# Statistical Morphology

R.M. Haralick, E. Dougherty*,
J. Ha, T. Kanungo, S. Karasu,
C.K. Lee, L. Rystrom, V. Ramesh, I. Phillips

Intelligent Systems Laboratory
Department of Electrical Engineering, FT-10
University of Washington
Seattle, WA 98195

* Rochester Institute of Technology
Center for Imaging Science
Rochester, New York 14623

**Abstract**

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 is 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 is necessary for evaluating 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. That is, 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. The characterization of a morphological operator, therefore, entails a similar characterization of the output.

There has been a recent interest in the area of statistical morphology and some results have been published in the literature. Morales and Acharya [MA92] analyzed the statistical characteristics of a morphological opening on grayscale signals perturbed by Gaussian noise. Stevenson and Arce [SA92] studied the effects of opening for a class of structuring elements. Atola, Koskinen and Neuvo [ALN93] studied the output distributions of one dimensional grayscale filtering. Costa and Haralick [CH92] came up with an empirical description of the output graylevel distributions of morphologically opened signals. Dougherty and Loce [DL93] 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 definitions 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 $\mathbf{Z}^2$, the dilation of $A$ by $B$ is the set of all possible vector sums of pairs of elements, one coming from $A$ and one coming from $B$ [HSZ87].

**Definition 2.1** The *dilation* of $A$ by $B$ is denoted by $A \oplus B$ and is defined by

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

Erosion is the morphological dual of dilation. If $A$ and $B$ are sets in $\mathbf{Z} \times \mathbf{Z}$, 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 2.2** The *erosion* of $A$ by $B$ is denoted by $A \ominus B$ and is defined as

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

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

**Definition 2.3** 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 closing smooths the contours in an image, fuses narrow breaks and long thin gulfs, eliminates small holes, and fill gaps in the contours.

**Definition 2.4** 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

Consider the discrete interval $D = [1, d]$ of length $d$. A line throwing process randomly throws discrete lines $K = [1, k]$ of length $k$ into the interval $D$. 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 $D$ with a 0. Next, each pixel in the interval $D$ is independently changed to a 1 with probability $q$. This set is now dilated with a line $K$ 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 $g_0, l_1, g_1, \ldots, l_N, g_N$. Note that these $g_i$'s and $l_i$'s are ordered and there are $N$ lines and $N + 1$ gaps. The first question implies finding $P(g_0, l_1, g_1, l_1, \ldots, l_N, g_N | q, k)$ 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,

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

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 $l$. That is,

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

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

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

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

1. If $t \geq l - k + 2$, $w(l, t, k) = 0$.
2. If $l = k$ and $t = 1$, $w(l, t, k) = 1$.
3. If $l \neq k$ and $t = 1$, $w(l, t, k) = 0$.
4. If $l > t \cdot k$ and $t = 1$, $w(l, t, k) = 0$.

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 in appendix B.

Using the definition of $w(l, t, k)$ we can express $P(l|q, k)$, the probability of observing a line of length $l$ given $q$ and $K$, 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 $P(l|q, k)$ is a probability, it should sum to 1. That is,

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

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

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

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. There are $N$ such lines. The inside summation is over the the number of possible events in the set before dilation. Let the length of the $n$th 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 $l_n - k + 1$. That is, 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 $w(l, t, k)$ the number of ways that $t_n$ events can be chosen such that the dilation results in an interval of length $l$. The proof that $w(l, t, k)$ is given by equation (1) is in the appendix.

# 4 Parameter Estimation

In many problems in computer vision, the form of the probability densities may be assumed or may be known as a prior 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.

## 4.1 Maximum Likelihood Estimation

Maximum likelihood estimation 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 $I = (g_0, l_1, g_1, l_2, g_2, \cdots, l_N, g_N)$ be an observation and let $p = (q, k)$ 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 $P(I|p)$. When viewed as a function of $p = (q, k)$, this conditional probability density is called a likelihood function. The likelihood function $L(I|p)$ of a line throwing process takes the form of Eq (5 ), that is,

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

The maximum likelihood estimation chooses, as its estimate, the value which maximizes the likelihood function.

If the necessary derivatives exist, one can obtain the maximum likelihood estimate by setting the derivatives of $L(I|p)$, or $\log L(I|p)$, 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 get a closed form solution from the derivative of $L(I|p)$ with respect to $q$. Therefore, we must obtain the maximum likelihood estimate $\hat{p} = (\hat{q}, \hat{k})$ which maximizes $L(I|p)$, numerically. In the next section, some graphs are shown for a particular observation $I$ and they might give some intuition to the characteristics of the likelihood function $L(I|p)$.

## 4.2 Behavior of the likelihood function $L(I|p)$

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

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

0000011111111100011111100111111011111110011111111111100000111111000111111111110000000000111111111000

Then the observation $I$ will be

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

In the figures plots of the likelihood function $L(I|p)$ 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 $p = .3$ and $k = 6$. This observation illustrates the fact that as the observation tends to get 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 shift tends to maintain itself even if the size of the observation sequence becomes larger.

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

# 5 Characterization of Morphological Operators

Suppose that the random model generation process of 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 one 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 relations.

$$\begin{aligned} \sum_{i=1}^{M} i m_i &\leq d - (N+1) \\ \sum_{i=1}^{M} m_i &= N \leq \lfloor M/2 \rfloor \end{aligned}$$

The first 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 $d$, the total length the interval. The second says that the sum of the number of lines of each different length must be $N$, the total number of lines which must be less than $\lfloor M/2 \rfloor$ to permit gaps to separate the different lines.

From section 3 we know that

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

We can rewrite the above equation in terms of $m_i$'s as follows. Note that $m_i = 0$ for $0 \leq i < k$.

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

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

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

Now we can 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 two dilations with lengths $k$ and $k_d$. Thus the probability equation for distribution of line lengths does not change except for the line length used for dilation is $k + k_d$ instead of $k$. This is summerized below. Note $m'_i$ denotes the number of lines of length $i$ in the output and $m_i$ denotes the number of lines of length $i$ in the input.

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

### 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 or equal to $k_e$ become smaller by a length $k_e - 1$ at the output. Thus if $m_i$ are the number of lines in the input with length $i$, and $m'_i$ are the number of lines of length $i$ in the output, we can say that $m'_i = 0$ for $0 \leq i < k_e$ and $m'_i = m'_{i+k_e}$ for $k_e \leq i \leq M - k_e$ 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 $i + j$ and thereby 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 after erosion all the lines of length $k, \ldots, k_e$ are eliminated, 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 $m_k, m_{k+1}, \ldots, m_{k_e-1}$ as follows.

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

where $k_e > k$ and the conditional probabilities are evaluated using equation (7) but with different values of $N$.

### 5.3 Opening

Results of opening are similar to that of erosion. After opening with a line of length $k_o$, all the lines in the input with a length smaller than $k_o$ are eliminated but rest of the lines of lengths greater than $k_o$ are left unchanged. Thus if $m_i$ are the number of lines in the input with length $i$, and $m'_i$ are the number of lines of length $i$ in the output, we can say that $m'_i = 0$ for $0 \leq i < k_o$ and $m'_i = m'_i$ for $k_o \leq i \leq M$ where $M$ is the maximum line length possible in the input.

$$P(m_{k_o}, \ldots, m_M | q, k, d, k_o) =$$

$$\sum_{n_k}\cdots\sum_{n_{k_o-1}} P(m_{k_o},\ldots,m_M|q,k,d,N-\sum_{l=k}^{k_o-1} n_l)$$
$$\cdot P(n_k,\ldots,n_{k_o-1}|q,k,k_o,d,N) \tag{10}$$

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 on dilation):

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

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

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

## 6 Summary

In order to statistically characterize 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 distributions of line-lengths, gap-lengths, number of line-lengths per unit length. 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. Maximum likelihood estimate of the input model parameters was derived and experimentally validated.

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

## References

[ALN93] J. Astola, L.Koskinen, and Y. Neuvo. Statistical properties of discrete morphological filters. In E.R. Dougherty, editor, *Mathematical Morphology in Image Processing*, pages 93–150. Marcel-Dekker, 1993.

[CH92] W.S. Costa and R.M. Haralick. Predicting expected grey level statistics of opened signals. In *Computer Vision and Pattern Recognition*, pages 554–559, June, 15-18 1992.

[DL93] E.R. Dougherty and R.P. Loce. Efficient design strategies for the optimal digital morphlogical filters: Probabilites, constraints, and structuring-element libraries. In E.R. Dougherty, editor, *Mathematical Morphology in Image Processing*, pages 43–92. Marcel-Dekker, 1993.

[HSZ87] R.M. Haralick, S.R. Sternberg, and X. Zhuang. Image analysis using mathematical morphology. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, PAMI–9(4):532–550, July 1987.

[MA92] A. Morales and R. Acharya. Non-linear multiscale filtering using mathematical morphology. In *Computer Vision and Pattern Recognition*, pages 572–578, June, 15-18 1992.

[SA92] R.L. Stevenson and G. Arce. Morphological filters: Statistical and further syntactic properties. *IEEE Trans. on Circuits and Systems*, CAS-34(11), 1992.

# A Tabulated values of $w(l, t, k)$

In the following table we give sample output of the $w(l, t, k)$. There are few things worth noting. 1) The non-zero numbers fall in a wedge shaped region. On one side of the wedge has slope is given by 1 unit increment in $l$ for every unit increment in $t$. On the other side there is $k$ units of increment of $l$ for every unit increment of $t$. 2) for the special case of $k = 2$ this wedge becomes the Pascal's triangle.

Table 1: The $w(l, t, k)$ values are tabulated here for various values of $l, t, k$. Here $k = 3$. $t$ increases from left to right starting from 0 at the left column. $l$ increases from top to bottom and starts with 0 at the top row.

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 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 | 1 | 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 |

# B Proof of $w(l, t, k)$

Let $L$ be discrete interval $[1, l]$ of length $l$ and $K$ be a structuring element of length $k$, i.e., $K = [1, k]$. Consider the point set $\Theta_t^l \subseteq [1, l]$ such that $\Theta_t^l \oplus K = [1, l]$ and $\#\Theta_t^l = t$. Now, since after dilation of $\Theta_t^l$

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 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 B.1** Let $\Theta_t^l \subseteq [1,l]$ be such that $\Theta_t^l \oplus K = [1,l]$ and $\#\Theta_t^l = t$. Then

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

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

**Proposition B.2** 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 $i > 0$,

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

Now we have $t-1$ points left. To be able to cover the whole line of length $l$ with $t$ points, then 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 that $l-k$, the final line will have gaps and hence will not cover a $l$ interval completely. Thus in order 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 $l-1, l-2, \ldots, l-k$.

**Proposition B.3** 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 $1 \leq i \leq k$

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

The following proposition states that two sets, which when dilated give rise to two intervals of different lengths, are different. That is they are mutually exclusive.

**Proposition B.4** 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 $i \neq j$

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

From the above two propositions we see that in order 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 B.5** Let

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

Then

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

Figure 1: Maximum likelihood function for dilation length $k = 1$ and $k = 2$.

Figure 2: Maximum likelihood function for dilation length $k = 3$ and $k = 4$.

Figure 3: Maximum likelihood function for dilation length $k = 5$ and $k = 6$.