CHAPTER 14

Statistical morphology

R. M. HARALICK, E. DOUGHERTY, J. HA, T. KANUNGO, S. KARASU, C. K. LEE, L. RYSTROM, V. RAMESH & I. PHILLIPS, *Department of Electrical Engineering, University of Washington*

SUMMARY *This paper first introduces a parametric model for the generation of stationary random correlated binary sequences. The parameters of the model include the probability that a pixel is a binary one pixel and the length of the structuring element which dilates the initially spatially uncorrelated sequence. The spatial statistics of such eroded, dilated, opened and closed correlated binary sequences are derived in terms of the spatial statistics of the input binary sequence. Understanding of such one-dimensional processing is a precondition for understanding what happens in the more interesting two-dimensional case.*

1 Introduction

Statistical morphology is concerned with the statistical characterization of the four morphological operations—dilation, erosion, opening and closing. By statistical characterization of a morphological operator, we mean the statistical characterization of the output in terms of the statistical characteristics of the input. Characterization of operators allows us to predict the characteristics of the output of an algorithm composed of a sequence of morphological operations in terms of the statistical characteristics of the input and the sequence of morphological operators used. Furthermore, such statistical analyses of morphological algorithms are necessary to evaluate the algorithm's performance.

In this paper, we describe what we have learned about one way to characterize the dilation and opening morphological operators in a one-dimensional setting, i.e. the input to each of these operators is assumed to be binary one-dimensional. The input is modeled as a union of randomly translated discrete lines of a fixed length. The line segments can overlap and result in line segments of various lengths. Thus, the final output appears as an unordered pattern of lines and gaps of various lengths. This input is characterized by giving its line and gap length distribution and the distribution of the number of line and gap segments of various lengths. Therefore, the characterization of a morphological operator entails a similar characterization of the output.

There has been recent interest in the area of statistical morphology and some results have been published in the literature. Morales and Acharya (1992) analyzed the statistical characteristics of a morphological opening on gray scale signals perturbed by Gaussian noise. Stevenson and Arcs (1992) studied the effects of opening for a class of structuring elements. Astola *et al.* (1993) studied the output distributions of one-dimensional gray scale filtering. Costa and Haralick (1992) came up with an empirical description of the output gray level distributions of morphologically opened signals. Dougherty and Loce (1993) used libraries of structuring elements to restore corrupted signals in the case when a noise model is available.

In the following section, we set up the notation and difinitions used in this paper. In Section 3, we give a formal statement of the random process used to generate random sequences. In Section 4, we give a maximum likelihood algorithm for estimating the model parameters. The four morphological operators are characterized in Section 5.

2 Notation and definitions

Dilation is the morphological transformation which combines two sets using vector addition of set elements. If A and B are sets in $\mathbb{Z}^2$, the dilation of A by B is the set of all the possible vector sums of pairs of elements, one coming from A and one coming from B (Haralick *et al.*, 1987).

Definition 1. The dilation of A by B is denoted by $A \oplus B$ and is defined by

$$A \oplus B = \{c \in \mathbb{Z}^2 \mid c = a + b \text{ for some } a \in A \text{ and } b \in B\}$$

Erosion is the morphological dual of dilation. If A and B are sets in $\mathbb{Z} \times \mathbb{Z}$, then the erosion of A by B is the set of all elements of x for which $x + b \in A$ for every $b \in B$.

Definition 2. The erosion of A by B is denoted by $A \ominus B$ and is defined as

$$A \ominus B = \{x \in \mathbb{Z}^2 \mid x + b \in A \text{ for every } b \in B\}$$

Opening an image with a disk structuring element smoothes the contour, breaks narrow isthmuses, and eliminates small islands and sharp peaks or capes.

Definition 3. The opening of a set B by a structuring element K is denoted by $B \circ K$ and is defined as

$$B \circ K = (B \ominus K) \oplus K$$

The morphological operation of closing smoothes the contours in an image, fuses narrow breaks and long thin gulfs, eliminates small holes and fill gaps in the contours.

Definition 4. The closing of a set B by a set K is denoted by $B \bullet K$ and is defined as

$$B \bullet K = (B \oplus K) \ominus K$$

3 Random model

3.1 Generation process

Let us consider the discrete interval $\mathbf{D}=[1, d]$ of length d. A line-throwing process randomly throws discrete lines $\mathbf{K}=[1, k]$ of length k into the interval $\mathbf{D}$. These lines may overlap and create lines of lengths l greater than k.

The line-throwing process is as follows. First, the process is initialized by marking all the pixels in the interval $\mathbf{D}$ with a 0. Next, each pixel in the interval $\mathbf{D}$ is independently changed to a 1 with probability q. This set is now dilated with a line $\mathbf{K}$ of length k and whose origin is the extreme left point (1).

There are two questions that need to be answered:

(1) What is the joint density function associated with the line and gap lengths?
(2) Given an estimate of the density function of the line lengths, how do we estimate the probability q?

3.2 Joint probability distribution of line and gap lengths

More formally, let the line lengths and gap lengths be denoted by l and g respectively. Let the observed line and gap lengths be $g_0, l_1, g_1, \ldots, l_N, g_N$. It should be noted that these g_i and l_i values are ordered, and there are N lines and $N+1$ gaps. The first question implies finding $P(g_1, l_1, g_1, l_1, \ldots, l_N, g_N | q, k)$ and the second question implies finding $E[q | g_0, l_1, g_1, \ldots, l_N, g_N, k]$.

Now, using Bayes theorem, we have

$$P(q | g_0, l_1, g_1, \ldots, l_N, g_N, k) = \frac{P(g_0, l_1, g_1, \ldots, l_N, g_N | q, k) P(q | k) P(k)}{\int_{q'=0}^{q'=1} P(g_0, l_1, g_1, \ldots, l_N, g_N | q', k) P(q' | k) P(k) \, \mathrm{d}q'}$$

To evaluate the above, we need to evaluate $P(l | q, k)$. Let $w(l, t, k)$ be the number of different point sets with t elements, which, when dilated by K, produce a line of length l, i.e.

$$w(l, t, k) = \#\{A | A \oplus K = [1, l], \#A = t\}$$

It can be shown that $w(l, t, k)$ can be expressed recursively as

$$w(l, t, k) = \sum_{j=1}^{k} w(l-j, t-1, k) \tag{1}$$

The initial conditions on $w(l, t, k)$ are as follows:

(1) if $t \geqslant l-k+2$, then $w(l, t, k)=0$;
(2) if $l=k$ and $t=1$, then $w(l, t, k)=1$;
(3) if $l \neq k$ and $t=1$, then $w(l, t, k)=0$;
(4) if $l>tk$ and $t=1$, then $w(l, t, k)=0$.

A program was written to compute the values of $w(l, t, k)$ and is tabulated in Appendix A. A proof that $w(l, t, k)$ is given by equation (1) is given in Appendix B.

Using the definition of $w(l, t, k)$, we can express $P(l | q, k)$, i.e. the probability of observing a line of length l given q and $\mathbf{K}$, as follows:

$$P(l | q, k) = \sum_{t=\lceil l/k \rceil}^{l-k+1} w(l, t, k) q^t (1-q)^{(l-t)} \tag{2}$$

Since this $P(l|q,k)$ term is a probability, it should sum to unity, i.e.

$$\sum_{l=k}^{\infty} P(l|q,k) = \sum_{l=k}^{\infty} \sum_{t=[l/k]}^{l-k+1} w(l,t,k) q^t (1-q)^{(l-t)} = 1 \tag{3}$$

The joint probability of observing the lines and gaps is thus given by

$$P(g_0, l_1, g_1, l_1, \ldots, l_N, g_N | q, k) = [(1-q)^{\Sigma_{n=0}^{N} g_n}] \prod_{n=1}^{N} \left[\sum_{t^n = \lceil l_n/k \rceil}^{l_n - k + 1} w(l_n, t_n, k) q^{t_n} (1-q)^{l_n - t_n} \right] \tag{4}$$

The reasoning behind the above probability is as follows. The first term is the probability of the gap events. The summation computes the sum of the gap lengths; this is the total length of gap events. The product is over each line segment and there are N such lines. The inside summation is over the number of possible events in the set before dilation. Let the length of the nth line segment be l_n. Then, the minimum number of events that gave rise to a line of length l_n is $\lceil l_n/k \rceil$ and the maximum number is $l_n = k+1$. In other words, if there are fewer than $\lceil l_n/k \rceil$ elements in the point set before dilation, after dilation with a structuring element of length k, the resultant line will have a length smaller than l_n. Similarly, if the initial point set had more than $l_n + k + 1$ points, after dilation, the result will have a length greater than l_n.

Next, q^{t_n} is the probability of t_n events turning on and $(1-q)^{l_n - t_n}$ is the probability of the rest of the events not turning on. Finally, $w(l,t,k)$ denotes the number of ways that t_n events can be chosen such that the dilation results in an interval of length l.

4 Parameter estimation

In many problems in computer vision, the form of the probability densities may be assumed or may be known as *a priori* knowledge, but the values of the associated parameters are usually not known. The process of determining the values of the parameters from observations is known as parameter estimation and the value of the parameter that results is called an estimate. In general, the parameters to be estimated may be scalars, vectors or matrices. People usually use two kinds of estimation methods in the parameter estimation, based on the availability of prior knowledge about the distribution of the parameters to be estimated.

Maximum likelihood estimation assumes that the parameters are constants which are simply unknown. This approach requires no prior knowledge of the probability of various values of the parameters. Bayesian estimation assumes that the parameters are random variables for which a prior density function is given. This approach is usually used when prior knowledge about the parameter values is available, and thus the parameter is treated as a random vector with an associated probability density function.

We also conducted experiments to estimate parameters of the generation process using a constrained optimization method, but the results were not satisfactory. Instead, we use a maximum likelihood estimation technique which gives much better results in comparison with the optimization technique.

4.1 Maximum likelihood estimation

Maximum likelihood estimation (MLE) can be thought of as a procedure to be used when one has no prior knowledge (or is willing to assume none) about the probability of various values of the parameter. In our situation, the parameters to be estimated are the probability q and the length k of the structuring element involved in a line-throwing process. Let $\mathbf{I}=(g_0, l_1, g_1, l_2, g_2 \ldots, l_N, g_N)$ be an observation and let $\mathbf{p}=(q, k)$ be the vector of parameters to be estimated. In the context of an estimation problem, this kind of observation is viewed as a random variable, which is characterized by the joint conditional probability density $P(\mathbf{I}|\mathbf{p})$. When viewed as a function of $\mathbf{p}=(q, k)$ this conditional probability density is called a likelihood function. The likelihood function $L(\mathbf{I}|\mathbf{p})$ of a line-throwing process takes the form

$$L(\mathbf{I}|\mathbf{p})=(1-q)^{\Sigma_{n=0}^{N} g_n} \prod_{n=1}^{N} \left[\sum_{t_n=\lceil l_n/k \rceil}^{l_n-k+1} w(l_n, t_n, k) q^{t_n} (1-q)^{l_n-t_n} \right] \tag{5}$$

The MLE chooses as its estimate the value which maximizes the likelihood function.

If the necessary derivatives exist, one can obtain the MLE by setting the derivatives of $L(\mathbf{I}|\mathbf{p})$, or $\log L(\mathbf{I}|\mathbf{p})$, to zero. However, q is a continuous variable which takes values between 0 and 1, and k is a discrete variable which takes positive integers. We do not know of a way to obtain a closed form solution from the derivative of $L(\mathbf{I}|\mathbf{p})$ with respect to q. Therefore, we must obtain the MLE $\hat{\mathbf{p}}=(\hat{q}, \hat{k})$ which maximizes $L(\mathbf{I}|\mathbf{p})$ numerically. In the next section, some graphs are shown for a particular observation $\mathbf{I}$, and these might give some intuition to the characteristics of the likelihood function $L(\mathbf{I}|\mathbf{p})$.

4.2 Bayesian estimation

Simple Bayesian methods determine the posterior probability density $P(p|I)$ based on the observed data and prior knowledge about the parameters. From Bayes' theorem, the posterior probability density can be expressed as

$$P(p|I)=\frac{P(I|p)}{P(I)} P(p)$$

where the probability that I is observed is calculated as

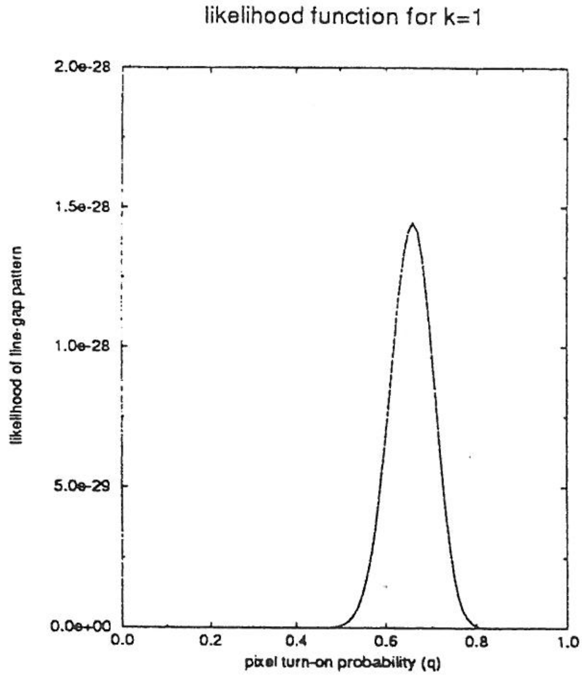
$$P(I)=\sum_k \int_{q=0}^{q=1} P(I|q, k) P(q, k) \, \mathrm{d}q$$

If the first term on the right-hand side is termed as the standardized likelihood, we can state Bayes' theorem in words as.

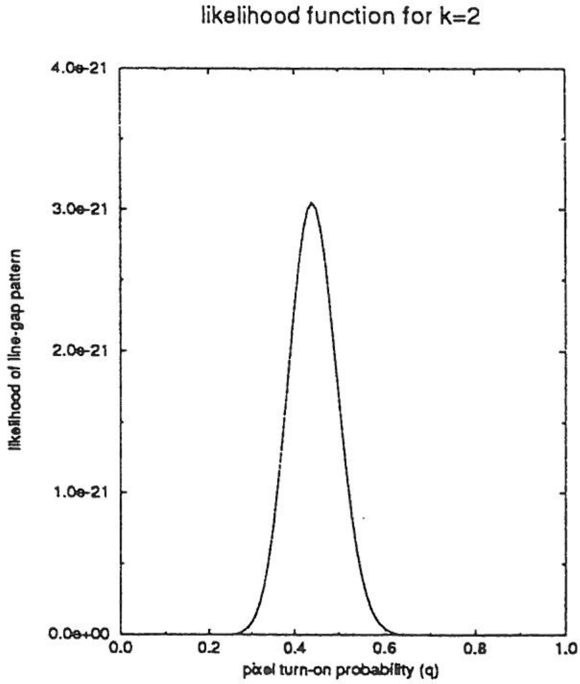
(posterior belief) = (standardized likelihood) × (prior belief)

The interpretation of this formula is simple and very commonsensical. Values of the unknown parameters which give rise to large values of the standardized likelihood will lead to higher posterior beliefs than will values of the unknown parameters which give rise to small values of the standardized likelihood. Furthermore, high values of prior belief correspond to parameter values which are likely to have led to the data we observed.

A sharply peaked likelihood function dominates over less sharply peaked prior densities and makes two posteriors obtained using two different but smooth priors become close together and vice versa. A special case is when the prior density $P(p)$ is uniform, i.e. when all values of p in the region of interest are equally likely. If this

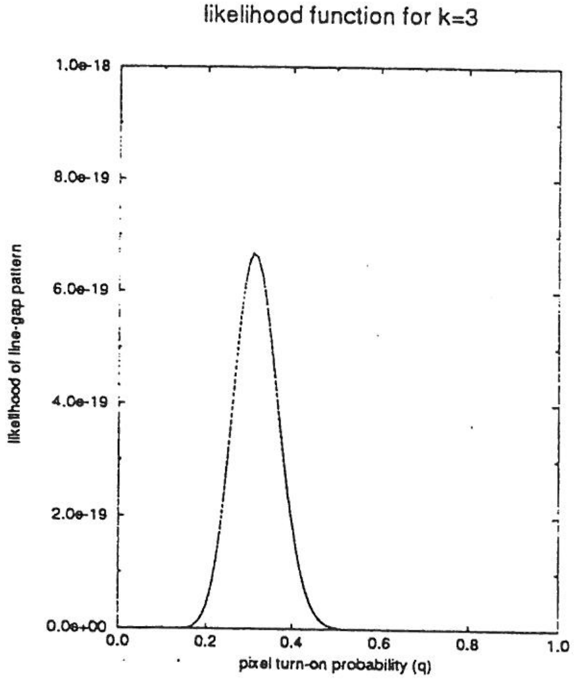


(a)

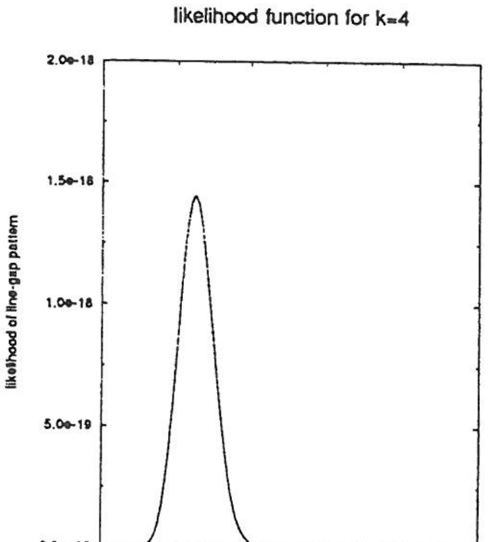


(b)

FIG. 1. Maximum likelihood function for dilation length (a) $k=1$ and (b) $k=2$.

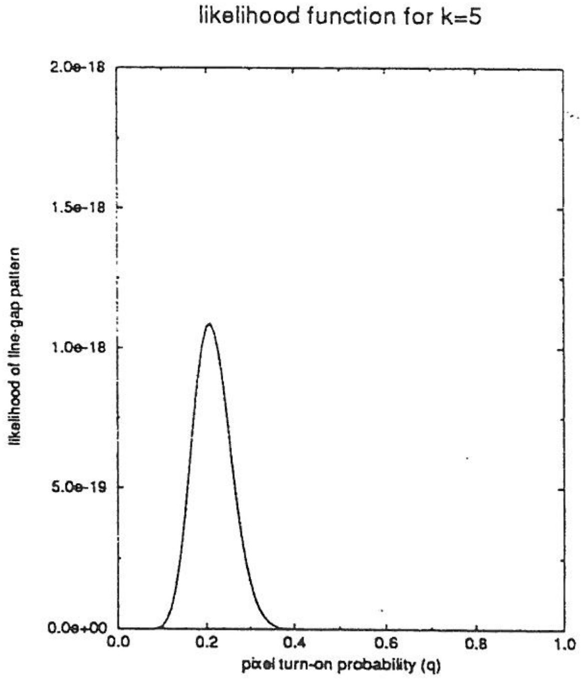


(a)

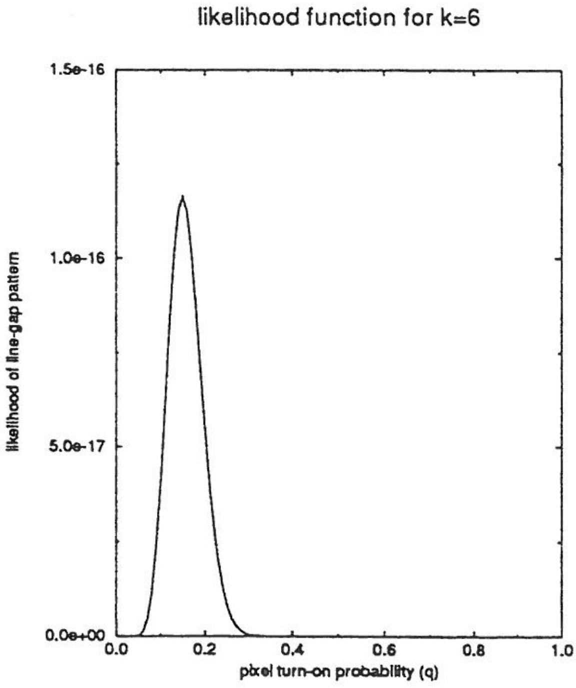


(b)

FIG. 2. Maximum likelihood function for dilation length (a) $k=3$ and (b) $k=4$.



(a)



(b)

FIG. 3. Maximum likelihood function for dilation length (a) $k=5$ and (b) $k=6$.

is the case, then maximizing the posterior density is equivalent to maximizing the likelihood function, and the Bayesian estimate is equal to the MLE.

4.3 Behavior of the likelihood function L(I/p)

Now, we will consider the behavior of the likelihood function $L(I|p)$. This might give us some insight before we actually carry out parameter estimation.

Let us consider the discrete interval $\mathbf{D}=[0,99]$. Assume that a line-throwing process occurs on $\mathbf{D}$ with p and k unknown. As explained earlier, a line-throwing process can be simulated by a morphological operation. Let us suppose that we have observed the following, as a result of the line-throwing process:

0000011111111100011111100111111011111110011111111111100000111111000111111111110000000000111111111000

Then, the observation $\mathbf{I}$ will be

$$\begin{aligned}\mathbf{I} &= (g_0, l_1, g_1, l_2, g_2, l_3 . g_3, l_4, g_4, l_5, g_5, l_6, g_6, l_7, g_7, l_8, g_8)\\ &= (5, 9, 3, 6, 2, 1, 7, 2, 12, 5, 6, 3, 11, 10, 9, 3)\end{aligned}$$

In the figures, plots of the likelihood function $L(\mathbf{I}|p)$ for various values of k are shown. Each plot is unimodal and shows a sharp peak. The maximum probability occurs at $k=6$ and $p \simeq 0.15$.

Although it would have been easy to choose an example where the observation leads to an estimate $\hat{p}$ and $\hat{k}$ that is close to the true parameter values p and k that were used to generate the observation, we have chosen in the above example an instance where the true values were $p=0.3$ and $k=6$. This observation illustrates the fact that, as the observation tends to become filled with binary 1 events, the estimated values for k will tend to shift higher and the estimated values for p will tend to shift lower. This shifting tends to maintain itself, even if the size of the observation sequence becomes larger.

Interesting questions in estimation concern the determination of the mean and variance of the estimators as a function of p, k, and n, the size of the observation sequence.

4.4 Regression estimation

An alternative to the maximum likelihood approach for estimating the two model parameters, pixel turn-on probability and dilation line segment length, is to use a linear model and select an estimator based on minimizing the fitting error over simulated data realizations. The advantage of the linear model is that it is well understood and may give reasonable results, evenin some highly non-linear situations. Often, non-linear functions of the realizations are created to allow additional flexibility in fitting, for instance, forming a histogram of the line segment lengths from the line segment length realization. We formed a linear estimator for the underlying 'seed probability' using the observed values and their squares in the regression. We assumed the dilation line segment length and interval length to be known.

The input to the estimator is a single realization of the form $l_1, g_2, \ldots, l_{N-2}, g_{N-1}, l_N$, which is simply a vector of the observed line and gap run lengths. From the observed data vector, a histogram of the observed counts is made, i.e. $H(i) = \#\{l | l_j = i, j = 1, 3, \ldots, N\}$. In addition to the histogram, another statistic, i.e. the sum of the observed gap lengths ($g_{\text{total}}=\Sigma_i g_i$), was formed from the realization. This was done because of the role of this statistic in the joint gap length run length density function. A linear estimator of the probability $\mathbf{p}$ can be formed as $\hat{\mathbf{p}}=(g_{\text{total}}, H)\mathbf{x}$, where $\mathbf{x}$ is a column vector determined from a least-squares regression procedure using simulated data. An additional constraint is that $\mathbf{p}$ is a probability and thus must lie between zero and one. This constraint was taken into account when solving the least-squares problem by using the methods described by Lawson and Hanson (1991). The results of this regression did not produce satisfactory estimates.

5 Characterization of morphological operators

Let us suppose that the random model generation process in Section 3 produces N line segments in an interval of length d, with a gap as the first event and a gap as the last event. Thus, there are $N+1$ gaps and the maximum length a line can have is $M=d-2$, in order to allow two gaps of length 1 on either side. Let $m_1, m_2, \ldots, m_M$ be the number of lines with lengths $1, 2, \ldots, M$ respectively. We have the following relationships:

$$\sum_{i=1}^{M} i m_i \leqslant d-(N+1)$$

$$\sum_{i=1}^{M} m_i = N \leqslant \lfloor M/2 \rfloor$$

The first expression states that the number of locations covered by the generated lines plus one location for each of the $N+1$ gaps must be less than d, i.e. the total length of the interval. The second expression says that the sum of the number of lines of each different length must be N, i.e. the total number of lines must be less than $\lfloor M/2 \rfloor$ to permit gaps to separate the different lines.

From Section 3, we know that

$$P(g_0, l_1, g_1, l_1, \ldots, l_N, g_N \mid q, k, d, N) = [(1-q)^{\Sigma_{n=1}^{N} g_n}] \prod_{n=1}^{N} \left[\sum_{t_n=\lceil l_n/k \rceil}^{l_n-k+1} w(l_n, t_n, k) q^{t_n} (1-q)^{l_n-t_n} \right]$$

We can rewrite the above equation in terms of m_i terms as follows (it should be noted that $m_i=0$ for $0 \leqslant i < k$) and $M=N-2$:

$$P(g_0, l_1, g_1, l_1, \ldots, l_N, g_N \mid q, k, d, N) = [(1-q)^{d-\Sigma_{i=k}^{M} i m_i}] \prod_{i=k}^{N} \left[\sum_{t_i=\lceil i/k \rceil}^{i-k+1} w(i, t_i, k) q^{t_i} (1-q)^{i-t_i} \right]^{m_i} \tag{6}$$

Now, since the order of gaps and lines does not affect the right-hand side of the above equation in any way, we can conclude the following:

$$P(m_k, m_{k+1}, \ldots, m_M \mid q, k, d, N) = [(1-q)^{d-\Sigma_{i=k}^{M} i m_i}] \prod_{i=k}^{M} \left[\sum_{t_i=\lceil i/k \rceil}^{i-k+1} w(i, t_i, k) q^{t_i} (1-q)^{i-t_i} \right]^{m_i} \tag{7}$$

We can now investigate the effect of morphological operations on the distributions of line and gap lengths.

5.1 Dilation

We are given a model of the input in terms of the pixel turn-on probability q and the dilation line length k. We want to consider what happens if the input is dilated with a line of length k_d. Since the input itself was a result of a dilation of a point set produced by randomly turning on each pixel with a probability q and dilating the result with a line of length k, we can now consider the two dilations as a single dilation operation with a structuring element of length $k+k_d-1$, instead of as two dilations with lengths k and k_d. Thus, the probability equation for the distribution of the line lengths does not change, except that the line length used for dilation is $k+k_d$ instead of k. This is summarized below.

$$P(m_{k+k_d}, m_{k+k_d+2}, \ldots, m_M | q, k, d, M)$$
$$= [(1-q)^{d-\Sigma_{i=k+k_d}^{M} i m_i}] \prod_{i=k+k_d}^{M} \left[\sum_{t_i=\lceil i/k \rceil}^{i-k+1} w(i, t_i, i) q^{t_i} (1-q)^{i-t_i} \right]^{m_i} \quad (8)$$

5.2 Erosion

In the case of erosion with a line of length k_e, all the lines in the input with lengths smaller than k_e are eliminated. Furthermore, all the lines with lengths greater than or equal to k_e become smaller by a length k_e-1 at the output. Thus, if m_i is the number of lines in the input with length i and m'_i is the number of lines of length i in the output, we can say that $m'_i=0$ for $0 \leqslant i < k_e$ and $m'_i = m_{i+k_e}$ for $k_e \leqslant i \leqslant M-k_e$, where M is the maximum line length possible in the input. This results in the shifting of the whole frequency plot to the left by k_e-1. Such a statement could not be made in the case of dilation, because of the fact that two lines of lengths i and j could merge and produce a line of length $i+j$, so decreasing the counts of m'_i and m'_j by one and increasing the count of m'_{i+j} by one.

Now, we give a more formal treatment for the output distribution for the case $k_e > k$. Since all the lines of length $k, \ldots, k_e$ are eliminated after erosion, we need to find the conditional probability of the event. Let us denote this conditional probability by $P(m_k, m_{k+1} \ldots, m_{k_e-1} | q, k, d)$. To find the final conditional probabilities, we need to sum the conditional probability $P(m_k, m_{k+1} \ldots, m_M | q, k, d)$ over the $m_k, m_{k+1}, \ldots, m_{k_e-1}$ terms as

$$P(m_1, \ldots, m_{M-k_e+1} | q, k, d, k_e)$$
$$= \sum_{n_k} \cdots \sum_{n_{k_e-1}} P\left(m_{k_e}, \ldots, m_M | q, k, d, N - \sum_{l=k}^{k_e-1} n_l \right) P(n_k, \ldots, n_{k_e-1} | q, k, k_e, d) \quad (9)$$

where $k_e > k$ and the conditonal probabilities are evaluated using equation (7) but with different values of N and where the summation is done over all non-negative integers $n_R, \ldots, n_{R_{e-1}}$ satisfying

$$\sum_{i=k}^{k_e-1} n_i \leqslant N$$

5.3 Opening

The results of opening are similar to those of erosion. After opening with a line of length k_0, all the lines in the input with a length smaller than k_0 are eliminated, but the rest of the lines longer than k_0 are left unchanged. Thus, if m_i is the number of lines in the input with length i and m'_i is the number of lines of length i in the output, we can say that $m'_i=0$ for $0 \leqslant i < k_0$ and $m'_i = m_i$ for $k_0 \leqslant i \leqslant M$, where M is the maximum line length possible in the input. It follows that

$$P(m_{k_0}, \ldots, m_M | q, k, d, k_0, N)$$
$$= \sum_{n_k} \cdots \sum_{n_{k_0}} P\left(m_{k_0}, \ldots, m_M | q, k, d, N - \sum_{l=k}^{k_0-1} n_l \right) P(n_k, \ldots, n_{k_0-1} | q, k, k_0, d, N) \quad (10)$$

where the conditional probabilities are evaluated using equation (7) but with different values of N.

5.4 Closing

Closing by a line of length k_c is a dilation by a line of length k_c, followed by an erosion by a line of length k_c. The dilation gives rise to a condition probability given by (see Section 5):

$$P(m_{k+k_c}, m_{k+k_c+1}, \ldots, m_M | q, k, k_c, d, M)$$
$$= [(1-q)^{d - \Sigma_{i=k+k_c}^{M} i m_i}] \prod_{i=k+k_c}^{M} \left[\sum_{t_i = \lceil i/k \rceil}^{i-k+1} w(i, t_i, i) q^{t_i} (1-q)^{i-t_i} \right]^{m_i} \quad (11)$$

This is followed by an erosion by a length k_c. However, since the minimum length in the input to erosion is k_c, no line segments in the input will be eliminated. Now, we can use the same formulation we used for computing the conditional probabilities in the case of erosion, i.e.

$$P(m_k, \ldots, m_M | q, k, k_c, d, M) = \sum_{n_k} \cdots \sum_{n_k + k_c - 1} P\left(m_{k+k_c-1}, \ldots, m_M | q, k, d, N - \sum_{l=k}^{k+k_c-1} n_l \right)$$
$$\times P(n_k, \ldots, n_{k+k_c-1} | q, k, k_c, d, N) \quad (12)$$

6 Summarizing remarks

To characterize statistically the dilation and opening morphological operators, we characterized the input statistically and then used the same representation to characterize the output. The statistical characterization of the input and output was in terms of the distributions of line lengths, gap lengths and number of line lengths per unit length. The output was characterized in a similar fashion. The parameters of the output were represented as functions of the parameters of the input and the parameters of the morphological operators. MLE of the input model parameters was derived and experimentally validated.

Acknowledgements

The authors would like to thank A. Badr, S. Chen, W. Costa, J. Suri and W. Wei for their comments and criticisms at various stages of this work.

Correspondence: R. M. Haralick, Intelligent Systems Laboratory, Department of Electrical Engineering, University of Washington, Seattle, WA 98195, USA.

TABLE A1. Values of $w(l, t, k)$ tabulated for various values of l, t and k

	$t=0$																
$l=0$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	1	2	1	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	3	3	1	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	2	6	4	1	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	7	10	5	1	0	0	0	0	0	0	0	0	0
	0	0	0	0	6	16	15	6	1	0	0	0	0	0	0	0	0
	0	0	0	0	3	19	30	21	7	1	0	0	0	0	0	0	0
	0	0	0	0	1	16	45	50	28	8	1	0	0	0	0	0	0
	0	0	0	0	0	10	51	90	77	36	9	1	0	0	0	0	0
	0	0	0	0	0	4	45	126	161	112	45	10	1	0	0	0	0
	0	0	0	0	0	1	30	141	266	266	156	55	11	1	0	0	0
	0	0	0	0	0	0	15	126	357	504	414	210	66	12	1	0	0
	0	0	0	0	0	0	5	90	393	784	882	615	275	78	13	1	0
	0	0	0	0	0	0	1	50	357	1016	1554	1452	880	352	91	14	1
	0	0	0	0	0	0	0	21	266	1107	2304	2850	2277	1221	442	105	15

Here, $k=3$; t increases from left to right, starting from 0 at the left-hand column; l increases from top to bottom and starts with 0 at the top row.

REFERENCES

ASTOLA, J., KOSKEINEN, L. & NEUVO, Y. (1993) Statistical properties of discrete morphological filters. *In:* E. R. DOUGHERTY (Ed.), *Mathematical Morphology in Image Processing* (New York, Marcel Dekker), pp. 93–150.

COSTA, W. S. & HARALICK, R. M. (1992) Predicting expected grey level statistics of opened signals, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, June, pp. 554–559.

DOUGHERTY, E. R. & LOCE, R. P. (1993) Efficient design strategies for the optimal digital morphlogical filters: probabilities, constraints, and structuring-element libraries. *In:* E. R. DOUGHERTY (Ed.), *Mathematical Morphology in Image Processing* (New York, Marcel Dekker), pp. 43–92.

HARALICK, R. M., STERNBERG, S. R. & ZHUANG, X. (1987) Image analysis using mathematical morphology, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-9, pp. 532–550.

LAWSON, C. L. & HANSON, R. J. (1991) *Solving Least Squares Problems* (Englewood Cliffs, NJ, Prentice-Hall).

MORALES, A. & ACHARYA, R. (1992) Non-linear multiscale filtering using mathematical morphology, *Computer Vision and Pattern Recognition*, June, pp. 592–578.

STEVENSON, R. L. & ARCE, G. (1989) Morphological filters: statistical and further syntactic properties, *IEEE Transactions on Circuits and Systems*, CAS-34, pp. 1292–1305.

Appendix A: tabulated values of $w(l, t, k)$

Table A1, we give sample output of the $w(l, t, k)$. There are a few things worth noting.

(1) The non-zero numbers fall in a wedge-shaped region. On one side of the wedge, the slope is given by 1 unit increment in l for every unit increment in t. On the other side, there are k units of increment of l for every unit increment of t.

(2) for the special case of $k=2$, this wedge becomes Pascal's triangle.

Appendix B: proof of $w(l, t, k)$

Let **L** be the discrete interval $[1, l]$ of length l and **K** be a structuring element of length k, i.e. $\mathbf{K}=[1, k]$. Consider the point set $\Theta_t^l \subseteq [1, l]$ such that $\Theta_t^l \oplus \mathbf{K}=[1, l]$ and $\#\Theta_t^l=t$. Now, since after dilation of Θ_t^l with a structuring element of length k we produce a line length of l, it is obvious that the last element in the set must fall precisely on the point $l-k+1$. If there is any element after this point, then it dilates past the end of l, which we postulate to be not allowed. This we state in the following proposition.

Proposition B1. Let $\Theta_t^l \subseteq [1, l]$ be such that $\Theta_t^l \oplus \mathbf{K}=[1, l]$ and $\#\Theta_t^l=t$. Then, we have

$$[l-k+1, l] \cap \Theta_t^l=\{l-k+1\}$$

Next, we claim that $\{l-k+1\}$ does not belong to Θ_t^{l-i} for $i>0$.

Proposition B2. Let $\Theta_{t-1}^{l-i} \subseteq [1, l-1]$ be such that $\Theta_{t-1}^{l-i} \oplus \mathbf{K}=[1, l-i]$ and $\#\Theta_{t-1}^{l-i}=t-1$, Then, for $i>0$, we have

$$\{l-k+1\} \notin \Theta_{t-1}^{l-i}$$

Now we have $t-1$ points remaining. To be able to cover the whole line of length l with t points, it is clear that the remaining $t-1$ points must cover either $l-1$ points, or $l-2$ points, or ... $l-k$ points. If the dilated line covers a length smaller than $l-k$, the final line will have gaps and, hence, will not cover an l interval of length l completely. Thus, to count the number of different point sets that, on dilation, can give rise to a line of length l, we need to count the number of different point sets that, on dilation, can give rise to lines of lengths of $l-1, l-2, \ldots \odot, l-k$.

Proposition B3. Let $\Theta_{t-1}^{l-i} \subseteq [1, l-1]$ be such that $\Theta_{t-1}^{l-i} \oplus \mathbf{K}=[1, l-i]$ and $\#\Theta_{t-1}^{l-i}=t-1$, Then, for $1 \leqslant i \leqslant k$, we have

$$\Theta_{t-1}^{l-i} \oplus \mathbf{K} \cup \{l-k+1\} \oplus \mathbf{K}=\mathbf{L}$$

The following proposition states that two sets are different if, when dilated, they give rise to two intervals of different lengths, i.e. they are mutually exclusive.

Proposition B4. Let $\Theta_{t-1}^{l-i} \subseteq [1, l-1]$ be such that $\Theta_{t-1}^{l-i} \oplus K=[1, l-i]$ and $\#\Theta_{t-1}^{l-i}=t-1$. Then, for $i \neq j$, we have

$$\Theta_{t-1}^{l-i} \neq \Theta_{t-1}^{l-i}$$

From the above two propositions, we can see that, to count the number of ways of constructing an interval of length l from t points, we could count the number of ways of constructing intervals of lengths $l-1, l-2, \ldots, l-k$, and sum the counts, since these are mutually exclusive events. This we state in the following proposition.

Proposition B5. Let

$$w(l, t, k)=\#\{\Theta_t^l \mid \Theta_{t-1}^l \oplus \mathbf{K}=[1, l] \text{ and } \#\Theta_t^l=t\}$$

Then, we have

$$w(l, t, k)=\sum_{i=1}^{k} w(l-i, t-1, k)$$