Robust multivariate density estimation under Gaussian noise

Jitka Kostková & Jan Flusser

Multidimensional Systems and Signal Processing
An International Journal

ISSN 0923-6082
Volume 31
Number 3

Multidim Syst Sign Process (2020)
31:1113-1143
DOI 10.1007/s11045-020-00702-7

Springer

Multidimensional Systems and Signal Processing (2020) 31:1113–1143
https://doi.org/10.1007/s11045-020-00702-7

Robust multivariate density estimation under Gaussian noise

Jitka Kostková[1] · **Jan Flusser**[1,2]

Received: 17 June 2019 / Revised: 13 January 2020 / Accepted: 20 January 2020 /
Published online: 30 January 2020

Abstract
Observation of random variables is often corrupted by additive Gaussian noise. Noise-reducing data processing is time-consuming and may introduce unwanted artifacts. In this paper, a novel approach to description of random variables insensitive with respect to Gaussian noise is presented. The proposed quantities represent the probability density function of the variable to be observed, while noise estimation, deconvolution or denoising are avoided. Projection operators are constructed, that divide the probability density function into a non-Gaussian and a Gaussian part. The Gaussian part is subsequently removed by modifying the characteristic function to ensure the invariance. The descriptors are based on the moments of the probability density function of the noisy random variable. The invariance property and the performance of the proposed method are demonstrated on real image data.

Keywords Multivariate density · Gaussian additive noise · Noise-robust estimation · Moments · Invariant characteristics

1 Introduction

Observation of random variables in a real-world environment is often corrupted by numerous degradation factors, among which an additive random noise is one of the most frequent ones. The noise may be introduced by measurement device imperfection, by storing and transmitting, and also due to the precision loss when pre-processing the data.

This work has been supported by the Czech Science Foundation (Grant No. GA18-07247S), by the Grant SGS18/188/OHK4/3T/14 provided by the Ministry of Education, Youth, and Sports of the Czech Republic (MŠMT ČR), by *Praemium Academiae*, and by the Joint Laboratory SALOME2. Thanks to Dr. Cyril Höschl IV for kind providing the test images used in the experiments in Section 8.2.

✉ Jitka Kostková
kostkova@utia.cas.cz

Jan Flusser
flusser@utia.cas.cz

1 Institute of Information Theory and Automation, Czech Academy of Sciences, Pod vodárenskou věží 4, 182 08 Prague 8, Czech Republic

2 Faculty of Management, University of Economics, Jarosovska 1117/II, 377 01 Jindrichuv Hradec, Czech Republic

Let X be the multivariate random variable to be observed and let N be an additive noise. As a result of the measurement, we actually observe realizations of a random variable $Z = X + N$ instead of X, which is observed only indirectly. If the signal-to-noise ratio is low, the corruption is so heavy that it is very difficult to deduce anything about the observed variable X from the sample data Z. This situation occurs frequently in many application areas such as signal and image processing, econometrics, experimental physics, geoscience, and many others.

A large amount of effort has been spent to develop methods that decrease the impact of the noise and allow to estimate either the entire X or at least some of its discriminative characteristics. These methods can be categorized into three main groups – denoising, deconvolution, and robust estimators.

Denoising methods aim to suppress the noise in the data and are usually based on linear or non-linear filtering of high-frequency components and data smoothing. The advantage of denoising methods is that they provide a complete estimate of X while the common disadvantage however may be artifacts and deformation or loss of high-frequency information contained in original X.

Deconvolution methods try to recover the *probability density function* (PDF) f_X of random variable X from the estimated PDF of the observed Z. Assuming the noise is independent of the data, it is well known that the PDF of the sum of two independent random variables is a convolution of the densities of the summands, i.e. in our case

$$f_Z(x) = (f_X * f_N)(x) = \int f_X(x-s) f_N(s)\, \mathrm{d}s. \tag{1}$$

Deconvolution methods could theoretically yield an accurate estimate of f_X but in reality they suffer from several drawbacks. Blind deconvolution methods, which do not require any prior knowledge of the noise density f_N, are numerically unstable, may converge to an incorrect solution and are very time-consuming. Non-blind methods are better in that sense but obtaining a good estimate of f_N may be in practice difficult or even impossible.

Robust estimators try to estimate certain characteristics of X such as mean value, variance, skewness and higher-order moments directly from the observed samples of Z. Standard formulas for sample moments do not perform well on noisy data. This is why some authors proposed not to resolve Eq. (1) but only to find some characteristics which are not affected by convolution. Such characteristics, called *convolution invariants*, must be the same for both f_Z and f_X and should be calculated directly from sample observations of Z. This modern approach [it was firstly proposed in Höschl IV and Flusser (2016)] may be very efficient whenever a complete knowledge of f_X is not necessary, typically in noisy signal classification and signal/image retrieval. In this paper, we develop this idea substantially.

The main idea of this paper is as follows. We assume that noise N is a multivariate Gaussian random variable with zero mean and a general covariance matrix, which is unknown. We introduce *projection operators*, acting on the space of all PDF's, that divide each PDF into two components. Based on the known parametric form of f_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 f_X without any noise estimation, denoising and deconvolution. Unlike Höschl IV and Flusser (2016), where the idea of invariant descriptors was used heuristically for univariate densities only, we present here a consistent theory for multivariate densities.

After providing the state-of-the-art review in Sect. 2, we formulate the problem formally in Sect. 3. Section 4 is dedicated to the construction of a projection operator and its relation to invariants. This theory is then used for the construction of moment-based invariants descriptors in r dimensions (r-D) in Sect. 5. In Sect. 6, one and two-dimensional cases are analyzed in detail and explicit invariant formula is derived. Finally, the experimental section verifies the theory on simulated as well as real data from image processing area and shows a potential application in image retrieval.

2 Literature review

2.1 Denoising methods

Majority of the articles on denoising comes from signal processing area, such as Motwani et al. (2004), Buades et al. (2005b). These methods often lead to subjective improvement but they can cause the loss of important information, the formation of artifacts, smoothing of the signal, etc. The simplest way of removing noise from the signal is a linear filtering, when the corrupted signal is convolved with some low-pass filter. This method, however, leads to deformation of the high-frequency components of the signal. Non-linear filtering methods, such as median filter and the anisotropic filter designed by Perona and Malik (1990), attempt to avoid the effect of signal blurring. One of the state-of-the-art methods is the *Non-Local Means* algorithm based on the self-similarity of signal patches (Buades et al. 2005a). Various methods rely on a transform domain filtering, e.g. wavelet-based denoising (Chen et al. 2013; Cho and Bui 2005), ridgelet- and curvelet-based denoising (Starck et al. 2002), and Fourier Wiener filtering (Khireddine et al. 2007). Other methods make use of minimization of some functionals, e.g. a total variation method (Chambolle and Lions 1997) and a method using higher order statistics (Teuber et al. 2012).

2.2 Deconvolution methods

Many books and papers dedicated to this topic have been published. The tutorial article (De Brabanter and De Moor 2012) and the book (Meister 2009) gave an introduction to deconvolution problems in non-parametric statistics (density estimation based on contaminated data, errors-in-variables regression, and image reconstruction). One of the approaches is to estimate the density of Z in the non-parametric form by a kernel estimator and then to use Fourier transform to recover the distribution of X (Carroll and Hall 1988; Stefanski and Carroll 1990; Fan 1992); another is Bayesian approach (Efron 2014) and wavelet approach (Pensky et al. 1999). The paper (Butucea et al. 2009) tries to estimate $\mathbb{E}[\psi(X)]$, where ψ is a known integrable function and the distribution of N is known. In Johannes et al. (2009), the authors deal with the estimation of deconvolution, when only an estimate of the distribution of N is available. The authors of Comte and Lacour (2011) suppose that the distribution of N is unknown and present an adaptive estimator. The goal of Kappus et al. (2014) is analogous, but they do not impose any assumption on the shape of the characteristic function of noise.

2.3 Convolution invariants

The use of convolution invariants for a noise-robust PDF estimation was firstly proposed in the paper Höschl IV and Flusser (2016), which was motivated by histogram-based image retrieval. The authors presented invariants defined for univariate densities only. Their invariants were based on moments of a histogram of the noisy graylevel image. However, the results of Höschl IV and Flusser (2016) cannot be easily extended to multidimensional signals and multivariate PDF's. It should be noted that convolution invariants have been thoroughly studied in a different context and domains (Flusser and Suk 1998; Flusser et al. 2003; Galigekere and Swamy 2006; Ojansivu and Heikkilä 2007; Zhang et al. 2010; Gopalan et al. 2012; Makaremi and Ahmadi 2012; Pedone et al. 2013; Flusser et al. 2015) but those results are not suitable for noisy PDF estimation due to a very specific convolution kernel shapes, that do not correspond to real-life noise PDFs.

3 Problem formulation

Let X and N be two r-D independent random variables with probability density functions f_X and f_N, respectively, where $N \sim \mathcal{N}(\mathbf{0}, \Sigma)$ is a normally distributed zero-mean random variable with a regular covariance matrix Σ. Then f_N has the well-known Gaussian shape

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

where $\mathbf{x} = (x_1, \ldots, x_r)^T$. Most frequently, but not necessarily, X is the multivariate random variable which represents the original non-corrupted data and N is an additive Gaussian noise.

Under the above assumptions, the PDF of the sum of these variables

$$Z = X + N \tag{3}$$

exists and is given by

$$f_Z = f_X * f_N. \tag{4}$$

Our aim is to design a functional (descriptor) I, which is invariant with respect to the noise. Since we construct this functional on the space of the PDFs, we require

$$I(f_X) = I(f_Z) = I(f_X * f_N) \tag{5}$$

for any normally distributed zero-mean random variable N with arbitrary (unknown) covariance matrix Σ.

To comply with Eq. (5) is, however, not the only desirable property of I. At the same time, I must be *discriminable*, which means

$$I(f_X) \neq I(f_Y) \tag{6}$$

for any X and Y such that they are not linked as $Y = X + N$ for any Gaussian N.

If we design such invariant I, it maps the original data as well as all its corrupted versions into a single point in an abstract feature space, while any two distinct data are mapped into distinct points regardless of their potential corruption. Such invariant feature space can be efficiently used for data description and classification.

4 Construction of the invariant

The main idea of constructing invariants to Gaussian convolution is based on projections of a PDF onto the set of all Gaussian functions. In this way, we extract the Gaussian component of the random variable. We will show that the ratio of the characteristic functions of the original random variable and of its Gaussian component possesses the desired invariant property. In the sequel, we introduce the necessary mathematical background.

Let us denote the set of all probability density functions with finite second-order central moments as $\mathcal{D}$ and the set of all zero-mean Gaussian probability density functions

$$\mathcal{S} = \{f_N | \Sigma \rhd 0\}, \tag{7}$$

where $\Sigma \rhd 0$ denotes the positive-definiteness of covariance matrix Σ. The set $\mathcal{S}$ exhibits the following basic properties.

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

It holds for any two Gaussian PDFs f_{N_1} and f_{N_2} with covariance matrices Σ_1 and Σ_2 that the result of convolution

$$f_{N_1} * f_{N_2} = f_N$$

is again a Gaussian PDF with covariance matrix $\Sigma = \Sigma_1 + \Sigma_2$.

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

$$Pf = f_N, \tag{8}$$

where f_N has the same covariance matrix as f. The operator P is well defined for all PDFs with a regular covariance matrix[1] and is idempotent, i.e. $P^2 = P$. Note that P is not linear, so it is not a projector in the common sense known from linear algebra.

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.

Lemma 2 *Let $f \in \mathcal{D}$ and $g \in \mathcal{S}$. Then*

$$P(f * g) = Pf * g. \tag{9}$$

Proof First, we investigate the right-hand side, where we have a convolution of two Gaussians with covariance matrices Σ_f and Σ_g, respectively. Thanks to Lemma 1, this is also a Gaussian with covariance matrix $\Sigma = \Sigma_f + \Sigma_g$.

On the left-hand side, $P(f * g)$ is by definition a Gaussian with covariance matrix Σ_{f*g}. 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, we have $\Sigma_{f*g} = \Sigma_f + \Sigma_g = \Sigma$, 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.

Theorem 1 *Let $f \in \mathcal{D}$ and let P be the projection operator defined above. Then the ratio of characteristic functions Φ of the densities f and Pf*

$$I(f) = \frac{\Phi(f)}{\Phi(Pf)} \tag{10}$$

is an invariant to convolution with a Gaussian probability density function, i.e. $I(f) = I(f * f_N)$ for any $f_N \in \mathcal{S}$.

[1] It is possible to extend this definition also to singular covariance matrices by admitting degenerated Gaussian densities in $\mathcal{S}$.

Proof First, note that I is well defined. Both characteristic functions always exist (they actually equal to the Fourier transform of the respective PDF) and the denominator is non-zero everywhere. The rest of the proof follows from the fact that P commutes with the Gaussian convolution (see Lemma 2). If $f_N \in \mathcal{S}$, then

$$I(f * f_N) = \frac{\Phi(f * f_N)}{\Phi(P(f * f_N))} = \frac{\Phi(f * f_N)}{\Phi(Pf * f_N)} = \frac{\Phi(f)\Phi(f_N)}{\Phi(Pf)\Phi(f_N)} = \frac{\Phi(f)}{\Phi(Pf)} = I(f).$$

□

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

Theorem 2 *Let f_1 and f_2 be two probability density functions and $I(f)$ be the invariant defined in Theorem 1. Then $I(f_1) = I(f_2)$ if and only if there exist $f_{N_1}, f_{N_2} \in \mathcal{S}$ such that $f_{N_1} * f_1 = f_{N_2} * f_2$.*

Proof Let us prove the forward implication first.

$$\begin{aligned} I(f_1) = I(f_2) \Rightarrow \frac{\Phi(f_1)}{\Phi(Pf_1)} = \frac{\Phi(f_2)}{\Phi(Pf_2)} &\Rightarrow \Phi(f_1)\Phi(Pf_2) = \Phi(f_2)\Phi(Pf_1) \\ &\Rightarrow \Phi(f_1 * Pf_2) = \Phi(f_2 * Pf_1) \Rightarrow f_1 * Pf_2 = f_2 * Pf_1. \end{aligned}$$

So, we have found $f_{N_1} = Pf_2$ and $f_{N_2} = Pf_1$. The backward implication follows directly from Theorem 1. □

In 1D, Theorem 2 can be formulated in a stronger way. $I(f_1) = I(f_2)$ if and only if there exists $f_N \in \mathcal{S}$ such that $f_N * f_1 = f_2$ or $f_N * f_2 = f_1$. This statement follows from the divisibility of 1D Gaussian functions but it cannot be extended into the multidimensional case.

Theorems 1 and 2 show that invariant I entirely and uniquely describes any PDF modulo convolution with a Gaussian. In particular, for any $f \in \mathcal{S}$ we have $I(f) = 1$.

5 Invariants from moments

Although theoretically the invariant $I(f)$ fully describes f, several problems can 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 $\Phi(f)$ and $\Phi(Pf)$. In this Section, we show that it can be accomplished by constructing moment-based invariants equivalent to $I(f)$.

We can rewrite Eq. (10) as

$$\Phi(Pf) \cdot I(f) = \Phi(f). \tag{11}$$

If the 1D characteristic function is n-times differentiable, then its k-th derivative ($k \le n$) is the moment m_k of the PDF up to a multiplicative constant. The same is true in multidimensional case. If the characteristic function $\Phi(f)$ has a Taylor expansion, then we can write, using a multi-index notation,

$$\sum_{\substack{\mathbf{k}=\mathbf{0} \\ |\mathbf{k}|\neq 0,\ \text{even}}}^{\infty} \frac{\mathbf{i}^{|\mathbf{k}|}}{\mathbf{k}!} m_{\mathbf{k}}^{(Pf)} \mathbf{u}^{\mathbf{k}} \cdot \sum_{\mathbf{p}=\mathbf{0}}^{\infty} \frac{\mathbf{i}^{|\mathbf{p}|}}{\mathbf{p}!} A_{\mathbf{p}} \mathbf{u}^{\mathbf{p}} = \sum_{\mathbf{q}=\mathbf{0}}^{\infty} \frac{\mathbf{i}^{|\mathbf{q}|}}{\mathbf{q}!} m_{\mathbf{q}}^{(f)} \mathbf{u}^{\mathbf{q}}. \tag{12}$$

where the $A_{\mathbf{k}}$'s are Taylor coefficients of $I(f)$. By equating the coefficients of the same power of $\mathbf{u}$ we get

$$\sum_{\substack{\mathbf{l}=\mathbf{0}\\|\mathbf{l}|\ \text{even}}}^{\mathbf{k}} \frac{\mathrm{i}^{|\mathbf{l}|}}{\mathbf{l}!} m_{\mathbf{l}}^{(Pf)} \frac{\mathrm{i}^{|\mathbf{k}|-|\mathbf{l}|}}{(\mathbf{k}-\mathbf{l})!} A_{\mathbf{k}-\mathbf{l}} = \frac{\mathrm{i}^{|\mathbf{k}|}}{\mathbf{k}!} m_{\mathbf{k}}^{(f)}, \tag{13}$$

which is equivalent to

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

Since $I(f)$ is an invariant, each $A_{\mathbf{k}}$ must be an invariant, too. Re-arranging the previous equation, we obtain a recursive formula for $A_{\mathbf{k}}$

$$A_{\mathbf{k}} = m_{\mathbf{k}}^{(f)} - \sum_{\substack{\mathbf{l}=\mathbf{0}\\|\mathbf{l}|\neq 0,\ \text{even}}}^{\mathbf{k}} \binom{\mathbf{k}}{\mathbf{l}} m_{\mathbf{l}}^{(Pf)} A_{\mathbf{k}-\mathbf{l}}. \tag{15}$$

For characteristic functions that do not possess a complete Taylor expansion, we may use the Taylor's Theorem. If the characteristic function has continuous derivatives up to the order $n+1$, then one can write the characteristic function as the Taylor expansion up to the n-th order plus the remainder. Consequently, the invariants up to the order n exist and follow Formula (15).

An intuitive meaning of the invariants $A_{\mathbf{k}}$ is the following one. They could be understood as moments of a "mother function" f_m, which is a function that has no Gaussian component and that "generates" f in the sense that there exist $g \in \mathcal{S}$ such that $f = g * f_m$. In general, f_m lies outside $\mathcal{D}$ or may not even exist but the invariants $A_{\mathbf{k}}$ can be, however, still applied correctly.

Another noteworthy point is that generally we have to calculate moments of both f and Pf in order to evaluate Eq. (15). In the next Section, we show how the construction of Pf and calculation of its moments can be avoided in one and two dimensions.

6 Invariants in one and two dimensions

In many practical applications, especially in signal and image processing, the PDFs we want to characterize are one dimensional or two dimensional functions. In this Section, we show how Eq. (15) can be further simplified in those cases.

In 1D, the recursive form of invariants (15) obtains the form

$$A_p = m_p^{(f)} - \sum_{\substack{k=2\\k\ \text{even}}}^{p} \binom{p}{k} (k-1)!!\, m_2^{k/2} A_{p-k}. \tag{16}$$

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, $\sigma^2 \equiv m_2^{(Pf)} = m_2^{(f)}$ which allows us to express all moments of Pf in terms of those of f. Thanks to this, Eq. (16) can be equivalently rewritten in a non-recursive form as

$$A_p = \sum_{\substack{k=0 \\ k \text{ even}}}^{p} (-1)^{k/2} \binom{p}{k} (k-1)!!\, m_2^{k/2} m_{p-k}^{(f)}. \tag{17}$$

In 2D, simplification of Eq. (15) is much more difficult. First, we need to express the moments of 2D Gaussian PDF explicitly. If we assume that the two components of our random variable N are independent, then we can constraint the covariance matrix of Pf to be diagonal. Then the general moment of Pf is simply

$$m_{pq}^{(Pf)} = (p-1)!!(q-1)!! m_{20}^{p} m_{02}^{q} \tag{18}$$

and we obtain similar formulas as in 1D case

$$A_{mn} = m_{mn}^{(f)} - \sum_{\substack{l=0 \\ l+k\neq 0, \\ l+k \text{ even}}}^{m} \sum_{k=0}^{n} \binom{m}{l} \binom{n}{k} (l-1)!!(k-1)!!\, m_{20}^{l/2} m_{02}^{k/2} A_{m-l,n-k} \tag{19}$$

in the recursive form and

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

in the explicit form.

However, the assumption of independent components and hence of the diagonal covariance matrix is not realistic in practice. We illustrate this by real data from signal processing. The signal was captured in two spectral bands, both corrupted by a thermal noise of the sensor. This noise can be approximatively modelled as an additive Gaussian noise. In Fig. 1, we can see the 2D histogram of the noise extracted from a real image by means of denoising algorithm and subtracting from the original. The normalized histogram is a sampled PDF of the noise. It is clearly apparent that there is a strong correlation between the noise in red and green channels. So, to make our method applicable in practice, we have to assume a general covariance matrix of Pf.

For a general covariance matrix, the formula for moments of a 2D Gaussian is much more complicated and is not commonly cited in the literature. It can be either deduced, after some manipulations, from the papers presenting general approaches to moment calculation (Isserlis 1918; Bar and Dittrich 1971; Von Rosen 1988; Blacher 2003; Schott 2003; Triantafyllopoulos 2003; Song and Lee 2015) or obtained directly from the definition by integration as shown in "Appendix A".

The moments of 2D Gaussian PDF are given as

$$\mathrm{m}_{mn}^{(Pf)} = \sum_{i=0}^{\lfloor \frac{m}{2} \rfloor} \sum_{\substack{j=0 \\ j\geq \frac{m-n}{2}}}^{i} (-1)^{i-j} \binom{m}{2i} \binom{i}{j} (m+n-2i-1)!!(2i-1)!! m_{11}^{m-2j} m_{20}^{j} m_{02}^{\frac{n-m}{2}+j}. \tag{21}$$

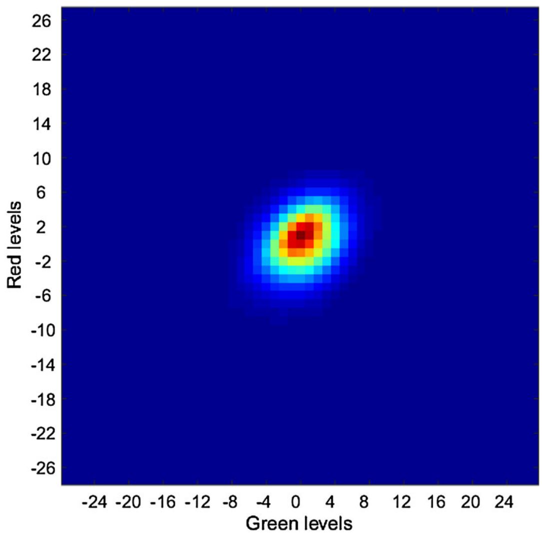


Fig. 1 2D histogram of the noise extracted from red and green channels of a real digital image. The on-chip postprocessing introduces a correlation about 0.33 between the noise in individual channels (Color figure online)

If we use Formula (21), the general recursive definition of the invariants (15) turns to the form

$$A_{mn} = m_{mn}^{(f)} - \sum_{\substack{l=0 \\ l+k\neq 0, \\ l+k \text{ even}}}^{m} \sum_{k=0}^{n} \binom{m}{l}\binom{n}{k} \sum_{i=0}^{\lfloor \frac{k}{2} \rfloor} \sum_{\substack{j=0 \\ j \geq \frac{k-l}{2}}}^{i} (-1)^{i-j} \binom{k}{2i}\binom{i}{j} (l+k-2i-1)!! \cdot (2i-1)!!\, m_{11}^{k-2j} m_{20}^{\frac{l-k}{2}+j} m_{02}^{j} A_{m-l,n-k}, \tag{22}$$

which can again be rewritten into a non-recursive formula

$$A_{mn} = \sum_{\substack{l=0 \\ l+k \text{ even}}}^{m} \sum_{k=0}^{n} (-1)^{\frac{k+l}{2}} \binom{m}{l}\binom{n}{k} \sum_{i=0}^{\lfloor \frac{k}{2} \rfloor} \sum_{\substack{j=0 \\ j \geq \frac{k-l}{2}}}^{i} (-1)^{i-j} \binom{k}{2i}\binom{i}{j} (l+k-2i-1)!! \cdot (2i-1)!!\, m_{11}^{k-2j} m_{20}^{\frac{l-k}{2}+j} m_{02}^{j} m_{m-l,n-k}^{(f)}. \tag{23}$$

The proof of equivalence of (22) and (23) can be found in "Appendix B".

7 Implementation

The formulas (22) and (23) are both efficient in the sense that they contain only the moments of f, which is the PDF of the observed noisy random variable. As we will see in the next section, in practice the theoretical PDF is often replaced with a normalized multidimensional histogram, which is in fact a sampled PDF and is easy to compute directly from the observed values. Neither the characteristic function $\Phi(f)$ nor the projection Pf are necessary to be constructed. This is the main advantage of the moment approach over the direct use of $I(f)$. Hence, in numerical implementation, we always use the moment-based invariants A_{mn} instead of $I(f)$.

The value of A_{mn} can be calculated either from (22) or (23). Since they are theoretically equivalent, the choice depends on the particular task. If all invariants up to a certain order are to be computed, the recursive formula is recommended. Formula (23) is more efficient if we want to compute a single invariant only. The complexity of both depends mainly on our ability to compute the moments of the PDF efficiently. Since the PDF may be of arbitrary shape, we calculate the moments from definition without any speed-up tricks.

In Eqs. (22) and (23), we may use either general or central moments, depending on the nature of the random variable we observe. If there is a systematic shift of the values, which is not important in our application, we use central moments that are not influenced by this shift. Regardless of what moments we employ, some invariants are trivial. $A_{00} = 1$ and $A_{20} = A_{02} = A_{11} = 0$ because of the normalization constraints. If we use central moments, then also $A_{10} = A_{01} = 0$. Trivial invariants are useless for the PDF description and should be removed from the feature vector to reduce its dimensionality.

When calculating the moments of a large-scale histogram, we face a threat of a precision loss due to rounding or even a floating-point overflow. This may happen particularly for higher-order moments and degrade the calculation of the invariants. However, in practice we usually obtain a sufficient characterization of the PDF from the invariants A_{mn} of reasonably low orders (in our experiments in the next section, the maximum order was 25) and we do not encounter any significant precision loss, but one has to be aware of this danger.

8 Experiments

In this section, we demonstrate the invariance property and performance of the proposed method on real data from image processing.

We can view an image as a realization of a random variable, the dimension of which is given by the number of spectral/color bands. Its normalized multidimensional histogram plays the role of a sampled multivariate PDF. The image has been corrupted by an additive Gaussian noise in all bands, which is assumed to be independent of the image content.

8.1 Invariance to simulated noise

In this experiment, we show the invariance property if the noise exactly follows the Gaussian model. We used real color photographs as the test data and we corrupted them by an artificial Gaussian noise generated from the PDF (2). We did not cut off the values below zero and above 255 in order to fulfill the assumption of normal distribution. We used two families of noise, each represented by 100 realizations. Medium noise (SNR about 32 dB) was generated such that the eigenvalues of the correlation matrix were set $\lambda_1 = 6$, $\lambda_2 = 3.5$ and the correlation coefficient was random. Heavy noise (SNR about 28) was generated in the same way using $\lambda_1 = 15$ and $\lambda_2 = 8$. To get 2D histograms, we used the RGB channels pairwise.

In Fig. 2, we can see the first test image. Figure 3 shows a segment of the original and noisy images, respectively, to illustrate the visual appearance of the noise. Figure 4 shows the 2D histogram of blue and green channels and the histogram of the same channels of the noisy image.

We calculated more than 300 invariants using Eq. (22) of a histogram of each noisy image and took the mean value of each invariant (separately for medium noise and heavy noise). Then we calculated relative errors between this mean value and the invariants calculated from the original "clear" histogram. The relative errors of all invariants are visualized in Fig. 5

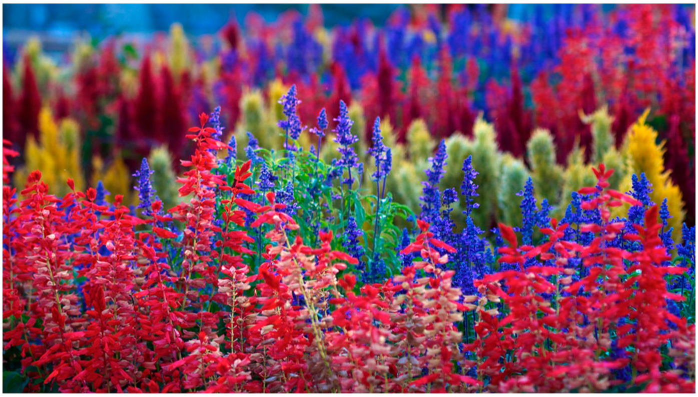


Fig. 2 The original image of a meadow

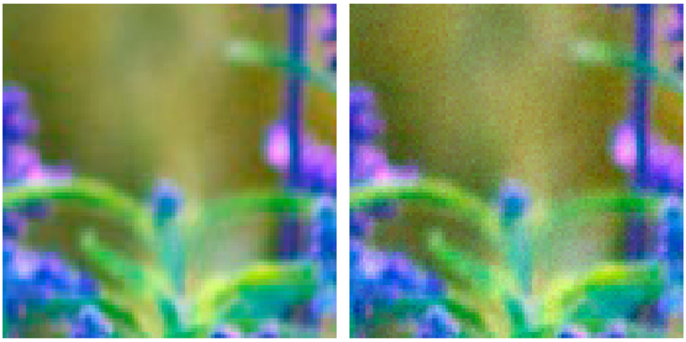


Fig. 3 The segment of the original image (left) and of the noisy image (right). The noise is visually apparent

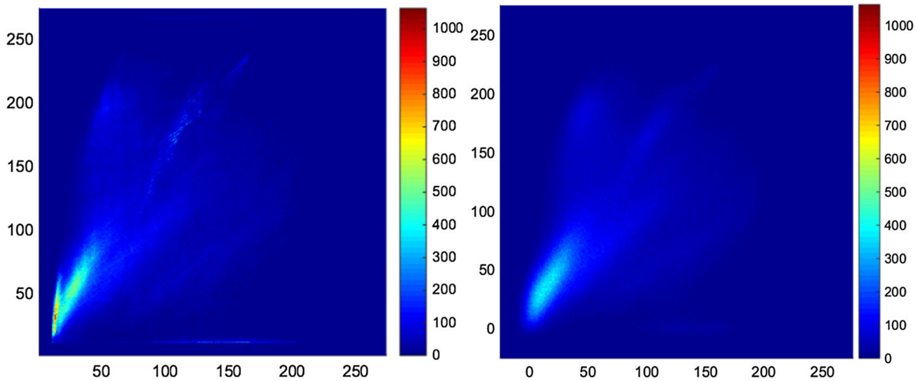


Fig. 4 2D histogram of the blue and green channels of the original clear image (left) and of the same image corrupted by an additive Gaussian noise (right). Note that the "noisy" histogram is actually a smoothed version of the original histogram

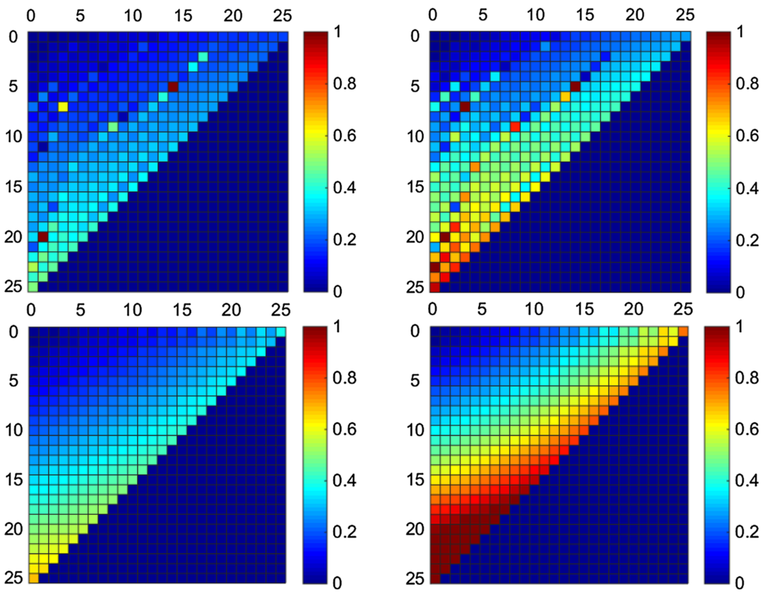


Fig. 5 Mean relative errors of the invariants (top) and of the moments (bottom) up to the order 25. The 25 × 25 matrix contains the color-encoded values of the mean errors of individual invariants/moments. Only the upper left triangle of the matrix is valid. Top left matrix shows invariants calculated from 100 instances of medium noise (see the text for the details on noise generation), top right matrix shows the same for heavy noise. The bottom matrices display the same for the moments (Color figure online)

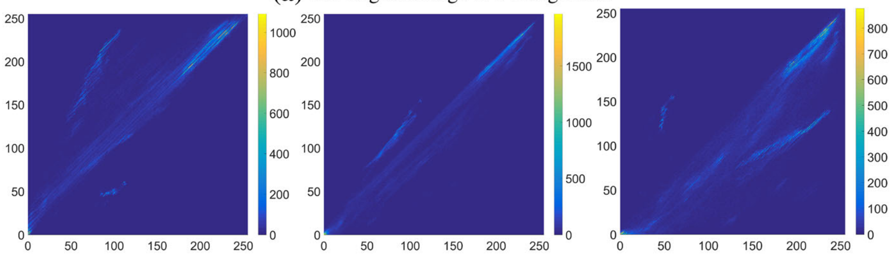
(top). We can see that almost all errors are reasonably low. Relative errors higher than 1% appear only for heavy noise in case of few invariants of orders between 20 and 25. To show the advantage of the proposed invariants over the plain moments, we calculated the same for the moments of the histograms, see Fig. 5 bottom. Comparing the corresponding values, we can see the errors of the moments are by one order higher since the moments do not posses the invariance property and are heavily influenced by noise. The errors of the invariants are, however, not zero as one could expect from the theory. Especially for higher-order invariants, we face precision loss in calculations. Another source of errors is that the theory assumes continuous Gaussian convolution kernel while in the discrete domain we work with sampled and truncated Gaussian.

We repeated this experiment on other test images and with various color band pairs. In most cases, the results were fully comparable to those described above (see Figs. 6, 7 and 8 for three other examples). However, we found a few examples where the relative errors of the invariants are significantly higher, even for medium noise. This occurs when the histogram is extremely sparse. In such a case, the sampling errors are more significant and the invariants properties derived in a continuous domain are violated (see Fig. 9 for an example).

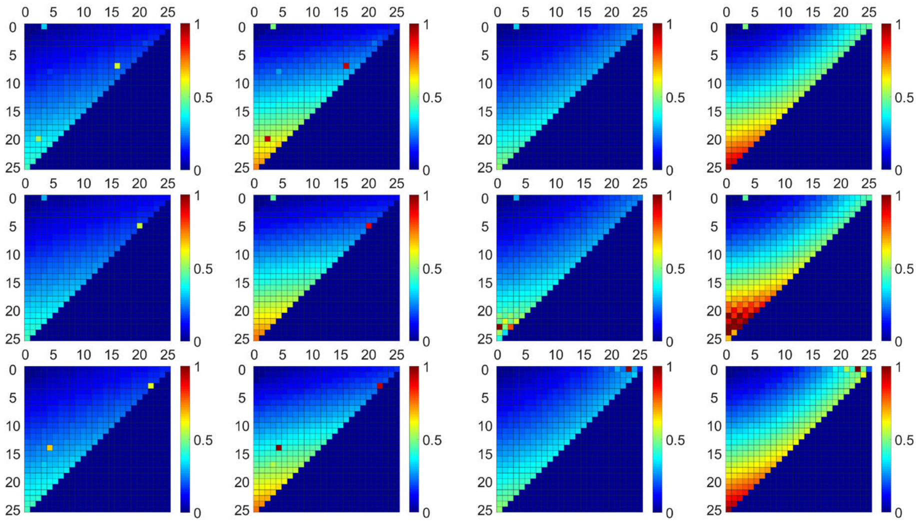
8.2 Invariance on real pictures

In the second experiment, we tested the invariance on real noisy images captured by a compact camera. The noise comes mainly due to physical processes on the CCD chip and has several components. Photon shot noise, thermal noise, readout noise and background noise are the main ones. An additive noise component can be reasonably modelled by a Gaussian random

(a) The original image of a living room.



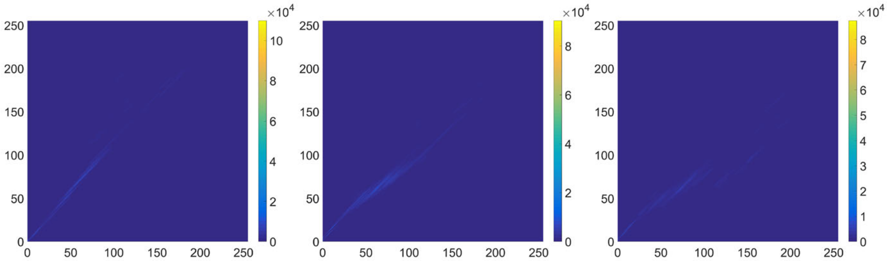
(b) 2D histograms of the original image - red-green (left), blue-green (middle), red-blue (right).



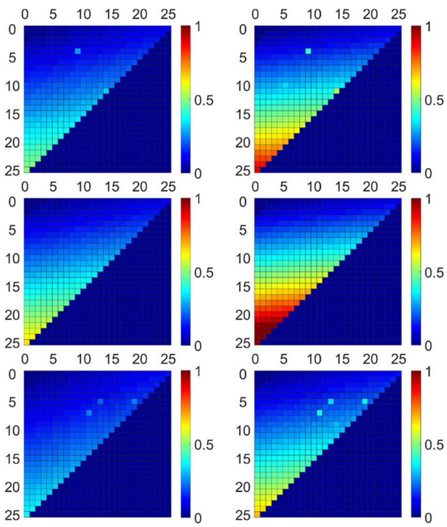
(c) MRE of the invariants. **(d)** MRE of the moments.

Fig. 6 Picture of a living room **a** with its 2D histograms **b** and mean relative errors (**c**, **d**). Mean relative errors of the invariants for medium noise (first column) and heavy noise (second column), and of the moments for the same noise (third and fourth column, respectively). Histograms of red–green (first row), blue–green (second row) and red–blue (third row) channels were used (Color figure online)

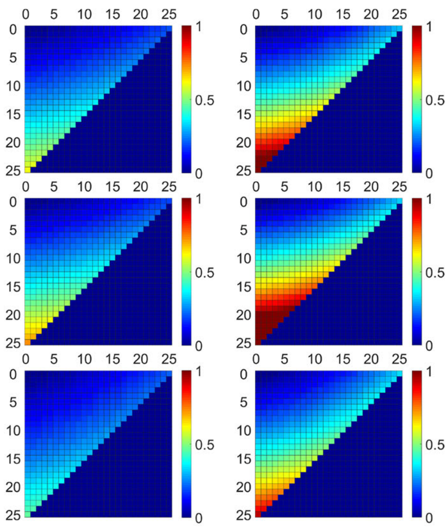
(a) The original image of a couple.



(b) 2D histograms of the original image - red-green (left), blue-green (middle), red-blue (right).

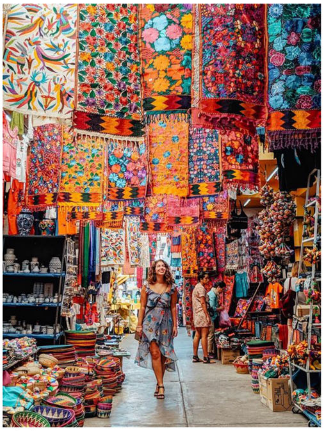


(c) MRE of the invariants.

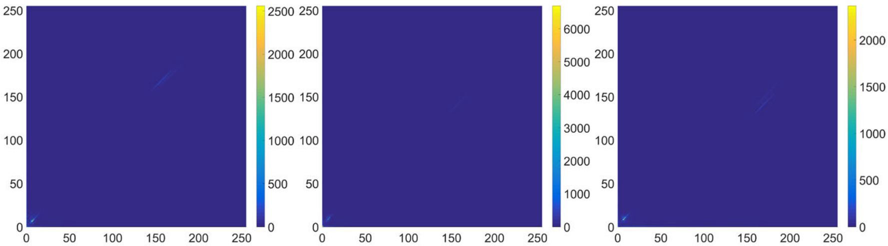


(d) MRE of the moments.

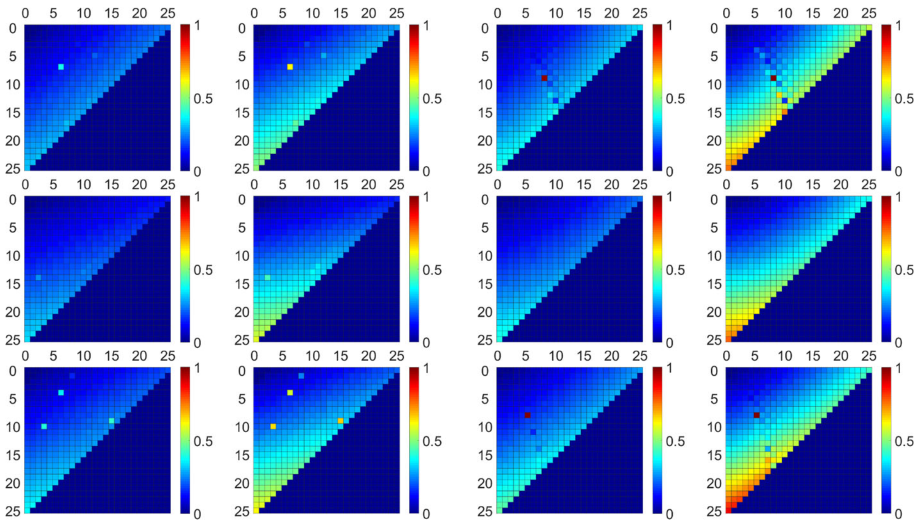
Fig. 7 Picture of a couple **a** with its 2D histograms **b** and mean relative errors (**c**, **d**). Mean relative errors of the invariants for medium noise (first column) and heavy noise (second column), and of the moments for the same noise (third and fourth column, respectively). Histograms of red–green (first row), blue–green (second row) and red–blue (third row) channels were used (Color figure online)



(a) The original image of a market.



(b) 2D histograms of the original image - red-green (left), blue-green (middle), red-blue (right).

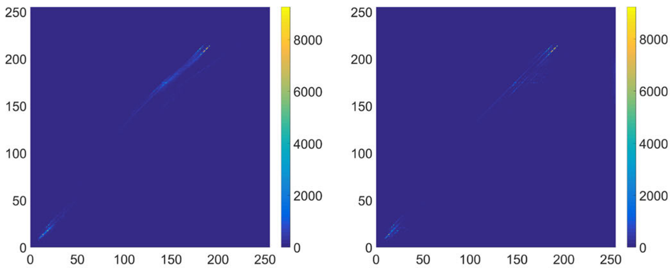


(c) MRE of the invariants.

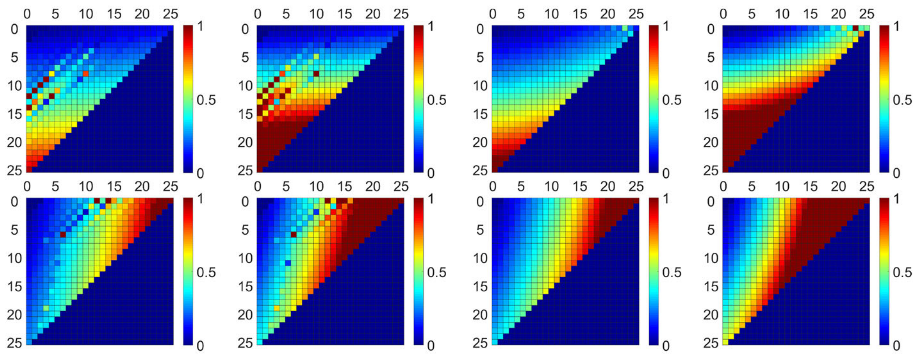
(d) MRE of the moments.

Fig. 8 Picture of a market **a** with its 2D histograms **b** and mean relative errors (**c**, **d**). Mean relative errors of the invariants for medium noise (first column) and heavy noise (second column), and of the moments for the same noise (third and fourth column, respectively). Histograms of red–green (first row), blue–green (second row) and red–blue (third row) channels were used (Color figure online)

(a) The original image of a mountain.



(b) 2D histograms of the original image - blue-green (left), red-blue (right).



(c) MRE of the invariants. **(d)** MRE of the moments.

Fig. 9 Picture of a mountain **a** with its 2D histograms **b** and mean relative errors (**c**, **d**). Mean relative errors of the invariants for medium noise (first column) and heavy noise (second column), and of the moments for the same noise (third and fourth column, respectively). Histograms of blue–green (first row) and red–blue (second row) channels were used (Color figure online)

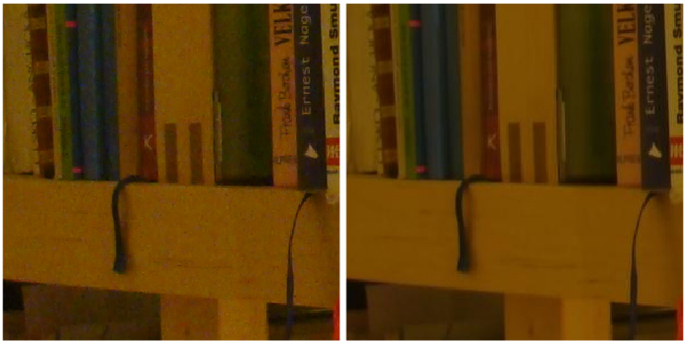


Fig. 10 A patch of a real noisy picture (left) and the same patch with the noise suppressed by time averaging over 10 images (right)

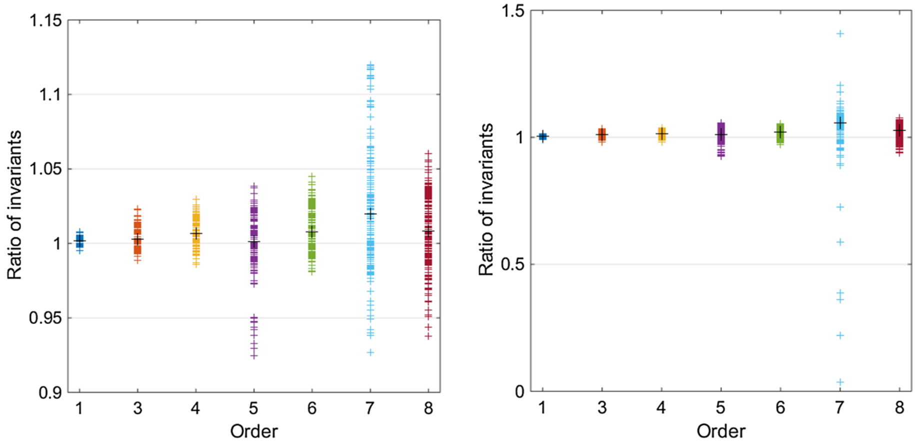
variable uncorrelated with the image values while the signal-dependent component (which is less significant here) follows Poisson distribution and cannot be handled by the proposed method. The main difference from the synthetic case is that the pixel values are always between 0 and 255, which cuts off the tails of the noise distribution and makes the PDF different from a Gaussian.

To obtain test images with visually apparent noise, we deliberately took pictures of the same scene in a dark environment using low exposure and high ISO. Such setup amplifies the noise, see Fig. 10 left for an example. Since it was not possible to capture the reference clear image directly, we estimated it by a time-averaging of twenty noisy images of the same scene, see Fig. 10 right.

We calculated three 2D histograms (R–G, R–B, and G–B) of each noisy image and the denoised image and computed the invariants. The ratio between the invariants is plotted in Fig. 11. Ideally, it should be close to one. We can, however, observe some oscillations around this theoretical value. This is caused by several factors. The actual noise distribution is not exactly normal (the normality hypothesis was in all cases rejected by the Pearson's test) and the convolution model between the clear and noisy histogram is not valid in boundary regions of the color space. Still, the invariants are relatively stable (especially comparing to moments and other common PDF characteristics) and provide a noise-robust description of the histogram, which can be used for instance in histogram-based image retrieval systems.

8.3 Application in image retrieval

The previous experiments indicated that one of the potential application areas of the proposed convolution invariants could be a content-based image retrieval (CBIR). CBIR methods often relies on histograms, because two images with similar histograms are mostly perceived as similar by humans (Pass and Zabih 1996; Wang and Healey 1998; Swain and Ballard 1991). Another attractive property of the histogram is 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. Current CBIR methods based on comparing histograms are sensitive to noise in the images. We already demonstrated that an additive noise leads to a histogram smoothing, which results in a drop of the retrieval performance because different histograms tend to be more and more similar to each other.



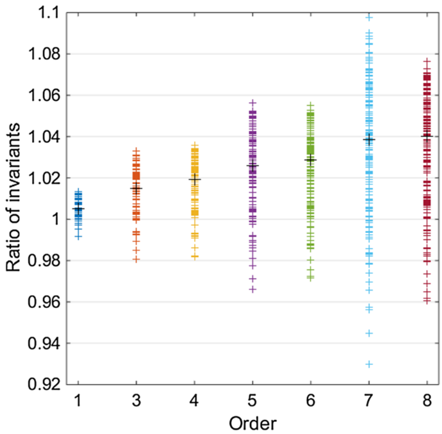


Fig. 11 The ratio of the invariants up to the order 8 of noisy and clear images for histogram of red and blue channels (left), red and green channels (right) and green and blue channels (bottom), respectively. Black crosses denote the median of the invariants (Color figure online)

We envisage the use of the proposed invariants as noise-robust descriptors of multidimensional histograms, similarly as the authors of Höschl IV and Flusser (2016) used the 1D convolution invariants for graylevel histogram recognition. The new invariants could be helpful in the case of noisy database and/or noisy query images (see Fig. 12 for the proposed method outline).

9 Conclusion

We proposed a new method for description of random variables, which is robust to an additive Gaussian noise. The method is based on the fact that the PDF of the noisy variable is a convolution of the PDF of the original unobservable variable and the PDF of the noise.

We constructed a projection operator onto the set of all Gaussian probability density functions, removed the Gaussian part of the functions and described the complement by

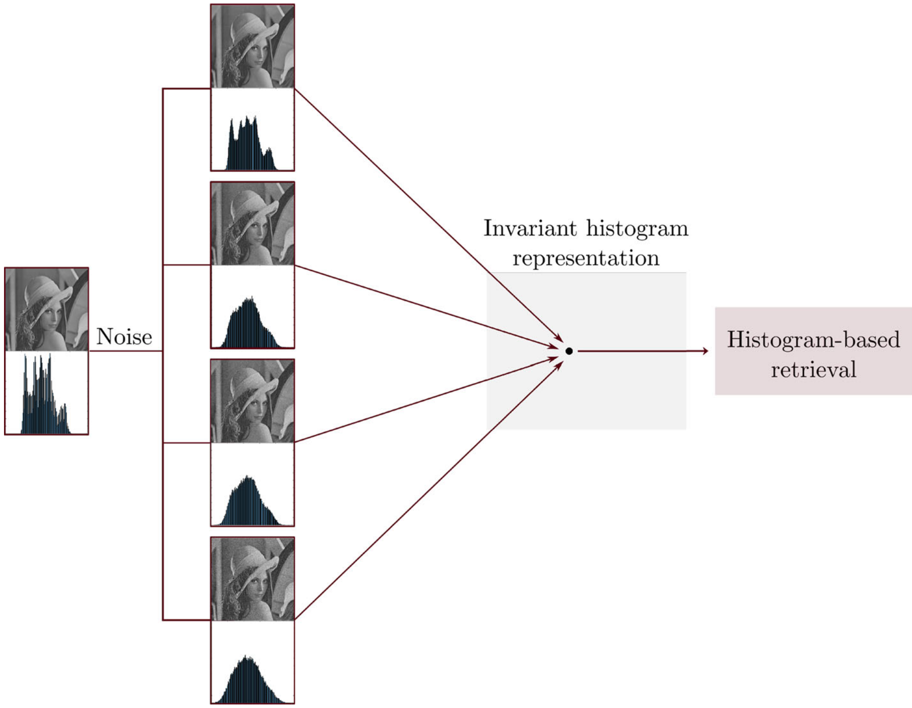


Fig. 12 Noise-robust CBIR. From left to right: original image and its histogram, noisy images with smoothed histograms, representation of the histograms by the proposed convolution invariants, image retrieval based on histogram similarity measured by the invariants. The actual implementation works with color images and multidimensional histograms

invariants composed of moments. The method does not require any estimation of the noise parameters, which makes it attractive for practical usage The 2D case was discussed in more details because of its importance in applications. The invariance property was demonstrated on experiments from image processing area.

A Explicit formula for Gaussian moments in two dimensions

In this Appendix, we present the derivation of the explicit formula for 2D central moments of the Gaussian probability density function $f_N(\mathbf{x})$ with the covariance matrix

$$\Sigma = \begin{pmatrix} \sigma_1 & \rho \\ \rho & \sigma_2 \end{pmatrix}.$$

It holds for the inverse matrix Σ^{-1} and its determinant

$$\Sigma^{-1} = \frac{1}{|\Sigma|}\begin{pmatrix} \sigma_2 & -\rho \\ -\rho & \sigma_1 \end{pmatrix} \equiv \begin{pmatrix} a & b \\ b & c \end{pmatrix}, \quad |\Sigma^{-1}| = ac - b^2 = \frac{1}{|\Sigma|}.$$

If $m + n$ is odd, the moments vanish due to the symmetry

$$\mathrm{m}_{mn}^{(f_N)} = 0.$$

For $m+n$ even we have

$$\begin{aligned}
\mathrm{m}_{mn}^{(f_N)} &= \frac{1}{2\pi\sqrt{|\Sigma|}} \iint\limits_{\mathbb{R}^2} x^m y^n \mathrm{e}^{-\frac{1}{2}(ax^2+2bxy+cy^2)}\,\mathrm{d}x\,\mathrm{d}y \\
&= \frac{1}{2\pi\sqrt{|\Sigma|}} \iint\limits_{\mathbb{R}^2} x^m \mathrm{e}^{-\frac{1}{2}x^2\left(a-\frac{b^2}{c}\right)} y^n \mathrm{e}^{-\frac{1}{2}\left(y+\frac{b}{c}x\right)^2 c}\,\mathrm{d}x\,\mathrm{d}y = \\
&= \left| \begin{array}{r} y+\frac{b}{c}x = u \\ x = v \end{array} \right| = \frac{1}{2\pi\sqrt{|\Sigma|}} \iint\limits_{\mathbb{R}^2} v^m \mathrm{e}^{-\frac{1}{2}v^2\left(a-\frac{b^2}{c}\right)} \left(u-\frac{b}{c}v\right)^n \mathrm{e}^{-\frac{1}{2}u^2 c}\,\mathrm{d}u\,\mathrm{d}v.
\end{aligned}$$

We can separate the integrals and use the formula for 1D moments of Gaussian function:

$$\begin{aligned}
&= \frac{1}{2\pi\sqrt{|\Sigma|}} \sum_{k=0}^{n} \binom{n}{k} \left(-\frac{b}{c}\right)^{n-k} \int\limits_{\mathbb{R}} u^k \mathrm{e}^{-\frac{1}{2}u^2 c}\,\mathrm{d}u \int\limits_{\mathbb{R}} v^{m+n-k} \mathrm{e}^{-\frac{1}{2}v^2\left(a-\frac{b^2}{c}\right)}\,\mathrm{d}v \\
&= \frac{1}{\sqrt{|\Sigma|}} \sum_{\substack{k=0,\\ k \text{ even}}}^{n} \binom{n}{k} \left(\frac{-b}{c}\right)^{n-k} \left(\frac{1}{a-\frac{b^2}{c}}\right)^{\frac{m+n-k+1}{2}} (m+n-k-1)!! \left(\frac{1}{c}\right)^{\frac{k+1}{2}} (k-1)!! \\
&= \sum_{\substack{k=0,\\ k \text{ even}}}^{n} \binom{n}{k} \left(\frac{-b}{|\Sigma^{-1}|}\right)^{n-k} \left(\frac{c}{|\Sigma^{-1}|}\right)^{\frac{m-n}{2}} |\Sigma|^{k/2} (m+n-k-1)!!(k-1)!! \\
&= \sum_{i=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{2i} \rho^{n-2i} \sigma_1^{\frac{m-n}{2}} \left(\sigma_1\sigma_2-\rho^2\right)^i (m+n-2i-1)!!(2i-1)!! \\
&= \sum_{i=0}^{\lfloor \frac{n}{2} \rfloor} \sum_{j=0}^{i} (-1)^{i-j} \binom{n}{2i} \binom{i}{k} (m+n-2i-1)!!(2i-1)!! \rho^{n-2j} \sigma_1^{\frac{m-n}{2}+j} \sigma_2^{j}. \qquad (24)
\end{aligned}$$

We may reduce the quadratic form in the exponent to a sum of squares in the following way

$$\frac{1}{2\pi\sqrt{|\Sigma|}} \iint\limits_{\mathbb{R}^2} y^n \mathrm{e}^{-\frac{1}{2}y^2\left(c-\frac{b^2}{a}\right)} x^m \mathrm{e}^{-\frac{1}{2}\left(x+\frac{b}{a}y\right)^2 a}\,\mathrm{d}x\,\mathrm{d}y.$$

Then using the substitution

$$\begin{aligned} x+\tfrac{b}{a}y &= u \\ y &= v \end{aligned}$$

another formula for moments of bivariate Gaussian distribution is obtained

$$\mathrm{m}_{mn}^{(f_N)} = \sum_{i=0}^{\lfloor \frac{m}{2} \rfloor} \sum_{j=0}^{i} (-1)^{i-j} \binom{m}{2i} \binom{i}{j} (m+n-2i-1)!!(2i-1)!! \rho^{m-2j} \sigma_1^{j} \sigma_2^{\frac{n-m}{2}+j}. \qquad (25)$$

When we compare these two results, it is obvious that the coefficients of negative powers must be zero. Hence, moments are composed of positive powers of the elements of covariance matrix only

$$\mathrm{m}_{mn}^{(f_N)} = \sum_{i=0}^{\lfloor\frac{m}{2}\rfloor}\sum_{\substack{j=0\\ j\geq\frac{m-n}{2}}}^{i}(-1)^{i-j}\binom{m}{2i}\binom{i}{j}(m+n-2i-1)!!\rho^{m-2j}\sigma_1^j\sigma_2^{\frac{n-m}{2}+j}. \tag{26}$$

B Proof of the equivalence

Let us show that Formulas (22) and (23) for convolution invariants are equivalent. The proof is done by induction.

Proof $A_{00} = 1$ in Formula (22) as well as in Formula (23).

Let us assume $(m, n),\ m + n > 0$. From the induction assumption, the explicit formula is valid for all indices (p, q), where $p \leq m,\ q \leq n$ and $(p, q) \neq (m, n)$.

$$\begin{aligned}
A_{mn} &= m_{mn} - \sum_{\substack{l=0\\ l+k\neq 0,\\ l+k \text{ even}}}^{m}\sum_{k=0}^{n}\binom{m}{l}\binom{n}{k}\sum_{i=0}^{\lfloor\frac{k}{2}\rfloor}\sum_{\substack{j=0\\ j\geq\frac{k-l}{2}}}^{i}(-1)^{i-j}\binom{k}{2i}\binom{i}{j}(l+k-2i-1)!!(2i-1)!!\\
&\quad\cdot m_{11}^{k-2j}m_{20}^{\frac{l-k}{2}+j}m_{02}^{j}A_{m-l,n-k} =\\
&= m_{mn} - \sum_{\substack{l=0\\ l+k\neq 0,\\ l+k \text{ even}}}^{m}\sum_{k=0}^{n}\binom{m}{l}\binom{n}{k}\sum_{i=0}^{\lfloor\frac{k}{2}\rfloor}\sum_{\substack{j=0\\ j\geq\frac{k-l}{2}}}^{i}(-1)^{i-j}\binom{k}{2i}\binom{i}{j}\\
&\quad\cdot (l+k-2i-1)!!(2i-1)!!\,m_{11}^{k-2j}m_{20}^{\frac{l-k}{2}+j}m_{02}^{j}\\
&\quad\cdot\sum_{\substack{s=0\\ s+t \text{ even}}}^{n-k}\sum_{t=0}^{m-l}(-1)^{\frac{s+t}{2}}\binom{m-l}{t}\binom{n-k}{s}\sum_{\alpha=0}^{\lfloor\frac{s}{2}\rfloor}\sum_{\substack{\beta=0\\ \beta\geq\frac{s-t}{2}}}^{\alpha}(-1)^{\alpha-\beta}\binom{s}{2\alpha}\binom{\alpha}{\beta}\\
&\quad(2\alpha-1)!!(s+t-2\alpha-1)!!\\
&\quad\cdot m_{11}^{s-2\beta}m_{20}^{\frac{t-s}{2}+\beta}m_{02}^{\beta}m_{m-l-t,n-k-s} =\\
&= m_{mn} - \sum_{\substack{l=0\\ l+k\neq 0,\\ l+k \text{ even}}}^{m}\sum_{k=0}^{n}\frac{m!}{l!(m-l)!}\frac{n!}{k!(n-k)!}\sum_{i=0}^{\lfloor\frac{k}{2}\rfloor}\sum_{\substack{j=0\\ j\geq\frac{k-l}{2}}}^{i}(-1)^{i-j}\binom{k}{2i}\binom{i}{j}(2i-1)!!\\
&\quad\cdot (l+k-2i-1)!!\\
&\quad\cdot m_{11}^{k-2j}m_{20}^{\frac{l-k}{2}+j}m_{02}^{j}\sum_{\substack{s=0\\ s+t \text{ even}}}^{n-k}\sum_{t=0}^{m-l}(-1)^{\frac{s+t}{2}}\frac{(m-l)!}{t!(m-l-t)!}\frac{(n-k)!}{s!(n-k-s)!}
\end{aligned}$$

$$\cdot \sum_{\substack{\alpha=0}}^{\lfloor \frac{s}{2} \rfloor} \sum_{\substack{\beta=0 \\ \beta \geq \frac{s-t}{2}}}^{\alpha} (-1)^{\alpha-\beta} \binom{s}{2\alpha} \binom{\alpha}{\beta} (2\alpha-1)!!(s+t-2\alpha-1)!! m_{11}^{s-2\beta} m_{20}^{\frac{t-s}{2}+\beta}$$

$$\cdot m_{02}^{\beta} m_{m-l-t,n-k-s}$$

$$= \left| \begin{smallmatrix} p=k+s \\ q=t+l \end{smallmatrix} \right| = m_{mn} - \sum_{\substack{l=0 \\ l+k \neq 0, \\ l+k \text{ even}}}^{m} \sum_{k=0}^{n} \frac{m!}{l!} \frac{n!}{k!} \sum_{i=0}^{\lfloor \frac{k}{2} \rfloor} \sum_{\substack{j=0 \\ j \geq \frac{k-l}{2}}}^{i} (-1)^{i-j} \binom{k}{2i} \binom{i}{j}$$

$$\cdot (l+k-2i-1)!!(2i-1)!!$$

$$\cdot m_{11}^{k-2j} m_{20}^{\frac{l-k}{2}+j} m_{02}^{j} \sum_{\substack{p=k \\ p+q \text{ even}}}^{n} \sum_{q=l}^{m} \frac{(-1)^{\frac{p+q-k-l}{2}}}{(q-l)!(m-q)!} \frac{1}{(p-k)!(n-p)!} \sum_{\alpha=0}^{\lfloor \frac{p-k}{2} \rfloor} \sum_{\substack{\beta=0 \\ \beta \geq \frac{p-q+l-k}{2}}}^{\alpha} (-1)^{\alpha-\beta}$$

$$\cdot \binom{p-k}{2\alpha} \binom{\alpha}{\beta} (2\alpha-1)!!$$

$$\cdot (p+q-k-l-2\alpha-1)!!\, m_{11}^{p-k-2\beta} m_{20}^{\frac{q-p+k-l}{2}+\beta} m_{02}^{\beta} m_{m-q,n-p} =$$

$$= m_{mn} - \sum_{\substack{l=0 \\ l+k \neq 0, \\ l+k \text{ even}}}^{m} \sum_{k=0}^{n} \sum_{\substack{p=k \\ p+q \text{ even}}}^{n} \sum_{q=l}^{m} \binom{m}{q} \binom{q}{l} \binom{n}{p} \binom{p}{k} (-1)^{\frac{p+q-k-l}{2}} m_{m-q,n-p}$$

$$\cdot \sum_{i=0}^{\lfloor \frac{k}{2} \rfloor} \sum_{\substack{j=0 \\ j \geq \frac{k-l}{2}}}^{i} (-1)^{i-j} \binom{k}{2i} \binom{i}{j}$$

$$\cdot (l+k-2i-1)!!(2i-1)!! \sum_{\alpha=0}^{\lfloor \frac{p-k}{2} \rfloor} \sum_{\substack{\beta=0 \\ \beta \geq \frac{p-q+l-k}{2}}}^{\alpha} (-1)^{\alpha-\beta} \binom{p-k}{2\alpha} \binom{\alpha}{\beta}$$

$$\cdot (2\alpha-1)!!(p+q-k-l-2\alpha-1)!!$$

$$\cdot m_{11}^{p-2j-2\beta} m_{20}^{\frac{q-p}{2}+j+\beta} m_{02}^{j+\beta} =$$

$$= m_{mn} - \sum_{\substack{p=0 \\ l+k \neq 0, p+q \neq 0 \\ l+k \text{ even}, p+q \text{ even}}}^{n} \sum_{k=0}^{p} \sum_{q=0}^{m} \sum_{l=0}^{q} \binom{m}{q} \binom{q}{l} \binom{n}{p} \binom{p}{k} (-1)^{\frac{p+q-k-l}{2}} m_{m-q,n-p}$$

$$\cdot \sum_{i=0}^{\lfloor \frac{k}{2} \rfloor} \sum_{\substack{j=0 \\ j \geq \frac{k-l}{2}}}^{i} (-1)^{i-j} \binom{k}{2i} \binom{i}{j}$$

$$\cdot (l+k-2i-1)!!(2i-1)!! \sum_{\substack{\alpha=0 \\ \beta \geq \frac{p-q+l-k}{2}}}^{\lfloor \frac{p-k}{2} \rfloor} \sum_{\beta=0}^{\alpha} (-1)^{\alpha-\beta} \binom{p-k}{2\alpha} \binom{\alpha}{\beta} (2\alpha-1)!!$$

$$\cdot (p+q-k-l-2\alpha-1)!!$$

$$\cdot m_{11}^{p-2j-2\beta} m_{20}^{\frac{q-p}{2}+j+\beta} m_{02}^{j+\beta} =$$

$$= m_{mn} - \sum_{\substack{p=0 \\ p+q\neq 0, \\ p+q \text{ even}}}^{n} \sum_{q=0}^{m} (-1)^{\frac{p+q}{2}} \binom{m}{q} \binom{n}{p} m_{m-q,n-p} \sum_{\substack{k=0 \\ k+l \text{ even} \\ k+l\neq 0}}^{p} \sum_{l=0}^{q} (-1)^{\frac{k+l}{2}} \binom{q}{l} \binom{p}{k}$$

$$\cdot \sum_{\substack{i=0 \\ j\geq \frac{k-l}{2}}}^{\lfloor \frac{k}{2} \rfloor} \sum_{j=0}^{i} (-1)^{i-j} \binom{k}{2i} \binom{i}{j}$$

$$\cdot (2i-1)!!(l+k-2i-1)!! m_{11}^{k-2j} m_{20}^{\frac{l-k}{2}+j} m_{02}^{j} \sum_{\substack{\alpha=0 \\ \beta\geq \frac{p-q-k+l}{2}}}^{\lfloor \frac{p-k}{2} \rfloor} \sum_{\beta=0}^{\alpha} (-1)^{\alpha-\beta} \binom{p-k}{2\alpha} \binom{\alpha}{\beta}$$

$$\cdot (2\alpha-1)!!(p+q-k-l-2\alpha-1)!! \, m_{11}^{p-k-2\beta} m_{20}^{\frac{q-p+k-l}{2}+\beta} m_{02}^{\beta} =$$

$$= m_{mn} - \sum_{\substack{p=0 \\ p+q\neq 0, \\ p+q \text{ even}}}^{n} \sum_{q=0}^{m} (-1)^{\frac{p+q}{2}} \binom{m}{q} \binom{n}{p} m_{m-q,n-p}$$

$$\cdot \Bigg[\sum_{\substack{k=0 \\ k+l \text{ even}}}^{p} \sum_{l=0}^{q} (-1)^{\frac{k+l}{2}} \binom{p}{k} \binom{q}{l} m_{l,k}^{(G)} m_{q-l,p-k}^{(G)}$$

$$- \sum_{\substack{\alpha=0 \\ \beta\geq \frac{p-q}{2}}}^{\lfloor \frac{p}{2} \rfloor} \sum_{\beta=0}^{\alpha} (-1)^{\alpha-\beta} \binom{p}{2\alpha} \binom{\alpha}{\beta} (2\alpha-1)!!(p+q-2\alpha-1)!! \, m_{11}^{p-2\beta} m_{20}^{\frac{q-p}{2}+\beta} m_{02}^{\beta} \Bigg] =$$

$$= \sum_{\substack{l=0 \\ l+k \text{ even}}}^{m} \sum_{k=0}^{n} (-1)^{\frac{k+l}{2}} \binom{m}{l} \binom{n}{k} \sum_{\substack{i=0 \\ j\geq \frac{k-l}{2}}}^{\lfloor \frac{k}{2} \rfloor} \sum_{j=0}^{i} (-1)^{i-j} \binom{k}{2i} \binom{i}{j} (l+k-2i-1)!!(2i-1)!!$$

$$\cdot m_{11}^{k-2j} m_{20}^{\frac{l-k}{2}+j} m_{02}^{j} m_{m-l,n-k}.$$

It remains to prove that for $p+q > 0,\ p+q$ even, it holds

$$\sum_{\substack{k=0 \\ k+l \text{ even}}}^{p} \sum_{l=0}^{q} (-1)^{\frac{k+l}{2}} \binom{p}{k} \binom{q}{l} m_{l,k}^{(f_N)} m_{q-l,p-k}^{(f_N)} = 0. \tag{27}$$

For $\frac{p+q}{2}$ being odd, the proof is trivial because every combination is present twice with the opposite signs. Thus, all terms vanish.

$$\begin{aligned}
k = a, l = b \Rightarrow \qquad & (-1)^{\frac{a+b}{2}} \binom{p}{a}\binom{q}{b} m_{b,a} m_{q-b,p-a} \\
k = p-a, l = q-b \Rightarrow \qquad & (-1)^{\frac{p+q-(a+b)}{2}} \binom{p}{p-a}\binom{q}{q-b} m_{q-b,p-a} m_{b,a} = \\
& -\left[(-1)^{\frac{a+b}{2}} \binom{p}{a}\binom{q}{b} m_{b,a} m_{q-b,p-a}\right]
\end{aligned}$$

For $\frac{p+q}{2}$ even we have

$$\sum_{\substack{k=0\\k+l \text{ even}}}^{p}\sum_{l=0}^{q}(-1)^{\frac{k+l}{2}}\binom{p}{k}\binom{q}{l} m_{l,k}^{(f_N)} m_{q-l,p-k}^{(f_N)} =$$

$$\begin{aligned}
&= \sum_{\substack{k=0\\k+l \text{ even}}}^{p}\sum_{l=0}^{q}(-1)^{\frac{k+l}{2}}\binom{p}{k}\binom{q}{l}\sum_{i=0}^{\lfloor\frac{k}{2}\rfloor}\sum_{\substack{t=0\\t\geq\frac{k-l}{2}}}^{i}(-1)^{i-t}\binom{k}{2i}\binom{i}{t}(l+k-2i-1)!! \\
&\cdot (2i-1)!!\, m_{11}^{k-2t} m_{20}^{\frac{l-k}{2}+t} m_{02}^{t} \\
&\cdot \sum_{\substack{s=0\\r\geq\frac{p-q-k+l}{2}}}^{\lfloor\frac{p-k}{2}\rfloor}\sum_{r=0}^{s}(-1)^{s-r}\binom{p-k}{2s}\binom{s}{r}(p+q-l-k-2s-1)!! \\
&\cdot (2s-1)!!\, m_{11}^{p-k-2r} m_{20}^{\frac{q-l-p+k}{2}+r} m_{02}^{r} = \\
&= \sum_{\substack{k=0\\k+l \text{ even} \wedge t\geq\frac{k-l}{2} \wedge r\geq\frac{p-q-k+l}{2}}}^{p}\sum_{l=0}^{q}\sum_{i=0}^{\lfloor\frac{k}{2}\rfloor}\sum_{t=0}^{i}\sum_{s=0}^{\lfloor\frac{p-k}{2}\rfloor}\sum_{r=0}^{s}\binom{p}{k}\binom{q}{l}\binom{k}{2i}\binom{i}{t}\binom{p-k}{2s}\binom{s}{r}(2i-1)!!(2s-1)!! \\
&\cdot (l+k-2i-1)!!(p+q-l-k-2s-1)!!(-1)^{\frac{k+l}{2}+i-t+s-r} \\
&\cdot m_{11}^{p-2t-2r} m_{20}^{\frac{q-p}{2}+r+t} m_{02}^{r+t} \\
&= \sum_{\substack{k=0\\k+l \text{ even} \wedge t\geq\frac{k-l}{2} \wedge r\geq\frac{p-q-k+l}{2}}}^{p}\sum_{l=0}^{q}\sum_{t=0}^{\lfloor\frac{k}{2}\rfloor}\sum_{i=t}^{\lfloor\frac{k}{2}\rfloor}\sum_{s=0}^{\lfloor\frac{p-k}{2}\rfloor}\sum_{r=0}^{s}\binom{p}{k}\binom{q}{l}\binom{k}{2i}\binom{i}{t}\binom{p-k}{2s}\binom{s}{r}(2i-1)!!(2s-1)!! \\
&\cdot (l+k-2i-1)!!(p+q-l-k-2s-1)!!(-1)^{\frac{k+l}{2}+i-t+s-r} m_{11}^{p-2t-2r} m_{20}^{\frac{q-p}{2}+r+t} m_{02}^{r+t}
\end{aligned}$$

$$= \left|\begin{matrix} k=0:p,\ t=0:\lfloor\frac{k}{2}\rfloor \Rightarrow \\ t=0:\lfloor\frac{p}{2}\rfloor,\ k=2t:p \end{matrix}\right| \begin{matrix} s=0:\lfloor\frac{p-k}{2}\rfloor,\ r=0:s \Rightarrow \\ r=0:\lfloor\frac{p-k}{2}\rfloor,\ s=r:\lfloor\frac{p-k}{2}\rfloor \end{matrix}\right| =$$

$$= \sum_{t=0}^{\lfloor \frac{p}{2} \rfloor} \sum_{k=2t}^{p} \sum_{l=0}^{q} \sum_{i=t}^{\lfloor \frac{k}{2} \rfloor} \sum_{r=0}^{\lfloor \frac{p-k}{2} \rfloor} \sum_{\substack{s=r \\ k+l \text{ even} \wedge t \geq \frac{k-l}{2} \wedge r \geq \frac{p-q-k+l}{2}}}^{\lfloor \frac{p-k}{2} \rfloor} \binom{p}{k}\binom{q}{l}\binom{k}{2i}\binom{i}{t}\binom{p-k}{2s}\binom{s}{r}(2i-1)!!(2s-1)!!$$

$$\cdot (l+k-2i-1)!!(p+q-l-k-2s-1)!!(-1)^{\frac{k+l}{2}+i-t+s-r} m_{11}^{p-2t-2r}$$

$$\cdot m_{20}^{\frac{q-p}{2}+r+t} m_{02}^{r+t}$$

$$= \left| \begin{array}{c} k = 2t : p,\ r = 0 : \lfloor \frac{p-k}{2} \rfloor \Rightarrow \\ r = 0 : \lfloor \frac{p-2t}{2} \rfloor,\ k = 2t : p-2r \end{array} \right| =$$

$$= \sum_{t=0}^{\lfloor \frac{p}{2} \rfloor} \sum_{r=0}^{\lfloor \frac{p-2t}{2} \rfloor} \sum_{k=2t}^{p-2r} \sum_{l=0}^{q} \sum_{i=t}^{\lfloor \frac{k}{2} \rfloor} \sum_{\substack{s=r \\ k+l \text{ even} \wedge t \geq \frac{k-l}{2} \wedge r \geq \frac{p-q-k+l}{2}}}^{\lfloor \frac{p-k}{2} \rfloor} \binom{p}{k}\binom{q}{l}\binom{k}{2i}\binom{i}{t}\binom{p-k}{2s}\binom{s}{r}(2i-1)!!(2s-1)!!$$

$$\cdot (l+k-2i-1)!!(p+q-l-k-2s-1)!!(-1)^{\frac{k+l}{2}+i-t+s-r} m_{11}^{p-2t-2r}$$

$$\cdot m_{20}^{\frac{q-p}{2}+r+t} m_{02}^{r+t}$$

$$= \sum_{t=0}^{\lfloor \frac{p}{2} \rfloor} \sum_{r=0}^{\lfloor \frac{p-2t}{2} \rfloor} m_{11}^{p-2t-2r} m_{20}^{\frac{q-p}{2}+r+t} m_{02}^{r+t} \sum_{k=2t}^{p-2r} \sum_{l=0}^{q} \sum_{i=t}^{\lfloor \frac{k}{2} \rfloor} \sum_{\substack{s=r \\ k+l \text{ even} \wedge t \geq \frac{k-l}{2} \wedge r \geq \frac{p-q-k+l}{2}}}^{\lfloor \frac{p-k}{2} \rfloor} (-1)^{\frac{k+l}{2}+i-t+s-r} \binom{p}{k}\binom{q}{l}$$

$$\cdot \binom{k}{2i}\binom{i}{t}\binom{p-k}{2s}\binom{s}{r}(2i-1)!!(2s-1)!!(l+k-2i-1)!!(p+q-l-k-2s-1)!!$$

$$= \left| \begin{array}{c} t + r = N \Rightarrow r = N - t \\ t = 0 : \lfloor \frac{p}{2} \rfloor \Rightarrow \\ N = t : \lfloor \frac{p}{2} \rfloor \end{array} \right| =$$

$$= \sum_{t=0}^{\lfloor \frac{p}{2} \rfloor} \sum_{N=t}^{\lfloor \frac{p}{2} \rfloor} m_{11}^{p-2N} m_{20}^{\frac{q-p}{2}+N} m_{02}^{N} \sum_{k=2t}^{p-2N+2t} \sum_{l=0}^{q} \sum_{i=t}^{\lfloor \frac{k}{2} \rfloor} \sum_{\substack{s=N-t \\ k+l \text{ even} \wedge N - \frac{p-k-q+l}{2} \geq t \geq \frac{k-l}{2}}}^{\lfloor \frac{p-k}{2} \rfloor} (-1)^{\frac{k+l}{2}+i+s-N} \binom{p}{k}\binom{q}{l}\binom{k}{2i}\binom{i}{t}$$

$$\cdot \binom{p-k}{2s}\binom{s}{N-t}(2i-1)!!(2s-1)!!(l+k-2i-1)!!(p+q-l-k-2s-1)!!$$

$$= \left| \begin{array}{c} t = 0 : \lfloor \frac{p}{2} \rfloor, N = t : \lfloor \frac{p}{2} \rfloor \Rightarrow \\ N = 0 : \lfloor \frac{p}{2} \rfloor, t = 0 : N \end{array} \right| =$$

$$= \sum_{N=0}^{\lfloor \frac{p}{2} \rfloor} \sum_{t=0}^{N} m_{11}^{p-2N} m_{20}^{\frac{q-p}{2}+N} m_{02}^{N} \sum_{k=2t}^{p-2N+2t} \sum_{l=0}^{q} \sum_{i=t}^{\lfloor \frac{k}{2} \rfloor} \sum_{\substack{s=N-t \\ k+l \text{ even} \wedge N - \frac{p-k-q+l}{2} \geq t \geq \frac{k-l}{2}}}^{\lfloor \frac{p-k}{2} \rfloor} (-1)^{\frac{k+l}{2}+i+s-N} \binom{p}{k}\binom{q}{l}\binom{i}{t}$$

$$\cdot \binom{k}{2i}\binom{p-k}{2s}\binom{s}{N-t}(2i-1)!!(2s-1)!!(l+k-2i-1)!!(p+q-l-k-2s-1)!! =$$

$$= \left| \begin{array}{c} k - 2t = m \Rightarrow k = m + 2t \\ k = 2t : p - 2N + 2t \Rightarrow \\ m = 0 : p - 2N \end{array} \right| =$$

$$= \sum_{N=0}^{\lfloor \frac{p}{2} \rfloor} m_{11}^{p-2N} m_{20}^{\frac{q-p}{2}+N} m_{02}^{N} \sum_{t=0}^{N} \sum_{\substack{m=0 \\ m+l \text{ even} \,\wedge\, N-\frac{p-q}{2} \geq \frac{l-m}{2} \geq 0}}^{p-2N} \sum_{l=0}^{q} \sum_{i=t}^{\lfloor \frac{m+2t}{2} \rfloor} \sum_{s=N-t}^{\lfloor \frac{p-m-2t}{2} \rfloor} (-1)^{\frac{m+l}{2}+t+i+s-N}$$

$$\cdot \binom{p}{m+2t} \binom{q}{l} \binom{m+2t}{2i} \binom{i}{t}$$

$$\cdot \binom{s}{N-t} \binom{p-m-2t}{2s} (2i-1)!!(2s-1)!!$$

$$\cdot (l+m+2t-2i-1)!!(p+q-l-m-2t-2s-1)!!$$

$$= \left| \begin{array}{c} i-t=j \Rightarrow i=j+t \\ i=t : \lfloor \frac{m+2t}{2} \rfloor \Rightarrow \\ j=0 : \lfloor \frac{m}{2} \rfloor \end{array} \right| =$$

$$= \sum_{N=0}^{\lfloor \frac{p}{2} \rfloor} m_{11}^{p-2N} m_{20}^{\frac{q-p}{2}+N} m_{02}^{N} \sum_{t=0}^{N} \sum_{\substack{m=0 \\ m+l \text{ even} \,\wedge\, N-\frac{p-q}{2} \geq \frac{l-m}{2} \geq 0}}^{p-2N} \sum_{l=0}^{q} \sum_{j=0}^{\lfloor \frac{m}{2} \rfloor} \sum_{s=N-t}^{\lfloor \frac{p-m-2t}{2} \rfloor} (-1)^{\frac{m+l}{2}+j+2t+s-N}$$

$$\cdot \binom{p}{m+2t} \binom{q}{l} \binom{m+2t}{2j+2t} \binom{j+t}{t}$$

$$\cdot \binom{s}{N-t} \binom{p-m-2t}{2s} (2j+2t-1)!!(2s-1)!!(l+m-2j-1)!!$$

$$\cdot (p+q-l-m-2t-2s-1)!!$$

$$= \left| \begin{array}{c} s-N+t=k \Rightarrow s=N-t+k \\ s=N-t : \lfloor \frac{p-2m-2t}{2} \rfloor \Rightarrow \\ k=0 : \lfloor \frac{p-2m-2N}{2} \rfloor \end{array} \right| =$$

$$= \sum_{N=0}^{\lfloor \frac{p}{2} \rfloor} m_{11}^{p-2N} m_{20}^{\frac{q-p}{2}+N} m_{02}^{N} \sum_{t=0}^{N} \sum_{\substack{m=0 \\ m+l \text{ even} \,\wedge\, N-\frac{p-q}{2} \geq \frac{l-m}{2} \geq 0}}^{p-2N} \sum_{l=0}^{q} \sum_{j=0}^{\lfloor \frac{m}{2} \rfloor} \sum_{k=0}^{\lfloor \frac{p-m-2N}{2} \rfloor} (-1)^{\frac{m+l}{2}+j+k+t} \binom{p}{m+2t} \binom{q}{l}$$

$$\cdot \binom{m+2t}{2j+2t} \binom{j+t}{t} \binom{N-t+k}{N-t} \binom{p-m-2t}{2N-2t+2k} (2j+2t-1)!!(2N-2t+2k-1)!!$$

$$\cdot (l+m-2j-1)!!(p+q-l-m-2N-2k-1)!!$$

$$= \sum_{N=0}^{\lfloor \frac{p}{2} \rfloor} m_{11}^{p-2N} m_{20}^{\frac{q-p}{2}+N} m_{02}^{N} \sum_{t=0}^{N} \sum_{\substack{m=0 \\ m+l \text{ even} \,\wedge\, N-\frac{p-q}{2} \geq \frac{l-m}{2} \geq 0}}^{p-2N} \sum_{l=0}^{q} \sum_{j=0}^{\lfloor \frac{m}{2} \rfloor} \sum_{k=0}^{\lfloor \frac{p-m-2N}{2} \rfloor} (-1)^{\frac{m+l}{2}+j+k+t} \binom{q}{l} p!$$

$$\cdot \frac{(2j+2t-1)!!}{(2j+2t)!(m-2j)!} \frac{(j+t)!}{t!j!} \frac{(N-t+k)!}{(N-t)!k!} \frac{(2N-2t+2k-1)!!}{(2N-2t+2k)!(p-m-2N-2k)!}$$

$$\cdot (l+m-2j-1)!!(p+q-l-m-2N-2k-1)!!$$

$$= \sum_{N=0}^{\lfloor \frac{p}{2} \rfloor} m_{11}^{p-2N} m_{20}^{\frac{q-p}{2}+N} m_{02}^{N} \sum_{t=0}^{N} \sum_{\substack{m=0 \\ m+l \text{ even} \,\wedge\, N-\frac{p-q}{2} \geq \frac{l-m}{2} \geq 0}}^{p-2N} \sum_{l=0}^{q} \sum_{j=0}^{\lfloor \frac{m}{2} \rfloor} \sum_{k=0}^{\lfloor \frac{p-m-2N}{2} \rfloor} (-1)^{\frac{m+l}{2}+j+k+t} \binom{q}{l} p!$$

$$\cdot \frac{1}{(2j+2t)!!(m-2j)!} \frac{(j+t)!}{t!j!} \frac{(N-t+k)!}{(N-t)!k!} \frac{1}{(2N-2t+2k)!!(p-m-2N-2k)!}$$

$$\cdot (l+m-2j-1)!!(p+q-l-m-2N-2k-1)!!$$

$$= \sum_{N=0}^{\lfloor \frac{p}{2} \rfloor} m_{11}^{p-2N} m_{20}^{\frac{q-p}{2}+N} m_{02}^{N} \sum_{t=0}^{N} \sum_{m=0}^{p-2N} \sum_{\substack{l=0 \\ m+l \text{ even } \wedge N-\frac{p-q}{2} \ge \frac{l-m}{2} \ge 0}}^{q} \sum_{j=0}^{\lfloor \frac{m}{2} \rfloor} \sum_{k=0}^{\lfloor \frac{p-m-2N}{2} \rfloor} (-1)^{\frac{m+l}{2}+j+k+t} \binom{q}{l} p!$$

$$\cdot \frac{1}{2^{j+t}(j+t)!(m-2j)!} \frac{(j+t)!}{t!j!} \frac{(N-t+k)!}{(N-t)!k!} \frac{1}{2^{N-t+k}(N-t+k)!(p-m-2N-2k)!}$$

$$\cdot (l+m-2j-1)!!(p+q-l-m-2N-2k-1)!!$$

$$= \sum_{N=0}^{\lfloor \frac{p}{2} \rfloor} m_{11}^{p-2N} m_{20}^{\frac{q-p}{2}+N} m_{02}^{N} \sum_{t=0}^{N} \sum_{m=0}^{p-2N} \sum_{\substack{l=0 \\ m+l \text{ even } \wedge N-\frac{p-q}{2} \ge \frac{l-m}{2} \ge 0}}^{q} \sum_{j=0}^{\lfloor \frac{m}{2} \rfloor} \sum_{k=0}^{\lfloor \frac{p-m-2N}{2} \rfloor} (-1)^{\frac{m+l}{2}+j+k+t} \binom{q}{l} p!$$

$$\cdot \frac{1}{2^j j!(m-2j)!} \frac{1}{2^{N+k}k!(p-m-2N-2k)!} \frac{1}{t!(N-t)!}$$

$$\cdot (l+m-2j-1)!!(p+q-l-m-2N-2k-1)!!$$

$$= \sum_{N=0}^{\lfloor \frac{p}{2} \rfloor} m_{11}^{p-2N} m_{20}^{\frac{q-p}{2}+N} m_{02}^{N} \sum_{m=0}^{p-2N} \sum_{\substack{l=0 \\ m+l \text{ even } \wedge N-\frac{p-q}{2} \ge \frac{l-m}{2} \ge 0}}^{q} \sum_{j=0}^{\lfloor \frac{m}{2} \rfloor} \sum_{k=0}^{\lfloor \frac{p-m-2N}{2} \rfloor} (-1)^{\frac{m+l}{2}+j+k} \binom{q}{l} p!$$

$$\cdot \frac{1}{(2j)!!(m-2j)!} \frac{1}{(2k)!!(p-m-2N-2k)!2^N N!}$$

$$\cdot (l+m-2j-1)!!(p+q-l-m-2N-2k-1)!! \sum_{t=0}^{N} (-1)^t \binom{N}{t} = \tag{28}$$

$$= \sum_{N=0}^{\lfloor \frac{p}{2} \rfloor} m_{11}^{p-2N} m_{20}^{\frac{q-p}{2}+N} m_{02}^{N} \sum_{m=0}^{p-2N} \sum_{\substack{l=0 \\ m+l \text{ even } \wedge N-\frac{p-q}{2} \ge \frac{l-m}{2} \ge 0}}^{q} \sum_{j=0}^{\lfloor \frac{m}{2} \rfloor} \sum_{k=0}^{\lfloor \frac{p-m-2N}{2} \rfloor} (-1)^{\frac{m+l}{2}+j+k} \binom{q}{l} p!$$

$$\cdot \frac{(2j-1)!!}{(2j)!(m-2j)!} \frac{(2k-1)!!}{(2k)!(p-m-2N-2k)!(2N)!!}$$

$$\cdot (l+m-2j-1)!!(p+q-l-m-2N-2k-1)!!(1-1)^N. \tag{29}$$

All the terms of (29) are zero if $N > 0$. If $N = 0$, there remains the last term only

$$\sum_{m=0}^{p} \sum_{\substack{l=0 \\ m+l \text{ even } \wedge q-p \ge l-m \ge 0}}^{q} \sum_{j=0}^{\lfloor \frac{m}{2} \rfloor} \sum_{k=0}^{\lfloor \frac{p-m}{2} \rfloor} (-1)^{\frac{m+l}{2}+j+k} \binom{p}{m} \binom{q}{l} \binom{m}{2j} \binom{p-m}{2k}$$

$$\cdot (2j-1)!!(2k-1)!!(l+m-2j-1)!!(p+q-l-m-2k-1)!!. \tag{30}$$

Now we prove that this term is zero as well. This term is equivalent to

$$\sum_{\substack{m=0 \\ m+l \text{ even} \\ q-p \geq l-m \geq 0}}^{p} \sum_{l=0}^{q} (-1)^{\frac{m+l}{2}} \binom{p}{m} \binom{q}{l} \sum_{j=0}^{\lfloor \frac{m}{2} \rfloor} (-1)^j \binom{m}{2j} (l+m-2j-1)!!(2j-1)!!$$
$$\cdot \sum_{k=0}^{\lfloor \frac{p-m}{2} \rfloor} (-1)^k \binom{p-m}{2k} (p+q-l-m-2k-1)!!(2k-1)!! = \Xi. \tag{31}$$

It can be shown for $l \geq m$ using the method of generating functions described in Gould and Quaintance (2012) that

$$\sum_{j=0}^{\lfloor \frac{m}{2} \rfloor} (-1)^j \binom{m}{2j} (l+m-2j-1)!!(2j-1)!! = \frac{l!}{(l-m)!!}. \tag{32}$$

We adopt the notation from Gould and Quaintance (2012) for double factorial binomial coefficients and we recall $(p+q)/2$ is even. The previous expression can be rewritten

$$\begin{aligned}
\Xi &= \sum_{\substack{m=0 \\ m+l \text{ even} \\ q-p \geq l-m \geq 0}}^{p} \sum_{l=0}^{q} (-1)^{\frac{m+l}{2}} \binom{p}{m} \binom{q}{l} \frac{l!}{(l-m)!!} \frac{(q-l)!}{(q-l-p+m)!!} = \\
&= \frac{q!}{(q-p)!!} \sum_{\substack{m=0 \\ m+l \text{ even} \\ q-p \geq l-m \geq 0}}^{p} \sum_{l=m}^{q-p+m} (-1)^{\frac{m+l}{2}} \binom{p}{m} \left(\binom{q-p}{l-m} \right) = \left| \begin{matrix} l-m = 2j \\ j = 0 : \frac{q-p}{2} \end{matrix} \right| = \\
&= \frac{q!}{(q-p)!!} \sum_{m=0}^{p} \sum_{j=0}^{\frac{q-p}{2}} (-1)^{j+m} \binom{p}{m} \left(\binom{q-p}{2j} \right) = \\
&= \frac{q!}{(q-p)!!} \sum_{m=0}^{p} (-1)^m \binom{p}{m} \sum_{j=0}^{\frac{q-p}{2}} (-1)^j \left(\binom{q-p}{2j} \right)
\end{aligned} \tag{33}$$

The inner sum is zero if $q > p$

$$\begin{aligned}
\sum_{j=0}^{\frac{q-p}{2}} (-1)^j \left(\binom{q-p}{2j} \right) &= \sum_{j=0}^{\frac{q-p}{2}} (-1)^j \frac{(q-p)!!}{(2j)!!(q-p-2j)!!} \\
&= \sum_{j=0}^{\frac{q-p}{2}} (-1)^j \binom{\frac{q-p}{2}}{j} = (1-1)^{\frac{q-p}{2}} = 0.
\end{aligned} \tag{34}$$

For the case $q = p$ ($q-p$ must be non-negative) the inner sum equals 1 and the expression (33) is

$$p! \sum_{m=0}^{p} (-1)^m \binom{p}{m} = p!(1-1)^p \tag{35}$$

which completes the proof because it is zero whenever $p > 0$.

The formula

$$\sum_{\substack{k=0 \\ k+l \text{ even}}}^{p} \sum_{l=0}^{q} (-1)^{\frac{k+l}{2}} \binom{p}{k} \binom{q}{l} m_{l,k}^{(f_N)} m_{q-l,p-k}^{(f_N)} = 0 \tag{36}$$

holds not only for $p+q$ even but for all p and q. If $p+q$ is odd, then $m_{q-l,p-k}^{(f_N)}$ is Gaussian moment of the odd order and all the terms in summation are zero. □

References

Bar, W., & Dittrich, F. (1971). Useful formula or moment computation of normal random variables with nonzero means. *IEEE Transactions on Automatic Control*, *16*(3), 263–265.

Blacher, R. (2003). Multivariate quadratic forms of random vectors. *Journal of Multivariate Analysis*, *87*(1), 2–23.

Buades, A., Coll, B., & Morel, J. M. (2005a). A non-local algorithm for image denoising. In *IEEE computer society conference on computer vision and pattern recognition, 2005, CVPR 2005* (Vol. 2, pp. 60–65). IEEE.

Buades, A., Coll, B., & Morel, J. M. (2005b). A review of image denoising algorithms, with a new one. *Multiscale Modeling and Simulation*, *4*(2), 490–530.

Butucea, C., Comte, F., et al. (2009). Adaptive estimation of linear functionals in the convolution model and applications. *Bernoulli*, *15*(1), 69–98.

Carroll, R. J., & Hall, P. (1988). Optimal rates of convergence for deconvolving a density. *Journal of the American Statistical Association*, *83*(404), 1184–1186.

Chambolle, A., & Lions, P. L. (1997). Image recovery via total variation minimization and related problems. *Numerische Mathematik*, *76*(2), 167–188.

Chen, G., Xie, W., & Zhao, Y. (2013). Wavelet-based denoising: A brief review. In *2013 fourth international conference on intelligent control and information processing (ICICIP)* (pp. 570–574). IEEE.

Cho, D., & Bui, T. D. (2005). Multivariate statistical approach for image denoising. In *IEEE International conference on acoustics, speech, and signal processing, 2005. Proceedings (ICASSP'05)* (Vol. 4, pp. 4–589). IEEE.

Comte, F., & Lacour, C. (2011). Data-driven density estimation in the presence of additive noise with unknown distribution. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *73*(4), 601–627.

De Brabanter, K., & De Moor, B. (2012). Deconvolution in nonparametric statistics. In *ESANN*.

Efron, B. (2014). *The Bayes deconvolution problem*. Stanford: Division of Biostatistics, Stanford University.

Fan, J. (1992). Deconvolution with supersmooth distributions. *Canadian Journal of Statistics*, *20*(2), 155–169.

Flusser, J., Boldyš, J., & Zitová, B. (2003). Moment forms invariant to rotation and blur in arbitrary number of dimensions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *25*(2), 234–246.

Flusser, J., & Suk, T. (1998). Degraded image analysis: An invariant approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*(6), 590–603.

Flusser, J., Suk, T., Boldyš, J., & Zitová, B. (2015). Projection operators and moment invariants to image blurring. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *37*(4), 786–802.

Galigekere, R. R., & Swamy, M. N. S. (2006). Moment patterns in the Radon space: Invariance to blur. *Optical Engineering*, *45*(7), 0770036.

Gopalan, R., Turaga, P., & Chellappa, R. (2012). A blur-robust descriptor with applications to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *34*(6), 1220–1226.

Gould, H., & Quaintance, J. (2012). Double fun with double factorials. *Mathematics Magazine*, *85*(3), 177–192.

Höschl, C, I. V., & Flusser, J. (2016). Robust histogram-based image retrieval. *Pattern Recognition Letters*, *69*(1), 72–81.

Isserlis, L. (1918). 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.

Johannes, J., et al. (2009). Deconvolution with unknown error distribution. *The Annals of Statistics*, *37*(5A), 2301–2323.

Kappus, J., Mabon, G., et al. (2014). Adaptive density estimation in deconvolution problems with unknown error distribution. *Electronic Journal of Statistics*, *8*(2), 2879–2904.

Khireddine, A., Benmahammed, K., & Puech, W. (2007). Digital image restoration by Wiener filter in 2D case. *Advances in Engineering Software*, *38*(7), 513–516.

Makaremi, I., & Ahmadi, M. (2012). Wavelet domain blur invariants for image analysis. *IEEE Transactions on Image Processing*, *21*(3), 996–1006.
Meister, A. (2009). *Deconvolution problems in nonparametric statistics. Lecture notes in statistics* (Vol. 193). Berlin: Springer.
Motwani, M. C., Gadiya, M. C., Motwani, R. C., & Harris, F. C. (2004). Survey of image denoising techniques. In *Proceedings of GSPX* (pp. 27–30).
Ojansivu, V., & Heikkilä, J. (2007). Image registration using blur-invariant phase correlation. *IEEE Signal Processing Letters*, *14*(7), 449–452.
Pass, G., & Zabih, R. (1996). Histogram refinement for content-based image retrieval. In *Proceedings 3rd IEEE workshop on applications of computer vision WACV'96* (pp. 96–102). IEEE.
Pedone, M., Flusser, J., & Heikkilä, J. (2013). Blur invariant translational image registration for N-fold symmetric blurs. *IEEE Transactions on Image Processing*, *22*(9), 3676–3689.
Pensky, M., Vidakovic, B., et al. (1999). Adaptive wavelet estimator for nonparametric density deconvolution. *The Annals of Statistics*, *27*(6), 2033–2053.
Perona, P., & Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *12*(7), 629–639.
Schott, J. R. (2003). Kronecker product permutation matrices and their application to moment matrices of the normal distribution. *Journal of Multivariate analysis*, *87*(1), 177–190.
Song, I., & Lee, S. (2015). Explicit formulae for product moments of multivariate Gaussian random variables. *Statistics and Probability Letters*, *100*, 27–34.
Starck, J. L., Candès, E. J., & Donoho, D. L. (2002). The curvelet transform for image denoising. *IEEE Transactions on Image Processing*, *11*(6), 670–684.
Stefanski, L. A., & Carroll, R. J. (1990). Deconvolving kernel density estimators. *Statistics*, *21*(2), 169–184.
Swain, M. J., & Ballard, D. H. (1991). Color indexing. *International Journal of Computer Vision*, *7*(1), 11–32.
Teuber, T., Remmele, S., Hesser, J., & Steidl, G. (2012). Denoising by second order statistics. *Signal Processing*, *92*(12), 2837–2847.
Triantafyllopoulos, K. (2003). On the central moments of the multidimensional Gaussian distribution. *Mathematical Scientist*, *28*(2), 125–128.
Von Rosen, D. (1988). Moments for matrix normal variables. *Statistics: A Journal of Theoretical and Applied Statistics*, *19*(4), 575–583.
Wang, L., & Healey, G. (1998). Using Zernike moments for the illumination and geometry invariant classification of multispectral texture. *IEEE Transactions on Image Processing*, *7*(2), 196–203.
Zhang, H., Shu, H., Han, G. N., Coatrieux, G., Luo, L., & Coatrieux, J. L. (2010). Blurred image recognition by Legendre moment invariants. *IEEE Transactions on Image Processing*, *19*(3), 596–611.

Jitka Kostková received the M.Sc. degree in Applied Mathematical Stochastic Methods from the Czech Technical University, Faculty of Nuclear Sciences and Physical Engineering, Prague, Czech Republic, in 2015. Currently, she is a Ph.D. student in Mathematical Engineering and tutors undergraduate courses on mathematical analysis at the same university. Jitka Kostková's research interest is focused on moments and moment invariants.

Jan Flusser received the M.Sc. degree in mathematical engineering from the Czech Technical University, Prague, Czech Republic, in 1985, the Ph.D. degree in computer science from the Czechoslovak Academy of Sciences in 1990, and the DrSc. degree in technical cybernetics in 2001. Since 1985 he has been with the Institute of Information Theory and Automation, Czech Academy of Sciences, Prague. In 1995–2007, he was holding the position of a head of Department of Image Processing. Since 2007 he has been a Director of the Institute. He is a full professor of computer science at the Czech Technical University, Faculty of Nuclear Sciences and Physical Engineering, and at the Charles University, Faculty of Mathematics and Physics, Prague, Czech Republic, where he gives undergraduate and graduate courses on Digital Image Processing, Pattern Recognition, and Moment Invariants and Wavelets. Jan Flusser's research interest covers moments and moment invariants, image registration, image fusion, multichannel blind deconvolution, and super-resolution imaging. He has authored and coauthored more than 200 research publications in these areas, including the monographs Moments and Moment Invariants in Pattern Recognition (Wiley, 2009) and 2D and 3D Image Analysis by Moments (Wiley, 2016). In 2007 Jan Flusser received the Award of the Chairman of the Czech Science Foundation for the best research project and won the Prize of the Academy of Sciences of the Czech Republic for the contribution toimage fusion theory. In 2010, he was awarded by the SCOPUS 1000 Award. He received the Felber Medal of the Czech Technical University for excellent contribution to research and education in 2015 and the Praemium Academiae of the Czech Academy of Sciences for outstanding researchers in 2017.