Robust Histogram Estimation Under Gaussian Noise

Jitka Kostková(✉) and Jan Flusser

The Czech Academy of Sciences, Institute of Information Theory and Automation, Pod Vodárenskou věží 4, 182 08 Prague 8, Czech Republic
{kostkova,flusser}@utia.cas.cz

Abstract. We present a novel approach to description of a multidimensional image histogram insensitive with respect to an additive Gaussian noise in the image. The proposed quantities, although calculated from the histogram of the noisy image, represent the histogram of the original clear image. Noise estimation, image denoising and histogram deconvolution are avoided. We construct projection operators, that divide the histogram into non-Gaussian and Gaussian part, which is consequently removed to ensure the invariance. The descriptors are based on the moments of the histogram of the noisy image. The method can be used in a histogram-based image retrieval systems.

Keywords: Gaussian additive noise · Multidimensional histogram · Invariant characteristics · Moments · Projection operator

1 Introduction

Real images are often corrupted by noise, which not only degrades their visual appearance but also significantly changes all quantitative descriptors. If the signal-to-noise ratio is low, the corruption may be so heavy that it is very difficult to deduce anything about the original scene from the acquired image.

In this paper, we pay attention to the influence of the noise on the image histogram. Histogram provides statistics of graylevel/color frequencies and has become a simple, yet powerful descriptor for image classification. Histogram has established itself as a meaningful image characteristic for content-based image retrieval (CBIR) [8,12,14] because histogram similarity is a salient property for human vision. Two images with similar histograms are mostly perceived as similar even if their actual content may be very different from each other. On the other hand, those images that have substantially different histograms are rarely rated by observers as similar. Another attractive property of the histogram is

This work has been supported by the Czech Science Foundation (Grant No. GA18-07247S), by the *Praemium Academiae*, and by the Grant SGS18/188/OHK4/3T/14 provided by the Ministry of Education, Youth, and Sports of the Czech Republic (MŠMT ČR).

M. Vento and G. Percannella (Eds.): CAIP 2019, LNCS 11678, pp. 421–432, 2019.
https://doi.org/10.1007/978-3-030-29888-3_34

that it does not depend on image translation, rotation and (if normalized to the image size) on scaling. Simple preprocessing can also make the histogram insensitive to linear changes of the contrast and brightness of the image.

The CBIR methods based on comparing histograms are sensitive to noise in the images. Additive noise leads to a histogram smoothing, the degree of which is proportional to the amount of noise (see Fig. 1 for illustration). This follows from the well-known theorem from probability theory. Given a random variable X (which represents the pixel values of the image) with probability density function (PDF) h_X (which is now the normalized histogram of the image) and additive noise N with PDF/histogram h_N, then for the PDF h_Z of noisy random variable $Z = X + N$ holds

$$h_Z(x) = (h_X * h_N)(x) = \int h_X(x - s) h_N(s) \, \mathrm{d}s, \tag{1}$$

assuming that the noise is independent of the image. The histogram smoothing immediately results in a drop of the retrieval performance because different histograms tend to be more and more similar to each other.

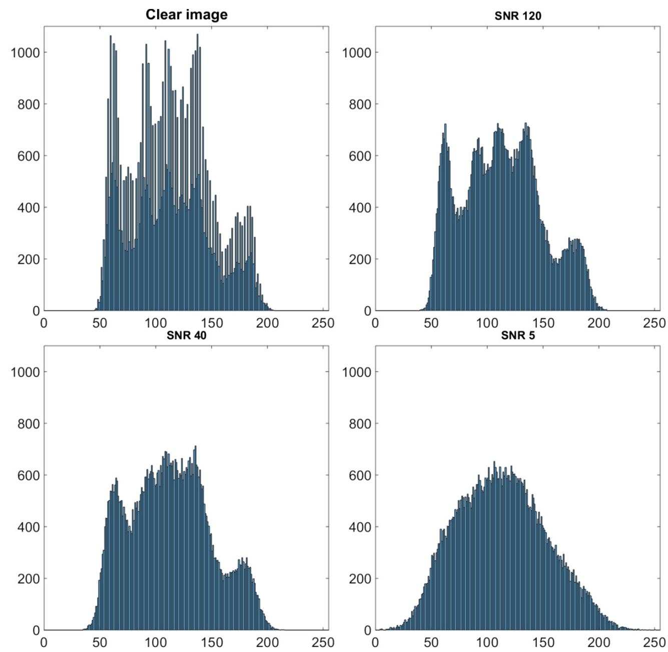

Fig. 1. Histogram smoothing due to image noise. Histogram of clear Lena image (top left). Histograms of noisy images: SNR = 120 (top right), SNR = 40 (bottom left), SNR = 5 (bottom right). For low SNR the histogram becomes unimodal and hard to distinguish from other smoothed histograms.

In digital photography, when using a CCD chip, the noise is unavoidable. It is apparent especially when taking a picture in low light using high ISO and/or long exposure time (see Fig. 2). The noise has several components. Photon shot noise, thermal noise, readout noise and background noise are the main ones. Additive noise component can be reasonably modelled by a Gaussian random variable uncorrelated with the image values. The signal-dependent component follows Poisson distribution. The method presented in this paper can handle only the former one.

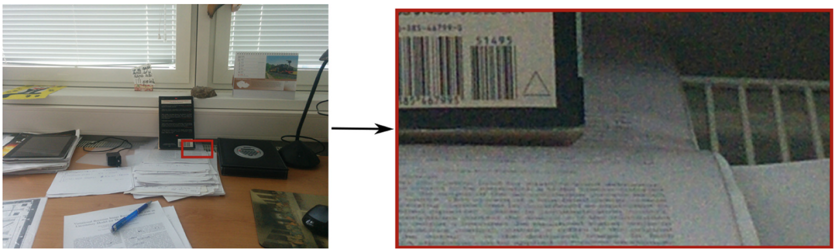

Fig. 2. A low-light scene (left). On the close-up, the noise is clearly apparent (right).

Although the noise in digital imaging is an issue we cannot ignore even in consumer photography, very little attention has been paid to developing noise-resistant histogram representation. The authors of the papers on CBIR have either skipped this problem entirely or rely on denoising algorithms applied to all images before they enter the database. A pioneer work on this field was published by Höschl and Flusser [3] who proposed a kind of *convolution moment invariants*. Their work was motivated by blur invariants applied to a different problem in the image domain [1,2,6,9,15]. The authors presented invariants w.r.t. convolution, calculated from the histogram moments. These invariants were, however, defined only for 1D histogram of a graylevel image and cannot be easily extended to multidimensional histograms of color and multispectral images.

In this paper, we present a new histogram representation, based on its moments, which is totally resistant (at least theoretically) to additive Gaussian noise. This histogram representation could be implemented in CBIR systems in the case of noisy database and/or noisy query images (see Fig. 3 for the method outline). Our method does not perform any denoising and cannot replace it in the applications where the noise should be suppressed in order to improve the visual quality of the image. Unlike [3], the proposed method works with *multidimensional* histograms, which makes it suitable for color images. Another remarkable feature is that the method does not assume an independent noise in individual channels/spectral bands.

The main idea of this paper is the following one. We introduce *projection operators*, acting on the histogram space, that divide each histogram into two

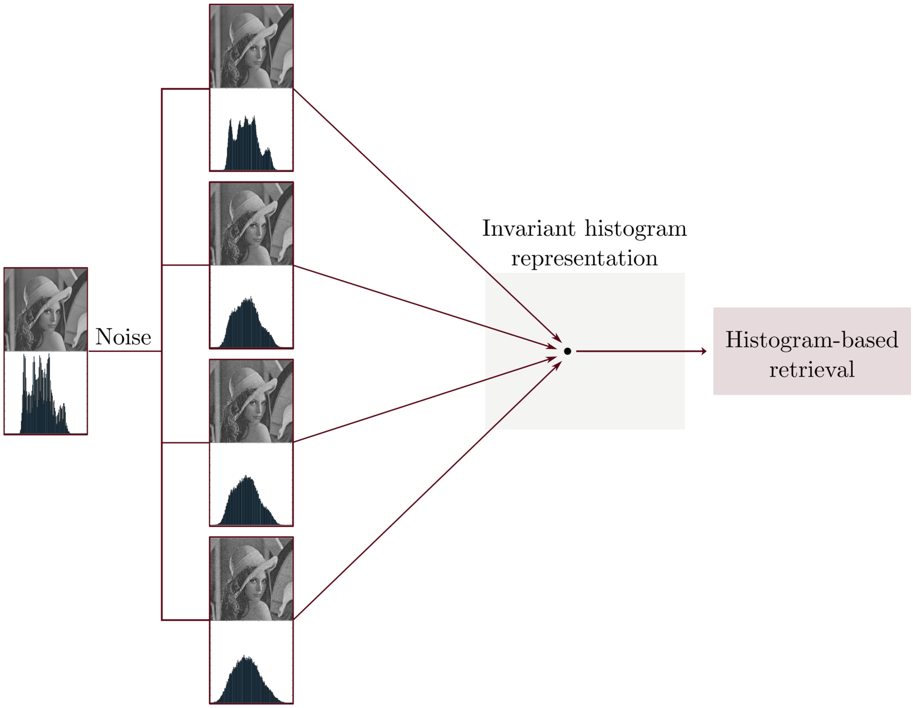

Fig. 3. The outline of the proposed method. From left to right: original image and its histogram, noisy images with smoothed histograms, representation of the histograms by invariant features (the core of the method), potential usage for noise-robust image retrieval. The actual implementation works with color images and the histograms are multidimensional.

components. Based on the known parametric form of h_N, we show that one of the components can be used to compute quantities, which are invariant with respect to convolution in Eq. (1). These quantities can be used directly to characterize h_X regardless of the amount of noise present.

2 Problem Formulation

Let $\mathbf{X}$ be an r-band original clear image with histogram $h_\mathbf{X}$ and $\mathbf{N} \sim \mathcal{N}(\mathbf{0}, \Sigma)$ be a normally distributed random noise, independent of $\mathbf{X}$, with regular covariance matrix Σ. The r-D noise histogram $h_\mathbf{N}$ has the well-known Gaussian shape

$$h_\mathbf{N}(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^r |\Sigma|}} \exp\left\{-\frac{1}{2}\mathbf{x}^T \Sigma^{-1} \mathbf{x}\right\} , \tag{2}$$

where $\mathbf{x} = (x_1, \ldots, x_r)^T$ is the vector of intensities. Under the above assumptions, histogram $h_\mathbf{Z}$ of the noisy image $\mathbf{Z} = \mathbf{X} + \mathbf{N}$ is given as

$$h_\mathbf{Z} = h_\mathbf{X} * h_\mathbf{N}. \tag{3}$$

Our aim is to design a histogram descriptor I, which is invariant with respect to the noise, i.e. we require

$$I(h_{\mathbf{X}}) = I(h_{\mathbf{Z}}) = I(h_{\mathbf{X}} * h_{\mathbf{N}}) \tag{4}$$

for any normally distributed zero-mean random noise $\mathbf{N}$ with arbitrary (unknown) covariance matrix Σ.

Complying with Eq. (4) is, however, not the only desirable property of I. At the same time, I must be *discriminable*, which means

$$I(h_{\mathbf{X}}) \neq I(h_{\mathbf{Y}}) \tag{5}$$

for any two images $\mathbf{X}$ and $\mathbf{Y}$ such that $\mathbf{Y}$ is not a noisy version of $\mathbf{X}$.

3 Construction of the Invariant

The main idea of constructing invariants to Gaussian convolution is based on projections of a histogram onto the set of all Gaussian functions and on its complement. In this way, we decompose any histogram into Gaussian and non-Gaussian components. We show that the ratio of these two components possesses the desired invariant property. In the sequel, we introduce the necessary mathematical background.

Let us denote the set of all probability density functions (normalized histograms) as $\mathcal{D}$ and the set of all zero-mean Gaussian probability density functions (including Dirac δ-function) as $\mathcal{S}$.

Lemma 1. *$\mathcal{S}$ is closed under convolution.*

Proof. It holds for any two Gaussian PDFs $h_{\mathbf{N}_1}$ and $h_{\mathbf{N}_2}$ with covariance matrices Σ_1 and Σ_2 that the result of convolution is again a Gaussian PDF

$$h_{\mathbf{N}_1} * h_{\mathbf{N}_2} = h_{\mathbf{N}}$$

with covariance matrix $\Sigma = \Sigma_1 + \Sigma_2$.

The set $\mathcal{S}$ with the operation convolution $(\mathcal{S}, *)$ is a commutative semi-group. However, $\mathcal{S}$ is not a vector space.

Let us define projection operator P that projects arbitrary PDF $h \in \mathcal{D}$ on the "nearest" Gaussian PDF in the sense of having the same covariance matrix. In other words, $P : \mathcal{D} \mapsto \mathcal{S}$ is defined as

$$P(h) = h_{\mathbf{N}}, \tag{6}$$

where $h_{\mathbf{N}}$ has the same covariance matrix as h. The operator P is well defined for all PDFs with a regular covariance matrix. It is idempotent, i.e. $P^2 = P$ and every $h \in \mathcal{D}$ can be uniquely written in the form $h = Ph + h_a$, where h_a is defined as $h - Ph$. In this way, the set $\mathcal{D}$ can be expressed as a direct sum $\mathcal{D} = \mathcal{S} \oplus \mathcal{A}$, where $\mathcal{S}$ is the range of P and $\mathcal{A}$ is the kernel. Note that P is not a projector in the common sense known from linear algebra because it is not a linear operator. Still, we call it projector because it keeps the above mentioned key properties of "algebraic" projection operators.

For our purposes, the crucial property of operator P is that it commutes with the convolution with functions from $\mathcal{S}$. This property is necessary for the construction of the invariant descriptors, as we demonstrate later.

Lemma 2. *Let $h \in \mathcal{D}$ and $h_{\mathbf{N}} \in \mathcal{S}$. Then*

$$P(h * h_{\mathbf{N}}) = Ph * h_{\mathbf{N}}. \tag{7}$$

Proof. First, we investigate the right-hand side, where we have a convolution of two Gaussians with covariance matrices Σ_h and $\Sigma_{h_{\mathbf{N}}}$, respectively. Thanks to Lemma 1, this is also a Gaussian with covariance matrix $\Sigma_h + \Sigma_{h_{\mathbf{N}}}$. On the left-hand side, $P(h * h_{\mathbf{N}})$ is by definition a Gaussian with covariance matrix $\Sigma_{h*h_{\mathbf{N}}}$. It is well known that central second-order moments of any PDF, which is a convolution of two other PDFs, are sums of the same moments of the factors. The same is true for entire covariance matrix. Hence, on the left-hand side we have $\Sigma_{h*h_{\mathbf{N}}} = \Sigma_h + \Sigma_{h_{\mathbf{N}}}$, which completes the proof.

Now we formulate the principal theorem of the paper that introduces the invariant descriptor of a probability density function as a ratio of certain characteristic functions. Characteristic function of a random variable is in fact the Fourier transform of its density [5].

Theorem 1. *Let $h \in \mathcal{D}$ and let P be the projector onto $\mathcal{S}$ defined above. Then the ratio of characteristic functions Φ of the densities h and Ph*

$$I(h)(\mathbf{u}) = \frac{\Phi(h)(\mathbf{u})}{\Phi(Ph)(\mathbf{u})} \tag{8}$$

is an invariant to convolution with a Gaussian probability density function: $I(h) = I(h * h_{\mathbf{N}})$ for any $h_{\mathbf{N}} \in \mathcal{S}$.

Proof. First of all, note that $I(h)$ is well defined for any h because both $\Phi(h)$ and $\Phi(Ph)$ exist and $\Phi(Ph)(\mathbf{u}) \neq 0$ for any $\mathbf{u}$. The proof of invariance follows from the fact that P commutes with the convolution (see Lemma 2). If $h_{\mathbf{N}} \in \mathcal{S}$, then

$$I(h * h_{\mathbf{N}}) = \frac{\Phi(h * h_{\mathbf{N}})}{\Phi(P(h * h_{\mathbf{N}}))} = \frac{\Phi(h * h_{\mathbf{N}})}{\Phi(Ph * h_{\mathbf{N}})} = \frac{\Phi(h)\Phi(h_{\mathbf{N}})}{\Phi(Ph)\Phi(h_{\mathbf{N}})} = \frac{\Phi(h)}{\Phi(Ph)} = I(h).$$

The following theorem claims that the invariant I is a complete description of h modulo convolution with a Gaussian.

Theorem 2. *Let $h_1, h_2 \in \mathcal{D}$ and let I be the invariant defined in Theorem 1. Then $I(h_1) = I(h_2)$ if and only if there exist $h_{\mathbf{N}_1}, h_{\mathbf{N}_2} \in \mathcal{S}$ such that $h_{\mathbf{N}_1} * h_1 = h_{\mathbf{N}_2} * h_2$.*

Proof.

$$I(h_1) = I(h_2) \Leftrightarrow \frac{\Phi(h_1)}{\Phi(Ph_1)} = \frac{\Phi(h_2)}{\Phi(Ph_2)} \Leftrightarrow \Phi(h_1)\Phi(Ph_2) = \Phi(h_2)\Phi(Ph_1)$$
$$\Leftrightarrow \Phi(h_1 * Ph_2) = \Phi(h_2 * Ph_1) \Leftrightarrow h_1 * Ph_2 = h_2 * Ph_1,$$

which implies that $h_{\mathbf{N}_1}, h_{\mathbf{N}_2} \in \mathcal{S}$ exist and are defined as $h_{\mathbf{N}_1} = Ph_2$, $h_{\mathbf{N}_2} = Ph_1$.

Theorems 1 and 2 show that invariant I entirely and uniquely describes any normalized histogram modulo convolution with a Gaussian.

Consider the equivalence relation on $\mathcal{D}$: $h_1 \sim h_2$ if and only if there exist functions $h_{\mathbf{N}_1}, h_{\mathbf{N}_2} \in \mathcal{S}$ such that $h_{\mathbf{N}_1} * h_1 = h_{\mathbf{N}_2} * h_2$. The factor set $\mathcal{D}/_\sim$ is the same as the orbit set of the semi-group action with $(\mathcal{S}, *)$. Invariant I is constant within each orbit but has distinct values for any two different orbits. In particular, for the Gaussian orbit $\mathcal{S}$ we have $I(h) = 1$.

4 Invariants from Moments

Although theoretically the invariant $I(h)$ fully describes the orbit, several problems may occur when dealing with finite-precision arithmetic. The division by small numbers leads to the precision loss. To speed up the computation, it would be better to avoid the explicit construction of $\Phi(h)$ and $\Phi(Ph)$. In this Section, we show that it can be accomplished by constructing moment-based invariants equivalent to $I(h)$. The idea of describing a histogram by its moments is reasonable. Moment-based representation yields an additional feature – the number of the moments used is a user-defined parameter by means of which we may control the trade-off between a high compression on one hand and an accurate histogram representation on the other hand [7].

We rewrite Eq. (8) as

$$\Phi(Ph)(\mathbf{u}) \cdot I(h)(\mathbf{u}) = \Phi(h)(\mathbf{u}). \tag{9}$$

If the characteristic function is n-times differentiable, then the nth derivative is a moment of the PDF up to a multiplicative constant. Assuming that both $\Phi(h)$ and $\Phi(Ph)$ have a Taylor expansion, then we can write, using a multi-index notation,

$$\sum_{\substack{\mathbf{k}=\mathbf{0} \\ |\mathbf{k}|\neq 0,\ \text{even}}}^{\infty} \frac{\mathrm{i}^{|\mathbf{k}|}}{\mathbf{k}!} m_{\mathbf{k}}^{(Ph)} \mathbf{u}^{\mathbf{k}} \cdot \sum_{\mathbf{p}=\mathbf{0}}^{\infty} \frac{\mathrm{i}^{|\mathbf{p}|}}{\mathbf{p}!} A_{\mathbf{p}} \mathbf{u}^{\mathbf{p}} = \sum_{\mathbf{q}=\mathbf{0}}^{\infty} \frac{\mathrm{i}^{|\mathbf{q}|}}{\mathbf{q}!} m_{\mathbf{q}}^{(h)} \mathbf{u}^{\mathbf{q}}, \tag{10}$$

where

$$m_{\mathbf{k}}^{(h)} = \int \mathbf{x}^{\mathbf{k}} h(\mathbf{x})\, \mathrm{d}\mathbf{x}. \tag{11}$$

By comparing coefficients of the same power of $\mathbf{u}$ we get

$$\sum_{\substack{\mathbf{l}=\mathbf{0} \\ |\mathbf{l}|\ \text{even}}}^{\mathbf{k}} \frac{\mathrm{i}^{|\mathbf{l}|}}{\mathbf{l}!} m_{\mathbf{l}}^{(Ph)} \frac{\mathrm{i}^{|\mathbf{k}|-|\mathbf{l}|}}{(\mathbf{k}-\mathbf{l})!} A_{\mathbf{k}-\mathbf{l}} = \frac{\mathrm{i}^{|\mathbf{k}|}}{\mathbf{k}!} m_{\mathbf{k}}^{(h)}, \tag{12}$$

which is equivalent to

$$\sum_{\substack{\mathbf{l}=\mathbf{0} \\ |\mathbf{l}| \text{ even}}}^{\mathbf{k}} \binom{\mathbf{k}}{\mathbf{l}} m_{\mathbf{l}}^{(Ph)} A_{\mathbf{k}-\mathbf{l}} = m_{\mathbf{k}}^{(h)}. \tag{13}$$

Each $A_{\mathbf{k}}$, being a Taylor coefficient of $I(h)$, must be an invariant. Rearranging the previous equation, we obtain a recursive formula for $A_{\mathbf{k}}$

$$A_{\mathbf{k}} = m_{\mathbf{k}}^{(h)} - \sum_{\substack{\mathbf{l}=\mathbf{0} \\ |\mathbf{l}|\neq 0,\ \text{even}}}^{\mathbf{k}} \binom{\mathbf{k}}{\mathbf{l}} m_{\mathbf{l}}^{(Ph)} A_{\mathbf{k}-\mathbf{l}}. \tag{14}$$

The intuitive meaning of invariants $A_{\mathbf{k}}$ is the following one. If the "mother PDF" h_r exists, then $I(h)$ is its characteristic function and $A_{\mathbf{k}}$'s are its moments. The invariants $A_{\mathbf{k}}$ can be, however, applied even if h_r does not exist. Another noteworthy point is that generally we have to calculate moments of both h and Ph in order to evaluate Eq. (14). In the next Section, we show how the construction of Ph and calculation of its moments can be avoided.

5 Invariants in One and Two Dimensions

Histogram is a function of either a single variable when working with graylevel images or of multiple variables for color/multispectral images. In this Section, we show how Eq. (14) can be further simplified in 1D and 2D cases.

In 1D, the invariants (14) obtain the form

$$A_p = m_p^{(h)} - \sum_{\substack{k=2 \\ k \text{ even}}}^{p} \binom{p}{k} (k-1)!!\, m_2^{k/2} A_{p-k}\,. \tag{15}$$

This simplification follows from the fact that the odd-order moments of a 1D Gaussian with standard deviation σ vanish and the even-order ones are given as $m_p = \sigma^p (p-1)!!$. Furthermore, $\sigma^2 \equiv m_2^{(Ph)} = m_2^{(h)}$ which allows to express all moments of Ph in terms of moments of h.

In 2D, simplification of Eq. (14) is much more difficult. First, we need to express the moments of 2D Gaussian explicitly. If we assume that the two components of our random variable $\mathbf{N}$ are independent, then we can constraint the covariance matrix of Ph to be diagonal. Then the moments of Ph are simply

$$m_{pq}^{(Ph)} = m_{20}^{p} m_{02}^{q} (p-1)!!(q-1)!! \tag{16}$$

and we obtain similar formula as in 1D case

$$A_{mn} = m_{mn}^{(h)} - \sum_{\substack{l=0 \\ l+k\neq 0, \\ l+k \text{ even}}}^{m} \sum_{k=0}^{n} \binom{m}{l} \binom{n}{k} (l-1)!!(k-1)!!\, m_{20}^{l/2} m_{02}^{k/2} A_{m-l,n-k}\,. \tag{17}$$

However, the assumption of independent noise components and hence of the diagonal covariance matrix is not always realistic in practice. In Fig. 4 we can see 2D histogram of red and green channels of the noise, extracted from a real RGB photograph. The noise is actually a background and thermal noise of the camera; the noise extraction was done by subtracting a time-averaged image. A correlation about 0.33 was found (and is also apparent visually) between the noise acting in the red and the green channel (the correlation was probably introduced by an in-built postprocessing/interpolation on the chip). So, to make our method applicable in practice, we have to assume a general covariance matrix of Ph.

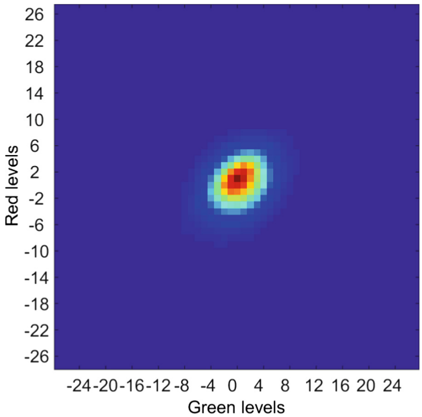

Fig. 4. 2D histogram of the noise extracted from red and green channels of a real digital image. The on-chip postprocessing introduced a correlation about 0.33 between the noise in both channels. (Color figure online)

For the general covariance matrix, the formula for moments of a Gaussian is not commonly cited in the literature.[1] Therefore we derived an explicit formula for 2D zero-mean case, which is very useful in the sequel:

$$\mathrm{m}_{mn}^{(Ph)} = \sum_{i=0}^{\lfloor \frac{m}{2} \rfloor} \sum_{\substack{j=0 \\ j \geq \frac{m-n}{2}}}^{i} (-1)^{i-j} \binom{m}{2i} \binom{i}{j} (m+n-2i-1)!!(2i-1)!! \cdot \\ \cdot m_{11}^{m-2j} m_{20}^{j} m_{02}^{\frac{n-m}{2}+j} . \tag{18}$$

[1] The reader is usually referred to the classical Isserlis' paper [4] or to some more recent papers [10,11,13] but no simple explicit formula can be found there.

If we use Formula (18), the recurrence (14) turns to the form

$$A_{mn} = m_{mn}^{(h)} - \sum_{\substack{l=0 \\ l+k\neq 0, \\ l+k \text{ even}}}^{m} \sum_{k=0}^{n} \binom{m}{l}\binom{n}{k} \sum_{i=0}^{\lfloor \frac{k}{2} \rfloor} \sum_{\substack{j=0 \\ j \geq \frac{k-l}{2}}}^{i} (-1)^{i-j} \binom{k}{2i}\binom{i}{j} \cdot \\ \cdot (l+k-2i-1)!!(2i-1)!!\, m_{11}^{k-2j} m_{20}^{\frac{l-k}{2}+j} m_{02}^{j} A_{m-l,n-k} \,. \tag{19}$$

Note that the formula contains only the moments of h. Neither the characteristic functions $\Phi(h)$ and $\Phi(Ph)$ nor the projection Ph itself are necessary to be constructed.

6 Experiment

In this experiment, we show the invariance property if the noise follows the Gaussian model. We used blue and green channels of a real RGB image as "clear" test data. We corrupted the image 100 times with a Gaussian noise generated according to Eq. (2), the covariance matrix of which was randomly generated in each realization. As one can see in Fig. 5, the histogram of noisy image is actually a smoothed version of the histogram of the clear image.

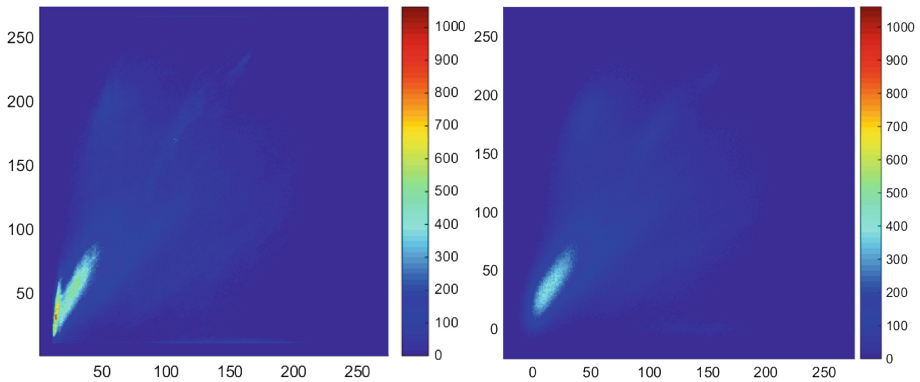

Fig. 5. 2D histogram of the blue and green channels of the original clear image (left) and of the same image corrupted by additive Gaussian noise (right). Note that the "noisy" histogram is actually a smoothed version of the original histogram. (Color figure online)

We calculated more than 300 invariants (19) of each noisy image histogram. In Fig. 6 left, we can see a cumulative graph of the ratio between the invariants of noisy and clear histograms (each invariant order is represented by a single cumulative value). The ratio is almost perfectly equal to one, even for higher orders. To show the advantage of the proposed invariants over the plain moments, we calculated the same for histogram moments, see Fig. 6 right. The errors of the moments are much higher since the moments do not posses the invariance property and are heavily influenced by noise. The error grows with the increasing moment order. This experiment clearly shows the quality of the proposed histogram descriptors.

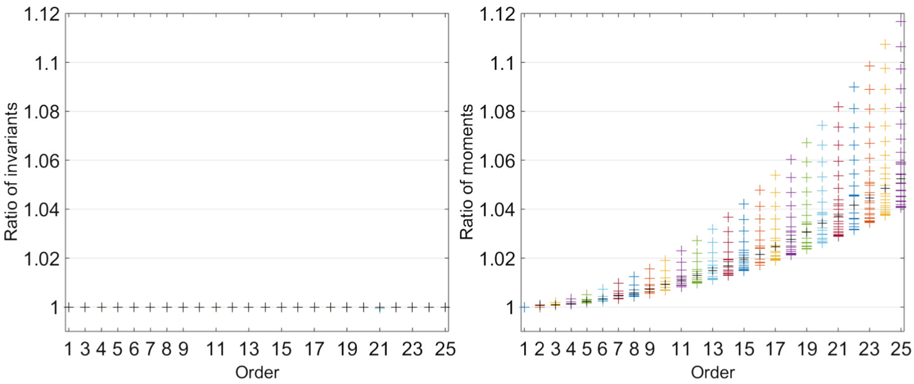

Fig. 6. Ratio of the invariants (left) and of the moments (right) between the 2D histograms of noisy and clear images. The invariants/moments of the same order have been cumulated. Black crosses denote the median of the invariants.

7 Conclusion

We proposed a new method for description of multidimensional histogram, which is robust to additive Gaussian noise in the source image. The method employs the fact that the histogram of noisy image is a convolution of the original histogram and the histogram of the noise. We proposed moment-based descriptors, which characterize the original histogram but can be computed directly from the histogram of the noisy image. The method does not require any denoising or estimation of the noise parameters, which makes it attractive for practical usage. Potential applications are in noise-robust histogram-based image retrieval and also in some areas outside image processing.

References

1. Flusser, J., Suk, T.: Degraded image analysis: an invariant approach. IEEE Trans. Pattern Anal. Mach. Intell. **20**(6), 590–603 (1998)
2. Flusser, J., Suk, T., Boldyš, J., Zitová, B.: Projection operators and moment invariants to image blurring. IEEE Trans. Pattern Anal. Mach. Intell. **37**(4), 786–802 (2015)
3. Höschl IV, C., Flusser, J.: Robust histogram-based image retrieval. Pattern Recogn. Lett. **69**(1), 72–81 (2016)

4. Isserlis, L.: On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. Biometrika **12**(1/2), 134–139 (1918)
5. Lukacs, E.: Characteristic Functions Griffin books of Cognate Interest. Hafner Publishing Company, New York (1970)
6. Makaremi, I., Ahmadi, M.: Wavelet domain blur invariants for image analysis. IEEE Trans. Image Process. **21**(3), 996–1006 (2012)
7. Mandal, M.K., Aboulnasr, T., Panchanathan, S.: Image indexing using moments and wavelets. IEEE Trans. Consum. Electron. **42**(3), 557–565 (1996)
8. Pass, G., Zabih, R.: Histogram refinement for content-based image retrieval. In: Proceedings 3rd IEEE Workshop on Applications of Computer Vision WACV 1996, pp. 96–102. IEEE (1996)
9. Pedone, M., Flusser, J., Heikkilä, J.: Blur invariant translational image registration for N-fold symmetric blurs. IEEE Trans. Image Process. **22**(9), 3676–3689 (2013)
10. Schott, J.R.: Kronecker product permutation matrices and their application to moment matrices of the normal distribution. J. Multivar. Anal. **87**(1), 177–190 (2003)
11. Song, I., Lee, S.: Explicit formulae for product moments of multivariate gaussian random variables. Stat. Probab. Lett. **100**, 27–34 (2015)
12. Swain, M.J., Ballard, D.H.: Color indexing. Int. J. Comput. Vis. **7**(1), 11–32 (1991)
13. Triantafyllopoulos, K.: On the central moments of the multidimensional gaussian distribution. Math. Sci. **28**(2), 125–128 (2003)
14. Wang, L., Healey, G.: Using Zernike moments for the illumination and geometry invariant classification of multispectral texture. IEEE Trans. Image Process. **7**(2), 196–203 (1998)
15. Zhang, H., Shu, H., Han, G.N., Coatrieux, G., Luo, L., Coatrieux, J.L.: Blurred image recognition by Legendre moment invariants. IEEE Trans. Image Process. **19**(3), 596–611 (2010)