both image and noise can be decomposed (relative to union). Both deterministic and nondeterministic cases are considered, and in each case the relevant model constraints are fully discussed. The type of filters that are naturally compatible with the image-noise models are analyzed, and the relation to general mean-square morphological optimization is explored.

1. INTRODUCTION

In the present paper we address the problem of finding an image representation that is applicable to restoration of binary images degraded by subtractive noise. Our intent is provide a paradigm for model-based restoration that is compatible with the very general statistical characterization of optimal mean-square morphological filters that has been given by Dougherty [1, 2]. The latter is applicable to those image processing problems that are concerned with estimation, for instance, restoration and compression; however, just as the optimal linear filter can best be employed by relating it to frequency content, as is done by the Wiener filter, there remains to be found relevant image decompositions that facilitate the discovery of optimal morphological filters. Such an approach was considered in [3] where we employed an opening decomposition to discover an optimal filter for a certain image-noise model. There, however, while the Wiener analogy held quite well, the resulting filters did not fit the framework of [1]. Specifically, the resulting optimal filters were not necessarily monotonic. The present model will fit the framework, and it will lead to a certain class of potential filters that are closely related to the image and the noise. We begin by presenting a general algebraic formulation of component-pass type filtering that will serve to frame the model to be subsequently discussed. We refer the reader to [3] for a general algebraic analysis of the Wiener approach (that could also be utilized herein).

2. AN ALGEBRAIC PARADIGM FOR COMPONENT FILTERING

We require a class $S$ of signals and with each signal $f$ in $S$ there must be an associated n-tuple $\Phi(f) = [f_k]$, called the spectrum of $f$, whose components come from some set $C$ of objects (which need not be numbers). This association forms a mapping

$$\Phi: f \rightarrow \Phi(f) = [f_1, f_2, ..., f_n]$$

(1)
from $S$ into $C^n$, and, since $[f_k]$ must serve as a representation of $f$, we demand there exist an inverse mapping $\Phi^{-1}$.

For any two signals $f$ and $g$ in $S$, there is defined a binary operation $\ll$, not necessarily commutative, such that $f \ll g$ need not lie in $S$. If $f \ll g$ lies in $S$, then we say $g$ is weakly compatible with $f$. Weak compatibility takes the place of linear-space closure, but more is required: some stronger condition must serve in lieu of basis-coefficient linearity.

To proceed further, we assume there exists an operation $\circ$ in $C$ and for any two elements in $C^n$ we induce $\circ$ componentwise by

$$[a_k] \circ [b_k] = [a_k \circ b_k] \quad (2)$$

We assume there exists a null element $0$ for $\circ$, so that $a \circ 0 = a$. We say $g$ is strongly compatible with $f$ if it is weakly compatible and

$$\Phi(f \ll g) = [f_k] \circ [g_k] \quad (3)$$

$S<f>$ denotes the class of strongly $f$-compatible elements of $S$.

We now have sufficient structure to provide an algebraic characterization of component filtering. First suppose $g$ is strongly $f$-compatible and that $f$ and $g$ do not share nonnull components, i.e., that the signal and noise components are separated. By rearranging the components, if necessary, there exists an index $q$ such that $f_k = 0$ for $k > q$ and $g_k = 0$ for $k \leq q$ (and note that we could have interchanged the ordering so that $f_k$ is null for small $k$ and $g_k$ is null for large $k$). We define the function $Q$ on $C^n$ by

$$Q([a_1, a_2, ..., a_n]) = [a_1, a_2, ..., a_q, 0, 0, ..., 0] \quad (4)$$

Assuming $f \ll g$ to be a degradation of $f$, total restoration is achieved by the mapping $\Phi^{-1}Q\Phi$.

Rather than filter in the component domain, we might, as with convolution and frequency, desire a spatial-domain filter that accomplishes the same task as $Q$. Such a filter would be a mapping $\Psi$ that completes the commutative diagram

$$\Phi$$

$$f \ll g \rightarrow [f_k \circ g_k]$$

$$\Psi \downarrow \quad \downarrow Q$$

$$f \rightarrow [f_k]$$

$$\Phi$$

that is, $\Psi = \Phi^{-1}Q\Phi$. From a strictly algebraic perspective, $\Psi$ always exists as $\Phi^{-1}Q\Phi$; however, what we really desire is that $\Psi$ come from some class of filters. Specifically, we require $\Psi$ to be a morphological filter of some given kind. Even with signal-noise separation, $\Psi$ in the desired form might not be obtainable exactly, so we will have to make do with some approximate
completion of the diagram.

More generally, the nonnull components of $f$ and $g$ are not disjoint. Thus, the mapping $Q$ of equation (4) does not exist (as formulated). Instead it takes the form $Q = [q_k]$ and filtering is accomplished by $Q(a_k) = q_k(a_k)$, where $q_k$ is the $k$th component function of $Q$. This is analogous to the frequency form of the Wiener filter, in which case $q_k(a_k) = w_k a_k$, where $w_k$ is the weight and $a_k$ is the frequency component. To measure the goodness of $Q$, there must exist some error measure $e$ between signals. Assuming signals to be random functions, goodness of $Q$ is measured by the expected value $E[e(f, Q^{-1} Q(f \neq g))]$.

3. THE HOLE SPECTRUM

Our goal is the development of a generalized spectral (canonical) representation that models holes in an image in a manner conducive to filling noise-created holes by a union of erosions. We begin by postulating a list of primitive shapes that will be assumed to generate both noise and generic holes. We consider a family of images, called shape primitives: $P = \{N_1, N_2, \ldots, N_n\}$. Each $N_j$ is assumed to be path-connected. If $S$ is any image, $<S>$ denotes the minimal rectangle containing $S$. $S$ is said to be $P$-representable if there exist points $x_{ij}$ such that

\[ i ) \quad S = <S> - \cup \{N_i + x_{ij}: i = 1, 2, \ldots, n; j = 1, 2, \ldots, h(i)\} \]

\[ i i ) \quad \text{If } C \text{ is any connected component of } <S> - S, \text{ then there exists a unique pair } (i, j) \text{ in the index set of the union such that } C = N_i + x_{ij}. \text{ Moreover, for any other index pair } (r, s), N_r + x_{rs} \text{ is disjoint from } C. \]

Condition (ii) imposes a minimality condition on the union in (i), thereby making the representation of (i) unique. $S$ is the class of $P$-representable images, and henceforth we assume all images to be $P$-representable. A hole is any connected component of $<S> - S$. It need not be enclosed by pixels in $S$; rather, it may be a penetration of the minimal containing rectangle $<S>$.

We postulate the existence of a subclass of $P$, $N = \{N_1, N_2, \ldots, N_p\}$, $1 \leq p \leq n$. The purpose of $N$ is to be a subfamily of shape primitives that serve as noise generators; specifically, we limit the noise holes of $S$ to be those holes in $<S>$ created by terms in (i) of the form $N_i + x_{ij}$, where $N_i$ lies in $N$. It might be that there is no limitation and $p = n$. We call $N$ the noise class. Note that generic holes, those belonging to $S$ properly and not having been created by noise, can arise from shapes in $N$.

For each $j$, we assume there exists some enclosure set $M_i$ containing $N_i$. The purpose of $M_i$ is to form a border $M_i - N_i$, so that the filling of a hole created by $N_i$ can be examined in terms of erosions applied to the border. The extent to which $M_i$ is greater than $N_i$ is a modeling question having to do with the nature of generic holes and constraints on the noise.

The hole spectrum of a $P$-representable image is defined by

\[ H[S] = \{[x_{1,1}, x_{1,2}, \ldots, x_{1,h(1)}], \ldots, [x_{n,1}, x_{n,2}, \ldots, x_{n,h(n)}]\} \]
where the \( x_{ij} \) are the translational points in the representation (i). \( H[S] \) and \( \langle S \rangle \) uniquely characterize \( S \). The n-tuple

\[
h[S] = [h(1), h(2), ..., h(n)]
\]

is called the **hole-amplitude spectrum (HAS)** of \( S \). It gives the number of times each shape primitive \( N_i \) is translated to form the representation (i). The set \( \langle S \rangle - S \) consists of "hole-like" components, is given by the union representation in (i), and is characterized by the hole spectrum.

Our filtering concern is with the image \( S \) corrupted by subtractive noise, i.e., new holes are created by the noise. Assuming all shape primitives lie in \( P \), a noise image \( N \) is postulated to be of the form \( N = \langle S \rangle - N^- \), where

\[
i ii) \quad N^- = U \{N_i + z_{ij}; i = 1, 2, ..., p; j = 1, 2, ..., t(i)\}
\]

The noise-corrupted image is defined to be \( S \cap N \). Since

\[
S \cap N = S - N^-
\]

the noisy image results from subtracting translated noise patterns from the uncorrupted image.

We wish to view \( S \cap N \) as the noisy version of \( S \); however, we are faced with a serious hurdle: we must be assured that \( S \cap N \) is \( P \)-representable. Thus, we must put constraints on the union forming \( N^- \). Put crudely, we must be careful that the intersection does not create holes that are not fit tightly by single translates of the \( N_i \), a condition required for \( P \)-representation. The problem here is the lack of a vector-space closure-like property. A practical way around the difficulty is to define a noise image of the type in (iii) to be **weakly compatible** with \( S \) if \( \langle S \cap N \rangle = \langle S \rangle \) and \( S \cap N \) is \( P \)-representable, and to limit our formal analysis to weakly \( S \)-compatible noise images. Relative to the algebraic paradigm of the preceding section, intersection corresponds to the abstract operation \( \langle \rangle \), with \( S \) being the \( P \)-representable images. The purpose for the requirement \( \langle S \cap N \rangle = \langle S \rangle \), which says that the minimal enclosing rectangle is not diminished, is to avoid reformulation of the representation (i), which might be required to maintain the uniqueness of the hole-spectrum representation if \( \langle S \cap N \rangle \) were a proper subset of \( \langle S \rangle \).

Even if \( N \) is weakly \( S \)-compatible, there is another potential problem with our representation theory. In subtracting \( N^- \) from \( S \) we may create **interaction holes**, these being holes created by enlarging generic holes during noise intersection. To appreciate the undesirability of such holes, consider the addition of a signal with noise in the frequency setting. Owing to the linearity of the basis coefficients, the spectrum of the sum is obtained by adding basis coefficients componentwise. We would like to have an analogous process with regard to hole spectra: basis coefficients for \( S \cap N \) should be formed by unioning signal and noise coefficients. Relative to the algebraic paradigm, we would like equation (3) to hold with union in place of \( \circ \). The following conditions on the noise will guarantee both weak compatibility and the type of noisy-image spectra we desire:

\[
i v) \quad \text{For each } N_i \text{ in } N, \text{ the border } M_i - N_i \text{ is strongly path-connected.}
\]
If \( N_i + z_{ij} \) and \( N_r + z_{rs} \) are translates in the union forming \( N^- \), then

\[
(N_r + z_{rs}) \cap (M_i + z_{ij}) = \emptyset
\]

v i) For each translate \( N_i + z_{ij} \) forming \( N^- \), \( M_i + z_{ij} < S \).

Condition (iv) makes the borders sufficiently impenetrable so that conditions (v) and (vi) accomplish their desired ends. Condition (v) is a separation condition assuring that the P-representability of \( S \cap N \) is not destroyed by overlapping or touching noise holes. Condition (vi) says that we ignore noise-primitive translates that miss \( S \) altogether (certainly realistic) and also that noise enclosures cannot hit the boundary of \( S \), which could destroy the P-representability of \( S \cap N \) or create interaction holes. Note another key consequence of the noise conditions: no translate \( N_r + z_{rs} \) forming \( N \) can "get lost" as a subset of some translate \( N_i + x_{ij} \) in (i), or vice versa. If a noise image of the form given in (iii) satisfies conditions (iv), (v), and (vi), it will be said to conform to \( S \).

Consider the hole spectrum of an image \( S \cap N \) obtained from \( S \) by conformable intersection noise. With \( N^- \) given by (iii), we need a representation of the generic holes. By (i), these holes comprise \( <S> \cap S \), which is of the form

\[
\text{v i i) } <S> \cap S = \bigcup (N_i + y_{ij} : i = 1, 2, ..., n; j = 1, 2, ..., s(i))
\]

Assuming \( S \) is P-representable and \( N \) conforms to \( S \),

\[
S \cap N = [<S> - (<S> - S)] - N^- = <S> \cap S - \bigcup (N_i + y_{ij}) - \bigcup (N_i + z_{ij})
\]

To simplify notation, we let \( Y_i \) and \( Z_j \) denote the pixel sets by which \( N_i \) is respectively translated in the signal and noise representations. Then the hole spectra take the following forms:

\[
H[S] = [Y_1, Y_2, ..., Y_n]
\]

\[
H[N] = [Z_1, Z_2, ..., Z_n]
\]

\[
H[S \cap N] = [Y_1 \cup Z_1, Y_2 \cup Z_2, ..., Y_n \cup Z_n]
\]

where it should be kept in mind that all spectra are relative to \(<S>\). Relative to equation (3), equation (12) can be rewritten as

\[
H[S \cap N] = H[S] \cup H[N]
\]

so that \( N \) is strongly compatible with \( S \). In sum, if \( N \) conforms to \( S \), then \( N \) is strongly compatible with \( S \) and we can proceed with component filtering. Of great practical importance is that the hole-amplitude spectrum components are combined by addition:

\[
h[S \cap N] = [s(1) + t(1), s(2) + t(2), ..., s(n) + t(n)]
\]
Let us now consider the spectral approach to noise filtering. Suppose there exists a separation index $q$ such that $s(i) = 0$ for $i \leq q$ and $t(i) = 0$ for $i > q$, so that the signal is separated from the noise, and the relevant spectra take the forms

$$H[S] = [\sigma, \sigma, \ldots, \sigma, Y_{q+1}, Y_{q+2}, \ldots, Y_n]$$

$$H[N] = [Z_1, Z_2, \ldots, Z_q, \sigma, \sigma, \ldots, \sigma]$$

(15)

In the spectrum domain, the noise is fully filtered by the function

$$Q([X_1, X_2, \ldots, X_n]) = [\sigma, \sigma, \ldots, \sigma, X_{q+1}, X_{q+2}, \ldots, X_n]$$

(16)

since $Q(H[S \cap N]) = H[S]$.

When signal and noise are not separated, we desire a function $Q$ that best restores the hole spectrum of $S$ from $S \cap N$, where bestness is defined relative to some criterion of goodness. Since our desire is restoration in the spatial domain, a good measure of error is the symmetric-difference

$$e[Q] = c[H^{-1}[Q(H[S \cap N])] - S] + c[S - H^{-1}[Q(H[S \cap N])]]$$

(17)

where $c$ denotes the number of pixels in a set. This approach is well-defined because we have assumed an image is reconstructable from its hole spectrum, thereby making $H^{-1}$ well-defined.

4. SPATIAL DOMAIN FILTERS

In analogy to the frequency-spectrum setting, we would like some spatial-domain filter that can, to one degree or another (and in a manner analogous to convolution), accomplish the task of $Q$. In the present setting, we desire a morphological filter $\Psi$ such that $\Psi(S \cap N) = S$. The problem is to find $\Psi$ so as to complete the following commutative diagram:

$$\begin{array}{ccc}
S \cap N & \xrightarrow{H} & H[S \cap N] \\
\Psi \downarrow & & \downarrow Q \\
S & \xrightarrow{H} & H[S]
\end{array}$$

(18)

which corresponds to diagram (5). Whether we are in the separated or nonseparated case, no filter $\Psi$ can be expected to complete the diagram exactly, so we use the measure of goodness

$$e[\Psi] = c[S - \Psi(S \cap N)] + c[\Psi(S \cap N) - S]$$

(19)

which is simply a reformulation of the spectral-restoration error $e[Q]$ of equation (17). Since we cannot expect to complete the diagram exactly, we should not expect $e[\Psi]$ to equal $e[Q]$, although ideally they would agree.

The most general problem, whether or not the noise and signal are separated by some index, is to seek $\Psi$ among all morphological filters; however, based upon our model, we will seek $\Psi$ from
a subclass of filters. We proceed to outline the model-based paradigm suitable to observed images $S \cap N$.

Although we could search for a suitable filter $\Psi$ among all possible morphological filters, this would mitigate the purpose of constructing an appropriate noise model. Thus, along with the class $P$ of shape primitives, we postulate the existence of a class $E = \{E_1, E_2, \ldots, E_m\}$ of structuring elements. Since these elements must restore images suffering subtractive noise degradation, to be of use they must satisfy some conditions that make them suitable for the task at hand. Hence, we require the elements of $E$ to satisfy the following two conditions:

1. For any $N_i$ lying in $N$,
   $$\cup\{(M_i \cap N_i) \cap E_j : j = 1, 2, \ldots, m\} > N_i$$

2. For $j = 1, 2, \ldots, m$, $E_j$ does not contain the origin, and there exist activated pixels $p_1$ and $p_2$ in $E_j$ such that the origin lies between $p_1$ and $p_2$.

Condition (viii) guarantees that holes created by subtracting noise primitives can be filled by eroding by elements of $E$, and condition (ix) assures that eroding $S$ or a subset of $S$ by an element of $E$ yields an output that lies within $<S>$. This latter condition is reasonable since $S \cap N < S$.

For any image $S$ and family $B$ of structuring elements, we define the morphological filter $\Psi_B(S)$ by

$$\Psi_B(S) = S \cup [\cup \{S \cap E_k : E_k \in B\}]$$

Since $S$ is equal to $S$ eroded by the origin, $\Psi_B$ is a union of erosions, and as such it is a morphological filter. There is no basis minimality condition imposed upon $E$; that is, it may be that there is an element of $E$ that is a proper subset of another element of $E$. Thus, if $B$ is selected from $E$, it may not be a basis; if not, simply eliminate redundant structuring elements.

Define an image $S$ to be $B$-closed if $B$ is a family of structuring elements for which $\Psi_B(S) = S$. Whether $S$ is $B$-closed or not, we refer to $\Psi_B(S)$ as the $B$-closure of $S$. E-closed images are of special interest, because if $S \cap N$ is a noisy image, $B$ is a subclass of $E$, and we apply $\Psi_B$, then $\Psi_B(S \cap N) < S$, so that $\Psi_B$ cannot create error pixels outside of $S$ so long as $S$ is E-closed. Note that by (ix), the E-closure of $S$ is a subset of $<S>$, so that $S < \Psi_E(S) < <S>$. The E-closure of $S$ need not be E-closed: $\Psi_B(S)$ is not necessarily equal to $\Psi_B(\Psi_B(S))$; that is, $\Psi_B$ is not generally idempotent. Nonetheless, for any integral power $k$, $\Psi_B^k(S) < <S>$. If we construct $E$ in a manner compatible with filling holes created by noise primitives, filters of the form $\Psi_B, B < E$, comprise a potentially fertile class.

If we employ the symmetric-difference restoration error, then error is basis dependent and is given by

$$e[B] = c[S - \Psi_B(S \cap N)] + c[\Psi_B(S \cap N) - S]$$

Our problem is combinatoric, in that we must search for the best among all subsets of E that are filter bases. As will be subsequently discussed, such a search fits into the general filter methodology developed in [1, 2]. If E has been selected in a manner highly compatible with filling holes created by primitives in N, then we have accomplished a good deal because the general search problem of [1] is dramatically reduced. Relevant search strategies will not be discussed herein, but have been developed and will be presented in subsequent publications.

5. THE NONDETERMINISTIC SETTING

In moving to the nondeterministic setting, we need to reinterpret the foundations of the hole-spectrum theory in a probabilistic light. Regarding the signal and noise assumptions, S ∩ N is a random image formed from intersecting two random images, S and N. To apply our theory, we make the modeling assumption that for each realization of S ∩ N, the realization of N conforms to the corresponding realization of S. In the random model the hole spectrum is an n-tuple of random sets, and the HAS is a random n-vector. For a given basis B, we consider the expected value of the error (21), namely, \( E[e(B)] \). Optimality is achieved by finding B that minimizes \( E[e(B)] \). Note that \( E[e(B)] \) corresponds to \( E[e(\Psi)] \), where \( e(\Psi) \) is given in equation (19), and hence to \( E[e(Q)] \), where \( e(Q) \) is given in equation (17). The latter point is important, since, just like with linear frequency-component optimization, there is a relation between \( E[e(Q)] \) and \( E[e(\Psi)] \). Ideally, if the commuting diagram (18) were completed, then we would have \( E[e(Q)] = E[e(\Psi)] \).

We now consider the manner in which our theory relates to the general MSE optimization theory developed in [1]. Because that theory treats window observations as a collection of random variables to estimate another random variable, the "true" value of the image at a given pixel, it yields morphological filters given by erosion expansions for which the basis elements are pixel dependent; that is, the optimal filter is spatially variant. As discussed in [1], the filter becomes spatially invariant if the image-noise process is stationary.

To proceed with a comparison of the hole-spectrum approach and the general MSE approach of [1], we suppose \( S \cap N \) to be stationary as a random process. Applying the methodology of [1] to our noise model, at an arbitrary pixel x mean-square error is

\[
\text{MSE}[x] = E[|\Psi(S \cap N)(x) - S(x)|^2]
\]

where \( S(x) = 1 \) if x lies in S and \( S(x) = 0 \) otherwise. Letting \( P[ \mid ] \) denote conditional probability, MSE can be rewritten as

\[
\text{MSE}[x] = P[\Psi(S \cap N)(x) = 1 \mid S(x) = 0]P[S(x) = 0] \\
+ P[\Psi(S \cap N)(x) = 0 \mid S(x) = 1]P[S(x) = 1] \\
= P[x \in \Psi(S \cap N) \cap S] + P[x \in S \cap \Psi(S \cap N)]
\]

If we let A denote the number of pixels in the image, by stationarity we obtain \( A \{\text{MSE}[x]\} = E[e(\Psi)] \). Thus, minimization of \( E[e(\Psi)] \) is equivalent to MSE minimization.
6. CONCLUSION

The image representation discussed herein facilitates model-based optimization for a certain class of image-noise processes over a relevant class of morphological filters. Because we do not require stationarity to obtain a spatially invariant filter, the methodology always yields a global Matheron basis. Moreover, whereas in [1] the restriction was that the structuring elements must lie in some observation window about x, here we restrict our attention to a set of model elements. Although we have dropped a local estimation approach common to basic statistical estimation, thereby circumventing stationarity questions, we have paid a price. First, we have given up hope of obtaining general probabilistic distributional characterizations. Second, we have placed strict restrictions on the signal, the noise, and their interaction. Third, we have restricted our attention to filtering a certain type of noise, rather than presenting a theory that applies equally well to any type of noise filtering, compression, or general restoration. Our gain, however, is that when the model is appropriate we have a filter-design methodology that is intuitive, avoids the taxing statistical analysis required in the more general approach, and is not sensitive to deviations from stationarity. From an engineering perspective, the present model-based approach provides the type of spectral design in morphological optimization that the Wiener-frequency approach provides in linear optimization.

7. REFERENCES

