CHAPTER 11

MOMENTS

Godfried Toussaint

ABSTRACT

This chapter introduces the basic principles behind the use of moments in various aspects of the pattern recognition problem.

1. Introduction

A graduate student named Sigfried in the 1970's was preparing to do research in character recognition for his masters thesis. He was aware of a variety of techniques described in the literature for extracting descriptors of shape, such as the contour tracing based methods used at M.I.T. for the design of reading machines for the blind, and he wanted to come up with something different. Sigfried had one year earlier been exposed to a course on communication theory and had some knowledge about random variables and Gaussian noise. He was familiar with some probability theory and knew what a *probability density function* was. He knew that the Gaussian (sometimes called *normal*) probability density function of one variable, say x, had a nice curvy "bell-shaped" form and that it was given by the expression:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]\right)$$

for which,

$$E[x] = \int xp(x)dx = \mu$$

and

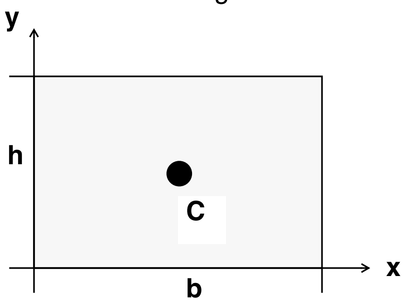
$$E[(x-\mu)^2] = \int (x-\mu)^2 p(x)dx = \sigma^2$$

where μ and σ^2 are two parameters referred to as the *mean* and the *variance*. The mean is the expected value of x denoted by $E[x]$ and the variance is the expected value of $(x-\mu)^2$. Sigfried also knew that the form of this *bell-shaped* function was *completely* determined by precisely and only these two parameters. Nevertheless, he had totally overlooked any possibility that two-dimensional counterparts and generalizations of such measures as μ and σ^2 might serve as ideal shape descriptors of bivariate probability density functions which in this context are better interpreted as a light intensity functions.

Sigfried had just read a couple of books on *lateral thinking* by Edward de Bono [Bo70], [Bo71], (see also the later edition [Bo78]). One of the principles of lateral thinking is to use randomness to arrive at new ideas and Sigfried decided to give this method a try in relation to finding a topic of research for his thesis since his advisor was too busy writing his own papers to be of

MOMENTS OF INERTIA OF COMMON GEOMETRIC SHAPES

Rectangle

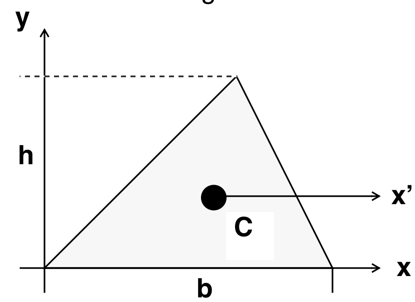


$I_x = (1/3)\ bh^3$

$I_y = (1/3)\ b^3h$

$J_c = (1/12)\ bh\ (b^2 + h^2)$

Triangle

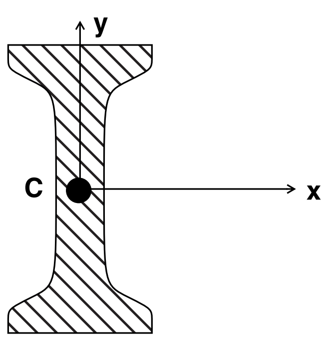


$I_{x'} = (1/36)\ bh^3$

$I_x = (1/12)\ bh^3$

SOME COMMON BEAMS

I-Beam



L-beam

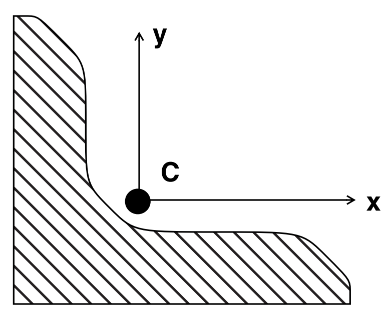


Fig. 1.1 A page from a book on *dynamics*.

much help. De Bono described various algorithms in his book that one could use to invoke the magic powers of randomness. Sigfried decided to browse through books in the library and adhering strictly to de Bono principles, by flips of the coin ended up in the engineering library. He was prepared to search even the Islamic Studies Library had the coins dictated such an outcome but he was pleased with the result. Another flip of the coin directed him to the civil engineering section and he hesitated but reminded himself to give de Bono an un-challenged chance. A further flip of the coin led him to open a book on *dynamics* and at this point he was visibly worried. Persistently and with an open mind he skimmed through the pages until his attention was suddenly and forcefully drawn to a page with figures such as those in Fig. 1.1. There under the title *Moments if Inertia of Common Geometric Shapes* were all sorts of shapes such as squares, triangles and circles with mathematical formulas printed next to them from which to calculate rotational moments of inertia about several different axes on the plane as well as points or axes orthogonal to the plane. Sigfried quickly noticed that different shapes had different formulas and gave different values even if the shapes had the same height, width or area. He was now getting excited. He turned to the next page and was electrified. Here was a page full of letter shapes, exactly his field of interest. In the book they were called I-beams, X-beams and L-beams and what have you, but to him they were just letters of the alphabet and they all had their formulas for the moments. Sigfried realized he had stumbled on a new and different method for extracting shape features from patterns. He put the dynamics book back on the shelf, kissed his worn-out paperback copy of de Bono's *Lateral Thinking* and rushed back to his office to tell his room mate he had found a topic for his thesis, feature extraction for character recognition by the *method of moments*.

2. Moments of Marginal Distributions or Collapsing-Projections

Consider a one dimensional function $g(x)$ normalized so that the area under the curve is equal to one. For our purposes it then behaves just like a probability density function. Also let us consider not the variance of x but the resulting expression one obtains when we remove μ from the equation yielding

$$E[(x)^2] = \int (x)^2 g(x) dx$$

This expression is seen to be quite similar to that of $E[x]$, the only difference being the presence of the number two. We call $E[x]$ and $E[x^2]$ the *first* and *second moments* of $g(x)$. In fact, we can generalize this notion and define the k-th moment of $g(x)$, for k=1,2,... as

$$m_k = \int (x)^k g(x) dx$$

In this way we now are able to obtain a *countably* infinite number of shape measurements for describing $g(x)$. Of course an infinite number of measurements yields an impractical system. Luckily, in practice, even if the shape of $g(x)$ is quite complex compared to the bell-shape of a Gaussian density function, not many moments are needed to adequately describe it and it is this property of the *moments* as well as their application-independent generality that makes them attractive.

How can we use the above-mentioned moments of one-dimensional functions to describe the shape of two-dimensional patterns? One way is to obtain from the two-dimensional pattern, one or more *collapsing projections* of the pattern onto carefully chosen axes. In psychology this approach is referred to as the *marginal distribution approach to visual form perception* [Zu70]. This approach is illustrated in Fig. 2.1 where the pattern P is shown along with three collapsed projections onto axes parallel to the x-axis, y-axis and $x+y$ axis. The collapsed projection along an axis can be viewed as a function that describes the length of the line segments obtained by intersecting lines orthogonal to the axis with the pattern P. Once a few of these collapsed projections are obtained, each of them can be normalized so that they resemble one-dimensional probability density functions (or distributions in the discrete case) and the moments of these distributions can then be computed. Sutherland [Su57] has conjectured that the visual system of the octopus contains a neural network that works in this manner with only *two* axes of projection, the vertical and the horizontal. The retina of the octopus thus favors discrimination along the vertical and horizontal direc-

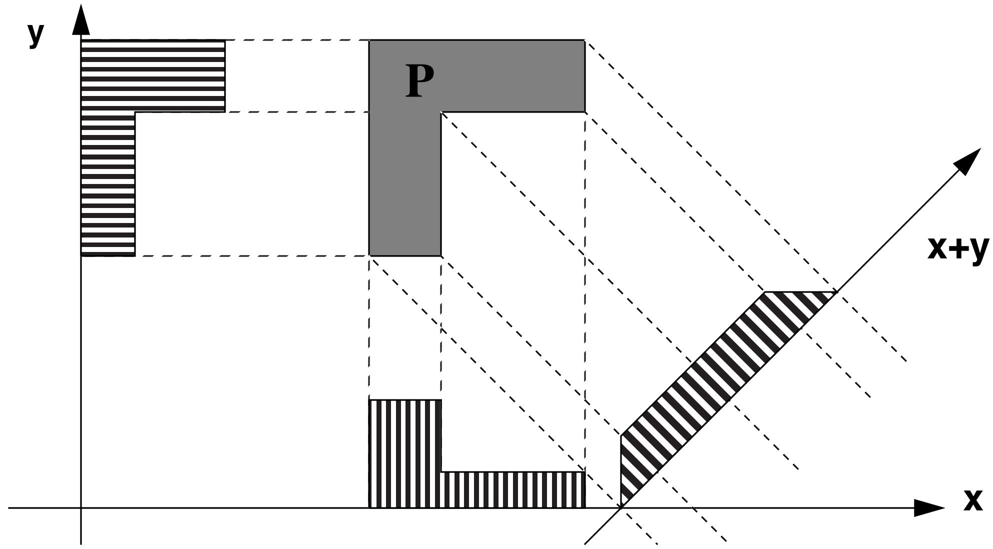


Fig. 2.1 A pattern P and three of its *collapsing projections* on the x, y and $x+y$ axes.

tions [MMS63].

3. Affine Transformations

Consider designing a character recognition system for machine printed characters. Two characters such as A and A are still considered A's and therefore quite similar even though one is bigger than the other. Such patterns would still be considered A's even if they differed in location, stretching, squeezing, and shearing. Rotation is a rather special transformation and depending on the context it may or may not be desirable to treat patterns that are invariant under rotation as belonging to the same pattern class. This may result, for example, in the system's inability to discriminate between M and W or p and d. Transformations such as these are special cases of *affine* transformations and are treated in detail in [Ry86]. Let $\mathbf{R}^2$ denote the two-dimensional Euclidean plane.

Definition: A mapping T of $\mathbf{R}^2$ onto $\mathbf{R}^2$ is called an *affine* transformation if there exists an invertible 2 by 2 matrix ***A*** and a vector ***b*** such that, for all points $p(x,y)$ in $\mathbf{R}^2$, $T(p(x,y)) = \boldsymbol{A}p + \boldsymbol{b}$.

Let us examine this definition in more detail. We have that:

$$\begin{bmatrix} T(x) \\ T(y) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

Multiplying ***A*** by $p(x,y)$ we obtain:

$$T(x) = a_{11}x + a_{12}y + b_1$$

$$T(y) = a_{21}x + a_{22}y + b_2$$

Now it is easy to see that for $\boldsymbol{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $b_1 = b_2$ we obtain the *translations*.

For $\boldsymbol{A} = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$ and $b_1 = b_2 = 0$ we obtain the *dilations*.

For $\boldsymbol{A} = \begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix}$ and $b_1 = b_2 = 0$ we obtain *stretching and squeezing*.

For $\boldsymbol{A} = \begin{bmatrix} 1 & a_{12} \\ 0 & 1 \end{bmatrix}$ and $b_1 = b_2 = 0$ we obtain *shearing*.

4. Moments of Area

Consider now a two dimensional light intensity function $g(x,y)$ (the pattern or visual input) normalized so that the volume under the function is equal to one. For our purposes it then behaves just like a bi-variate probability density function. For such a function of two variables the *moments* are defined as follows:

$$m_{jk} = \iint x^j y^k f(x,y)\,dx\,dy$$

for j, k = 0,1,2,....

In this way we now are able to obtain a *countably* infinite number of shape measurements for describing $g(x,y)$. As in the one dimensional case, an infinite number of measurements yields an impractical system. Luckily, in practice, even if the shape of $g(x,y)$ is quite complicated not many moments are needed to adequately describe it. For a detailed description of the theory required to obtain such moments that are invariant to affine transformations as well as rotations see

[Hu62]. Here we describe the main ideas for the discrete case.

Accordingly, let the two dimensional pattern or visual scene be represented as an *n* by *n* square array of grey level pixels. The pixels are assumed to have unit height and width and therefore their grey level can be specified by their integer coordinates x_t, y_t, t=1, 2,..., *n*. The *jk-th* moments therefore become:

$$M_{jk} = \sum_{l=1}^{n} \sum_{m=1}^{n} (x_l)^j (y_m)^k f(x_l, y_m)$$

for j, k = 0, 1, 2, 3, etc.

If we assume further for simplicity that the pixels are binary valued and take on the values zero and one then we have:

$$M_{jk} = \sum_A (x)^j (y)^k$$

where the summation extends over all the elements in *A*, i.e., all the "black" pixels or "ones" in the array. Note that we have dropped the subscripts from x_t, y_t, so that we have just plain *x* and *y* representing the coordinates of the pixels.

With this framework it is now possible, by suitably combining different moments, to design (define) shape features or measurements that are invariant to certain affine transformations. Such features however are not invariant to some typical kinds of noise such as "salt-and-pepper" noise.

From the previous equation we see that M_{00} = the area of the pattern, i.e., the number of black pixels. Let us denote the area by *A*. The *center-of-gravity* (also called *average* or *arithmetic mean*) of the pattern is given by $\bar{x} = M_{10}/M_{00}$ and $\bar{y} = M_{01}/M_{00}$. The moments become *invariant to translation* when they are computed with respect to the center of gravity, i.e.,

$$\overline{M}_{jk} = \sum_A (x - \bar{x})^j (y - \bar{y})^k$$

These moments need not be computed from "scratch" if the fundamental moments M_{jk} and other lower order moments are known. The reader should verify for example that $\overline{M}_{20} = M_{20} - (M_{10})^2/M_{00}$ and $\overline{M}_{02} = M_{02} - (M_{01})^2/M_{00}$.

Consider the *standard deviation* in the *x* direction of a set of black pixels:

$$\sigma_x = \sqrt{\frac{1}{M} \sum_{i=1}^{M} (x_i - \bar{x})^2}$$

Since $M = M_{00}$ it follows that:

Analogous expressions hold for the standard deviation in the *y* direction. Therefore the mo-

$$\sigma_x = \sqrt{\frac{\overline{M}_{20}}{M_{00}}}$$

ments can be made invariant to *stretching, squeezing and dilation* by computing the following:

$$\hat{M}_{jk} = \sum_A \left(\frac{x - \bar{x}}{\sigma_x}\right)^j \left(\frac{y - \bar{y}}{\sigma_y}\right)^k$$

For efficient algorithms for computing moments see [JB91] and [LS91] as well as [Za87]. For efficient parallel algorithms see [Ch90] but with a corrected complexity analysis in [Pa91].

Another very efficient approach to computing the area moments of a shape is to approximate the boundary of the shape by a polygon and to compute the area moments using simple formulae that depend only on the coordinates of the vertices of the approximating polygon [St86], [Bo89], [Le91]. Typically the number of such vertices is far fewer than the number of pixels in the original shape that are used in standard methods.

5. Moments of Perimeter

In the moments of area all the pixels of a pattern are used in the computation. However, it is reasonable to suspect that since discrimination information lies more on the boundary of a pattern than in its interior portions [AA56], satisfactory results may be obtained by using only pixels on the perimeter or boundary of the shape. Indeed psychologists have determined that the second, third and fourth moments of the perimeter of a shape are powerful predictors of discrimination performance [Zu70]. Accordingly, we may compute the moments of a spatially differentiated picture $g'(x, y)$ (for example the magnitude of the gradient of the picture) instead of the original picture $g(x,y)$ to obtain the moments of perimeter:

$$m_{pq} = \int x^p y^q [g'(x, y)] dx dy$$

Alternately, if the gradient has been thresholded to a binary picture in which the black pixels are the resulting boundary pixels, the moments can then be computed using only the black pixels resulting in considerable savings in computation since the number of computations is now not proportional to the size of the picture nor the area of the pattern but rather the perimeter of the pattern.

If the shape is given as a simple polygon or a chain code then efficient algorithms for computing moment invariants can be found in [Ch93], [BF86], [SNA90] and [Si93].

6. Moments for Preprocessing

For a method of normalizing hand printed characters see [Ca70].

7. Generalizations of π as Predictors of Discrimination Performance

In a recent paper Ball [Ba73] proposed and investigated a generalization of the notion of π. From the observation that π signifies the ratio of the circumference to diameter of a circle, and the fact that this ratio is independent of the size of a circle, he defined a generalization of π for an arbitrary shape as

$$\mathrm{pi}_1(\text{shape}) = \text{perimeter of shape} / \text{width of shape},$$

where the width of a shape he defined as the maximum value of the distance between two points on its boundary (this is more commonly termed the *diameter* of the shape). To obtain the pleasing property that pi_1(shape) takes on a maximum value for a circle, Ball had to restrict himself to convex shapes.

Other generalizations of π, some of which have the nice property of taking one of their extreme values for the circle and which are not restricted to convex shapes, are possible [To74]. One such generalization is given by

$$\mathrm{pi}_2(\text{shape}) = P^2/4A,$$

where P and A are the perimeter and the area of the shape, respectively. Two additional generalizations are given by:

$$\mathrm{pi}_3(\text{shape}) = A/2I \text{ and}$$

$$\mathrm{pi}_4(\text{shape}) = P/2(2I)^{1/2},$$

where I is the moment of inertia of the shape, assumed to have unit mass, taken about a perpendicular axis through its center of mass. For a circle $I=A/2\pi$ and hence pi_3(shape) and pi_4(shape) reduce to π. These measures find application to psychology [AA56], [Zu65], [Zu70] and in machine classification of shapes [Al62].

In the psychology literature these measures are known as *gestalt variables*, i.e., variables which do not provide a description from which the shape can be re-constructed, but which do abstract important information or characteristics of the shape as a whole. The gestalt variable $\mathrm{pi}_2(\text{shape}) = P^2/4A$ has been attractive in ordering shapes along a compactness-dispersion continuum, not only because it is size-invariant, but because it may also be easily transformed in various ways to suit the user.

The measures pi_3(shape) and pi_4(shape) belong to a much larger class of moment invariants which have been investigated by researchers interested in perception by animals, humans and machines. In psychology they are used as predictors of discrimination performance [Zu65]. In fact, an early theory of human visual perception put forward by Hofstaetter [Ho39] held that discrimination of shape was a result of direct computation of the various moments of the shape by some neural network in the visual cortex.

8. References

[AA56] Attneave, F. and Arnoult, M. D., "The quantitative study of shape and pattern perception," *Psychological Bulletin*, vol. 53, 1956, pp. 452-471.

[AD85] Abu-Mostafa, Y. S. and Psaltis, D., "Image normalization by complex moments," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. PAMI-7, January 1985, pp. 46-55.

[Al62] Alt, F. L., "Digital pattern recognition by moments," *Journal of the ACM,* vol. 9, 1962, pp. 240-258.

[Ba73] Ball, D. G., "A generalization of π," *Mathematical Gazette*, December 1973, pp. 298-302.

[BF86] Bamieh, B. and de Figueiredo, R., "A general moment invariants/attributed-graph method for three dimensional object recognition from a single image," *IEEE Journal of Robotics and Automation*, vol. RA-2, 1986, pp. 31-41.

[Bo70] de Bono, E., *Lateral Thinking: Creativity Step by Ste*p, Harper & Row, 1970.

[Bo71] de Bono, E., *The Use of Lateral Thinking*, Pelican, 1971.

[Bo78] de Bono, E., *Teaching Thinking*, Pelican Books, 1978.

[Bo89] Bockman, S. F. "Generalizing the formula for areas of polygons to moments," *American Mathematical Monthly,* February 1989, pp. 131-132.

[Ca70] Casey, R. G., "Moment normalization of hand printed characters," *IBM Journal of Research and Development,* September 1970, pp. 548-557.

[Ch90] Chen, K., "Efficient parallel algorithms for the computation of two-dimensional image moments," *Pattern Recognition*, vol. 23, 1990, pp. 109-119.

[Ch93] Chen, C.-C., "Improved moment invariants for shape discrimination," *Pattern Recognition,* vol. 26, No. 5, 1993, pp. 683-686.

[DBM83] Dudani, S. A., Breeding, K. J. and McGhee, R. B., "Aircraft identification by moment invariants," *IEEE Transactions on Computers,* vol. C-26, January 1983, pp. 39-45.

[HC78] Hogg, R. V. and Craig, A. T., *Introduction to Mathematical Statistics*, Macmillan Publishing Co., Inc., 1978.

[Ho39] Hofstaetter, P. R., "Ueber die Scherzung von Gruppeneigenschaften," *Z. Psychol.,* vol. 145, 1939, pp. 1-44.

[Hu62] Hu, M. K., "Visual pattern recognition by moment invariants," *IRE Transactions on In-*

formation Theory, vol. IT-8, February 1962, pp. 179-187.

[JB91] Jiang, X. Y. and Bunke, H., "Simple and fast computation of moments," *Pattern Recognition*, vol. 24, No. 8, 1991, pp. 801-806.

[KH90] Khotanzad, A. and Hong, Y. H., "Rotation invariant image recognition using features selected via a systematic method," *Pattern Recognition,* vol. 23, No. 10, 1990, pp. 1089-1101.

[Le91] Leu, J.-G., "Computing a shape's moments from its boundary," *Pattern Recognition,* vol. 24, No. 10, 1991, pp. 949-958.

[LS91] Li, B.-C. and Shen, J., "Fast computation of moment invariants," *Pattern Recognition*, vol. 24, No. 8, 1991, pp. 807-813.

[MMS63] Mackintosh, N. J., Mackintosh, J. and Sutherland, N. S., "The relative importance of horizontal and vertical extents in shape discrimination by octopus," *Animal Behavior,* vol. 11, 1963, pp. 355-358.

[NT70] Nagy, G. and Tuong, N., "Normalization techniques for hand printed numerals," *Communications of the ACM,* vol. 13, No. 8, August 1970, pp. 475-481.

[Pa91] Pan, Y., "A note on efficient parallel algorithms for the computation of two-dimensional image moments," *Pattern Recognition,* vol. 24, No. 9, 1991, p. 917.

[Ry86] Ryan, P. J., *Euclidean and Non-Euclidean Geometry: An Analytic Approach*, Cambridge University Press, Cambridge, 1986.

[Si93] Singer, M. H., "A general approach to moment calculation for polygons and line segments," *Pattern Recognition,* vol. 26, No. 7, 1993, pp. 1019-1028.

[SNA90] Strachan, N. J. C., Nesvadba, P. and Allan, A. R., "A method for working out the moments of a polygon using an integration technique," *Pattern Recognition Letters,* vol. 11, 1990. pp. 351-354.

[St86] Stone, M. G., "A mnemonic for areas of polygons," *American Mathematical Monthly,* vol. 93, 1986, pp. 479-480.

[Su57] Sutherland, N. S., "Visual discrimination of orientation by octopus," *British Journal of Psychology,* vol. 48, 1957, pp. 55-71.

[To74] Toussaint, G. T., "Generalisations of π: some applications," *The Mathematical Gazette*, December 1974, pp. 291-293.

[Za87] Zakaria, M. F., Vroomen, L. J., Zombor-Murray, P. J. A. and Kessel, J. M. H. M., "Fast algorithm for the computation of moment invariants," *Pattern Recognition*, vol. 20, No. 6, 1987, pp. 639-644.

[Zu65] Zusne, L., "Moments of area and of the perimeter of visual form as predictors of discrimination performance," *Journal of Experimental Psychology*, vol. 69, 1965, pp. 213-220.

[Zu70] Zusne, L., *Visual Perception of Form*, Academic Press, New York, 1970.