Moment-Based Image Normalization With High Noise-Tolerance

Matthias Gruber and Ken-Yuh Hsu

Abstract—In this paper the effects of noise with nonzero mean on existing moment-based image normalization methods are studied. Several modifications to reduce noise sensitivity are presented and tested. They involve nonlinear mapping and fractional- and negative-order moments.

Index Terms—Moments, image normalization, image centroid, pattern recognition, noise suppression.

1 Introduction

THE normalization of images with respect to position, orientation, and scale can be used as a preprocessing step in pattern recognition procedures to limit the range of variations within different classes of patterns thus allowing the use of classification criteria that are sharper and need not be invariant under geometric transformations of the processed data. Such a normalization step can be implemented on the basis of moments.

The moment-concept has been introduced to pattern recognition by Hu [1]. Since then, a variety of new moment-types and moment-based methods have been developed and used [2], [3], [4], and [5]. Moments are attractive because their computation is algorithmically simple and uniquely defined for any image-function; it can be carried out in parallel and therefore very fast, and, since moments are global quantities, all available information is used making moment-based methods less vulnerable to losses or changes of pattern details than methods that use (few) particular features.

However, moments become very noise-sensitive with increasing order [5]. Hence, the lowest possible orders should be used in a moment-based procedure. As for image-normalization, the classical method (in [3]), here referred to as Method "I," involves (integral) moments up to an order of two; a modified and more robust algorithm involving only zero- and first-order moments (Method "II") has been presented in [4]. We will describe how a nonlinear mapping and fractional- and negative-order moments can be used to further reduce the noise-sensitivity.

2 Existing Methods

Moment-based image normalization methods can be formulated in a very simple and transparent way by means of ordinary and rotational moments which are defined by

$$M_{pq} = \int\int x^p y^q f(x, y) dx dy \qquad (1)$$

and

$$R_{st} = \int\int r^s e^{-it\varphi} f(r, \varphi) r dr d\varphi \qquad (2)$$

respectively, with $p, q, s, t \in N_0$ as order indices, and (x, y) and (r, φ) as Cartesian and polar coordinates; f shall be a non-negative continuous image function with bounded support so that integration within the available image area, defined as $|x|, |y| \leq 1$ for geometric and $r \leq 1$ for rotational moments, is sufficient to gather all signal information.

The two existing methods mentioned earlier adjust the coordinate system in a first step by moving its origin into the classical centroid $(\xi_1, \eta_1) = (M_{10}/M_{00}, M_{01}/M_{00})$ of the image, which constitutes a unique reference-point. Subsequent operations are then carried out in the adjusted coordinate system.

Expressions to determine orientation and scale exploit the particular changes that rotational moments experience under rotations $f(\varphi) \rightarrow f(\varphi + \phi)$ and scale changes $f(r) \rightarrow f(kr)$ of an image.

$$R_{st}(f(\varphi + \phi)) = e^{it\phi} \cdot R_{st}(f(\varphi)) \qquad (3)$$

$$R_{st}(f(kr)) = k^{-s-2} \cdot R_{st}(f(r)) \qquad (4)$$

The standard orientation is defined as the one at which a certain rotational moment with non-zero repetition t is real positive, and the angle ϕ by which a given image is rotated from it is obtained by

$$\phi = \frac{1}{t} \arctan\left(\frac{\Im(R_{st})}{\Re(R_{st})}\right), \qquad t \neq 0 \qquad (5)$$

Methods "I" and "II" use moments R_{22} and R_{02}, respectively in (5). As for scale, the factor k by which a given image is magnified with respect to its normalized version can be derived from (4); assuming that $f(r)$ is the normalized version, we obtain

$$k = \kappa \cdot \left(\frac{R_{i0}}{R_{j0}}\right)^{\frac{1}{j-i}} \qquad (6)$$

with κ being a constant that is composed of moments of $f(r)$. Method "II" uses the most simple form of (6) involving R_{00} and R_{10} while Method "I" uses R_{00} and R_{20}.

After computing ϕ and k the image is normalized through a rotation of $-\phi$, a dimensional scaling by a factor $1/k$, and a shift of all image points that moves the origin back into the center of the image. The whole procedure is unaffected by intensity-changes $f \rightarrow c \cdot f$, since they cause a common factor in all moments which is canceled in above fractional terms.

3 Effects of Noise

Above methods yield perfect results only under ideal noise-free conditions. In the presence of noise the involved moments suffer errors which in turn falsify the normalization. ϕ and k are thereby affected directly through (5) and (6), and indirectly through changes of the centroid position.

Earlier studies of noise effects [4], [5] using models of uncorrelated random disturbances with mean zero have shown that errors increase with increasing moment-orders and that the error in ϕ is unbiased up to quadratic-order expansion terms. However, zero-mean noise models are often unrealistic, especially when dealing with non-negative image functions.

• *The authors are with Electro-Optical Engineering, National Chiao Tung University, 1001 Ta Hsueh Road Hsinchu, Taiwan, R.O.C.*
M. Gruber can be contacted at FernUniversitaet Hagen Optische Nachrichtentechnik Feithstr. 140 58084 Hagen Germany,
E-mail: matthias.gruber@FernUni-Hagen.de

Manuscript received May 10, 1995; revised June 24, 1996. Recommended for acceptance by R. Chin.
For information on obtaining reprints of this article, please send e-mail to: transpami@computer.org, and reference IEEECS Log Number P96078.

To address this problem it is sufficient to study the disturbing effects of an additive constant ν on the normalization process. The additive approach allows a separation of moments into signal and noise components: $R_{st} = S_{st} + N_{st}$ and $M_{pq} = S'_{pq} + N'_{pq}$. After integration N'_{pq} becomes

$$N'_{pq} = \begin{cases} \dfrac{4\nu}{(p+1)(q+1)} & \text{for } p, q \text{ even} \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

and we note that expressions M_{10} / M_{00} and M_{01} / M_{00} for the centroid are only affected in the denominators which pick up a term $N'_{00} = 4\nu$. This amounts to a down-scaling by a factor $M_{00} / (M_{00} + 4\nu)$ compared to an image without bias ν. In other words, the centroid is located closer to the image center. Since there is no simple general expression for N_{st} in a shifted coordinate-system, we use instead the term for the centered one

$$N_{st} = \begin{cases} \dfrac{2\pi\nu}{s+2} & \text{for } t = 0 \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

which is a good approximation for the relevant moment-orders.

According to (5), (6), and (8), only k is directly affected by a bias. To quantify this effect we need an estimate for S_{s0} which shall be computed from an averaged signal function: We assume that the average of all centered but otherwise unnormalized signals of the *same* size $r_{sig.}$ (radius of the smallest circle around the origin that completely covers the signal) is a binary function with a certain non-zero value inside the $r_{sig.}$-circle and 0 outside. Assuming further that all sizes up to $r_{sig.} = \rho$ occur with equal probability, the general averaged signal is the average of all these binary functions up to radius ρ, i.e., a cone-shaped function with radius ρ and a certain height σ. Then S_{s0} becomes

$$S_{s0} = \frac{2\pi\sigma \cdot \rho^{s+3}}{(s+2)(s+3)} \tag{9}$$

and

$$k = \kappa \cdot \left(\frac{S_{i0} + N_{i0}}{S_{j0} + N_{j0}} \right)^{\frac{1}{j-1}} \tag{10}$$

Table 1 lists several numerical values of k for different signal-to-noise-ratios S_{00} / N_{00}; thereby $\rho = 0.5$ and κ is set so that $k = 1$ in the noise free case. Apparently, the induced error is sizable and gets larger for higher order moments. An amount that still allows correct recognition of the signal is often already surpassed for comparatively weak noise levels. Therefore, an efficient noise-suppression is indispensable.

Method "II" performs a clipping prior to all further operations. If there is a gap between the intensities of the signal and those of the background noise (i.e., noise at locations without significant signal-intensity) as it is usually the case for binary signals, and if the clipping-threshold is set properly in between, background-noise will be perfectly suppressed. In the analog case, however, significant parts of the signals often have intensities comparable to those of the background-noise, which means, noise will be left and / or parts of the signal lost after a clipping. Hence, values ξ_1, η_1, ϕ and k can be significantly changed. Even more problematic: The threshold needs to be adjusted for every image to be normalized, and this requires information which is usually unknown. This problem is avoided in the noise suppression scheme we present now.

4 Modifications

4.1 Nonlinear Mapping

For analog signals we propose as first operation a nonlinear mapping that, rather than completely eliminating low intensities, weights them instead overproportionally weaker than higher ones. This will enhance the total weight of the signal relative to that of the background-noise and can be implemented without a threshold. In addition, if the mapping is of type $f \to f^m$, $m > 1$ it preserves the invariance of the normalization with respect to absolute intensities.

How should m be chosen? The larger m is, the better will background noise be suppressed, but at the same time, the stronger will signal-noise (i.e., noise at locations with significant signal-intensity) be amplified. So, there must be an optimum m with minimum overall error. Its value depends on the extent with which these two types of disturbances occur.

4.2 Fractional- and Negative-Order Moments

Following an idea by Reddi [6] we will now generalize the moment-concept to include moments with fractional and negative orders and investigate their usefulness for image normalization. We focus on generalized rotational moments, which shall be defined as in (2) but with real-valued s. Such moments generally exist for $s > -2$.

4.2.1 Coordinate-Origin

Methods "I" and "II" adjust the coordinate-system by shifting its origin into the centroid (ξ_1, η_1). This point can also be defined through the condition $R_{11} = 0$ and approached asymptotically through a stepwise shifting of the coordinate-system along the x- and y-axis by successively decreasing distances in directions opposite to the signs of the current values of $\Re(R_{11})$ and $\Im(R_{11})$, respectively. Using the generalized condition

$$R_{s1} = 0, \text{ with } s \in R \tag{11}$$

this iterative method still defines a reference-point, denoted (ξ_s, η_s), whose relative position is invariant with respect to orientation and scale of the signal. We show now that it is unique for

$$0.18 \approx 3 - 2 \cdot \sqrt{2} < s < 3 + 2 \cdot \sqrt{2} \approx 5.82 \tag{12}$$

R_{s1} consists of a real and an imaginary part, X_s and Y_s, which have to be 0 simultaneously to satisfy (11). Expressed as functions of a shift (ξ, η) of the origin, X_s and its partial derivatives with respect to ξ and η read

$$X_x(\xi, \eta) = \int\int x r^{s-1} f(x+\xi, y+\eta) dxdy \tag{13}$$

$$\frac{\partial}{\partial \xi} X_s(\xi, \eta) = -\int\int \frac{sx^2 + y^2}{r^{3-s}} f(x+\xi, y+\eta) dxdy \quad \leq 0 \quad \forall \xi, \eta, f \tag{14}$$

$$\frac{\partial}{\partial \eta} X_s(\xi, \eta) = -\int\int \frac{(s-1)xy}{r^{3-s}} f(x+\xi, y+\eta) dxdy \tag{15}$$

X_s is certainly positive / negative when the signal is completely located in the half-plane with pos. / neg. x. Since it decreases monotonously when the origin is shifted in x-direction (cf.,(14) the value 0 is taken on exactly one time for any (fixed) η. Hence, the set of points ($X_s(\xi, \eta) = 0$) is a bounded continuous function of η. Similarly, ($Y_s(\xi, \eta) = 0$) constitutes a bounded continuous function of ξ and both curves must have at least one finite intersection point.

A sufficient condition that there are no further ones is that curves ($X_s = 0$) and ($Y_s = 0$) run nowhere steeper than 45° against the η- and ξ-axis, respectively. For X_s this is equivalent to the demand $|\partial/\partial\xi_s(X_s)| \geq |\partial/\partial\eta_s(X_s)|$, and if the relation is to hold for any f it must be satisfied for the weighting kernels of the involved integrals i.e.,

$$\left|\frac{sx^2 + y^2}{r^{3-s}}\right| > \left|\frac{(1-s)xy}{r^{3-s}}\right| \tag{16}$$

which is equivalent to (12). For Y_s we obtain the same relation.

In practice, the iterative approach towards (ξ_s, η_s) has to be halted after a finite number of steps, usually when $|X_s|$ and $|Y_s|$ fall below a certain acceptance threshold t_a. This introduces a residual error of coordinates ξ_s, η_s, which can be significant despite a small t_a if ($X_s = 0$) and ($Y_s = 0$) intersect at a very sharp angle. To avoid such a problem we limit s to $0.33 < s < 5.67$ in which case the intersection angle cannot be sharper than 30°.

For the large majority of images where $r_{sig.}$ is significantly smaller than one, the reference-point location will be more noise resistant for small s because then the marginal zone of the image, which carries only noise, has a lower relative weight in integral (2).

4.2.2 Scale-Factor

According to (10) and Table 1, the bias related error of k is order-dependent. Now we look for the real-valued i, j that minimize the error for the averaged signal defined earlier. If N_{i0}, $N_{j0} \neq 0$ the value of k can only remain unchanged if

$$\frac{N_{i0}}{N_{j0}} = \frac{S_{i0}}{S_{j0}} \tag{17}$$

Inserting terms (8) and (9) this condition becomes $\rho^{i-j} = (i+3)/(j+3)$, which, after substituting $i = u + v - 3$, $j = v - 3$, can be transformed to

$$\rho^u = 1 + \frac{u}{v} \tag{18}$$

Equation (18) can be interpreted as an intersection point of functions ρ^u and $1 + u/v$. Besides the trivial solution $u = 0 \Leftrightarrow i = j$, which cannot be used in (10), there can only be another intersection point if $v < 0$. A negative v also lets the two functions intersect at a sharper angle so that (18) is approximately fulfilled for a wider range of u-values.

Based on these considerations we expect normalization errors to be lower for negative i, j. This suggestion is also supported by the fact that in moments with negative radial orders the region around the origin, where the signal information is concentrated, is weighted higher than the marginal regions, which carry only noise.

TABLE 1
INCREASE OF K DUE TO A BIAS

S_{00}/N_{00} → method ↓	50	25	12.5	6.25	2
"I"	1.033	1.064	1.124	1.230	1.556
"II"	1.013	1.026	1.051	1.099	1.283

5 PERFORMANCE TESTS

To substantiate our earlier statements we present now some test results of the proposed modifications. In these tests

- the *signals* are images of the ten numerals. We use low-pass-filtered versions of binary raw-patterns to obtain intensity histograms similar to those found in images from CCD-cameras.
- *Signal-noise* (symbolized by S) is created by modulating the signals with slowly varying random disturbances to mimic the effect of non-uniform illumination, a real-world S-noise source. It is quantified through the modulation-factor *mod* which indicates (in percentages) the maximum possible intensity change-rate in both direction.
- *Background-noise* (B) consists of a few randomly distributed speckles with a maximum intensity of 85 % of the maximum signal intensity, and it is quantified through the signal-to-noise-ratio (SNR) S_{00}/N_{00}.

In an experiment to determine the optimum m for the nonlinear mapping we generated randomly positioned signals of fixed orientation and size $r_{sig.}$, added S-noise (mod = 15) and / or B-noise (SNR = 6), and computed the statistical errors of ξ_s, η_s, ϕ and k (using orders $s = 0.33$, $i = -0.75$, $j = -0.25$). Table 2 lists the mean distance between noisy and noise-free reference-points, and the variances of k and ϕ for different m. As expected, errors due to S-noise increase with increasing m while errors due to B-noise decrease. If both types of noise occur, a certain finite m (here $m = 3$) yields the lowest overall error.

TABLE 2
NORMALIZATION ERRORS FOR DIFFERENT M DUE TO SIGNAL- AND/OR BACKGROUND-NOISE

mean error of ref.-point location (in % of $r_{sig.}$)

m→ noise-type ↓	1	2	3	4	5
S	1.92	3.98	6.22	8.52	10.94
B	13.30	6.56	4.12	2.96	2.36
S+B	13.30	7.48	7.24	8.86	11.06

$\text{Var}(k) \times 10^{-2}$

S	0.6	1.6	2.8	4.1	5.7
B	5.9	4.1	3.4	3.0	2.8
S+B	6.0	4.5	4.4	5.1	6.2

Var(ϕ) (in deg.)

S	1.09	2.01	2.83	3.53	4.16
B	18.45	7.59	4.46	3.07	2.37
S+B	18.46	7.61	4.89	4.48	4.68

For details, see text.

TABLE 3
MEAN DISTANCE BETWEEN NOISY AND NOISE FREE (ξS, ηS) FOR DIFFERENT S DUE TO BACKGROUND NOISE

absolute mean distance (in % of $r_{sig.}$)

mod → s↓	128	32	8	2	1	0.5
1	2.2	8.6	32.1	94.7	148	203

relative mean distance

1	1.0	1.0	1.0	1.0	1.0	1.0
0.5	0.45	0.49	0.49	0.61	0.70	0.82
0.33	0.37	0.37	0.41	0.51	0.55	0.71
0.25	0.34	0.35	0.35	0.42	0.54	0.67

For details see text.

TABLE 4
MEAN DISTANCE BETWEEN NOISY AND NOISE FREE (ξS, ηS) FOR DIFFERENT S DUE TO SIGNAL NOISE

absolute mean distance (in % of $r_{sig.}$)

mod → s↓	10	21	33	46	61	77
1	0.9	2.0	2.9	4.3	6.2	7.5

relative mean distance

1	1.0	1.0	1.0	1.0	1.0	1.0
0.5	1.28	1.26	1.29	1.24	1.24	1.27
0.33	1.41	1.43	1.46	1.46	1.41	1.43
0.25	1.50	1.48	1.52	1.48	1.50	1.49

For details see text.

In order to quantify the noise-sensitivity of (ξ_s, η_s) signals with size $r_{sig.}$ = 1/4 and random orientation were generated and then shifted by 3/4 from their normalized positions in random directions. After adding noise we determined the noisy reference-points and their distances from the noise-free ones. Tables 3 and 4 show the mean distances for B- and S-noise, respectively. For small *s*, errors due to S-noise increase slightly but the effect is by far overcompensated through the decrease of the error due to B-noise. Hence, small *s* provide significantly higher robustness.

In an experiment to find the moment-orders *i*, *j* that minimize the error of *k* we added B-noise (SNR = 10) to normalized signals and computed mean and variance of *k*. The results (Table 5) indicate that small negative orders provide the highest noise-tolerance.

TABLE 5
MEAN E(K) AND VARIANCE VAR(K) OF K FOR DIFFERENT ORDERS I, J IN THE PRESENCE OF BACKGROUND NOISE

i	0	0	–0.5	–0.75	–1	–2
j	2	1	0	–0.25	–0.5	–1.5
E(*k*)	1.138	1.102	1.051	1.034	1.029	0.999
Var(*k*)	0.068	0.049	0.035	0.028	0.032	0.052

We conclude that a normalization algorithm with high noise-tolerance should use a mapping $f \rightarrow f^3$, moment $R_{-\frac{1}{3}0}$ to define the reference-point, and moments $R_{-\frac{1}{4}0}$ and $R_{-\frac{3}{4}0}$ to compute scale-factor *k*.

ACKNOWLEDGMENT

This work was supported by the National Science Council of Taiwan, Republic of China (NSC 83-0417-E-009-011) and the Deutsche Forschungsgemeinschaft of Germany (DFG 446 TAI-113). The authors are grateful for Prof. A.W. Lohmann's helpful comments.

REFERENCES

[1] M. Hu, "Visual Pattern Recognition by Moment Invariants," *IRE Trans. Info. Theory*, vol. 8, pp. 179–187, 1962.

[2] M. Teague, "Image Analysis via the General Theory of Moments," J.O.S.A., vol. 70, no. 8, pp. 920-930, 1980.

[3] Y. Abu-Mostafa and D. Psaltis, "Image Normalization by Complex Moments," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 7 no. 1, pp. 46–55, 1985.

[4] Y. Sheng and H. Arsenault, "Noisy-Image Normalization Using Low-Order Radial Moments of Circular Harmonic Functions," J.O.S.A., vol. 4 no. 7, pp. 1,176–1,184, 1987.

[5] C. Teh and R. Chin, "On Image Analysis by the Methods of Moments," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 10, no. 4, pp. 496–513, 1988.

[6] S. Reddi, "Radial and Angular Moment Invariants for Image Identification," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 3 no. 2, pp. 240–242, 1981.

0162-8828/97$10