Blur Invariants for Image Recognition

Jan Flusser[1] · Matěj Lébl[1] · Filip Šroubek[1] · Matteo Pedone[2] · Jitka Kostková[1]

Abstract

Blur is an image degradation that makes object recognition challenging. Restoration approaches solve this problem via image deblurring, deep learning methods rely on the augmentation of training sets. Invariants with respect to blur offer an alternative way of describing and recognising blurred images without any deblurring and data augmentation. In this paper, we present an original theory of blur invariants. Unlike all previous attempts, the new theory requires no prior knowledge of the blur type. The invariants are constructed in the Fourier domain by means of orthogonal projection operators and moment expansion is used for efficient and stable computation. Applying a general substitution rule, combined invariants to blur and spatial transformations are easy to construct and use. Experimental comparison to Convolutional Neural Networks shows the advantages of the proposed theory.

Keywords Blurred image · Object recognition · Blur invariants · Projection operators · Moments

1 Introduction

Real digital images often represent degraded versions of the original scene. If we want to recognize objects in such images, the recognition algorithms either must be able to suppress or remove the degradation, or they should be sufficiently robust with respect to them. One of the most common degradations is *blur*, which usually performs smoothing and

✉ Jan Flusser
flusser@utia.cas.cz

Matěj Lébl
lebl@utia.cas.cz

Filip Šroubek
sroubekf@utia.cas.cz

Matteo Pedone
matteo.pedone@oulu.fi

Jitka Kostková
kostkova@utia.cas.cz

1 Czech Academy of Sciences, Institute of Information Theory and Automation, Pod Vodárenskou věží 4, 18208 Prague 8, Czech Republic

2 Department of Computer Science and Engineering, Center for Machine Vision and Signal Analysis, University of Oulu, 90014 Oulu, Finland

suppression of high-frequency details in images. The blur may appear in images for several reasons caused by different physical phenomena. Motion blur is caused by a mutual movement of the camera and the scene and we have to deal with it for instance in licence plate recognition of fast moving cars. Out-of-focus blur appears due to incorrect focus setting and complicates face recognition in non-collaborative scenarios, character recognition, template matching and registration (including stereo analysis) of blurred scenes and numerous similar tasks. Medium turbulence blur is the result of light dispersion and random fluctuations of the refractive index during the acquisition and appears namely in ground-based astronomical imaging.

Capturing an ideal scene f by an imaging device with the point-spread function (PSF) h, the observed image g is modeled as a convolution:

$$g(\mathbf{x}) = (f * h)(\mathbf{x}). \tag{1}$$

This linear image formation model, even if it is very simple, is a reasonably accurate approximation of many imaging devices and acquisition scenarios and it holds if not globally then at least locally.

1.1 Three Approaches to Blurred Images

There exist basically three different approaches to the recognition of blurred images (see Fig. 1).

The oldest (and the least efficient) one relies on image restoration/deconvolution techniques. Image restoration inverts (1) to estimate f from its degraded version g, while the PSF may be partially known or unknown (Campisi & Egiazarian, 2007; Kundur & Hatzinakos, 1996; Rajagopalan & Chellappa, 2014). As soon as f has been recovered, the objects can be described by any standard features and recognized by traditional classifiers. However, image restoration is a time-consuming ill-posed problem and without any additional constraints infinitely many solutions satisfying (1) exist. To choose the correct one, it is necessary to incorporate a prior knowledge into regularization terms and other constraints. If the prior knowledge is not available, the restoration methods frequently converge to solutions which are far from the ground truth even in the noise-free case. Due to its inefficiency, the restoration approach is not considered in this paper.

"Handcrafted" features invariant to blur originally appeared in the 1990's, when some researchers realized that a complete restoration of f is not necessary for object recognition and can be avoided, provided that an appropriate image representation is used. These so called *blur invariants* are then used as an input of a standard classifier. Roughly speaking, blur invariant I is a functional fulfilling the constraint $I(f) = I(f * h)$ for any h from a certain set $\mathcal{S}$ of admissible PSFs. The main drawback of all blur invariants published so far is that they require a parametric model (or other equivalent prior knowledge) of the PSF and lack a unified mathematical framework. Thus, for each class of PSFs, the invariants have to be derived "from scratch" and the type of the PSF must be examined in the applications, which is difficult and time consuming. On the other hand, if the model of the PSF is available, numerous successful applications have been reported in the literature (see Flusser et al. 2016b, Chapter 6 for a survey and further references thereof).

In accordance with the current trends in computer vision, a deep learning approach using convolutional neural networks (CNN) has been tested in several papers. A detailed theoretical analysis of robustness of CNN-based classification with respect to blur has not been conducted yet, so most authors just reported their empirical observations. They all agreed that the blur in query images significantly decreases the CNN performance if the network has been trained on clear images only and that data augmentation (i.e. the extension of the training set with artificially blurred samples, which is a kind of "brute force") helps to mitigate the performance drop. However, the augmentation should be relatively large-scaled, going across various blur types and size, because CNNs are not able to learn one type/size of blur and recognize another one.

1.2 The Manuscript's Contribution

The original contribution of this paper is twofold.

- *Unified theory of blur invariants.* We discover a unified theoretical background of blur invariants which is presented in this paper for the first time. We show that all previously published blur invariants are particular cases of a general theory, which provides this topic with a roof. Three key theorems, referred here as Theorems 3, 4, and 6 are formulated and proved here *regardless of the particular PSF type*. We proved that the key role is played by the orthogonality of the projection operator, which implies blur-invariant properties. This is a significant theoretical contribution of this paper, which has an immediate practical consequence. If we want to derive blur invariants w.r.t. a new class of PSFs, Theorems 3 and 6 offer the solution directly.
- *Comparison of blur invariants and CNN.* We compare the proposed blur invariants to a deep learning and CNN-based solutions and demonstrate the pros and cons of both. We explain when and why each of them fails/wins. Last but not least, we envisage a "fusion" of both approaches to reach an optimal recognition performance and time complexity.

2 Related Work

2.1 Blur Invariants

Unlike geometric invariants, which can be traced over two centuries back to Hilbert (1993), blur invariants are a relatively new topic. The problem formulation and the basic idea appeared originally in the 1990's in the series of papers by Flusser et al. (1995, 1996a) and Flusser and Suk (1998). The invariants presented in these pioneer papers were found heuristically without any theoretical background. The authors observed that certain moments of a symmetric PSF vanish. They derived the relation between the moments of the blurred image and the original and thanks to the vanishing moments of the PSF they eliminated the non-zero PSF moments by a recursive subtraction and multiplication. They did it for axially symmetric (Flusser et al., 1995, 1996a) and centrosymmetric (Flusser & Suk, 1998) PSFs. These invariants, despite their heuristic derivation and the restriction to centrosymmetric PSFs, have been adopted by many researchers in further theoretical studies (Zhang et al., 2010; Wee & Paramesran, 2007; Dai et al., 2010a, b, 2013; Liu et al., 2011a; Zuo et al., 2010; Kautsky & Flusser, 2011; Zhang

Fig. 1 Three approaches to the recognition of blurred images. Image restoration via deconvolution (left), description and recognition by blur invariants (middle), and a deep learning approach (right)
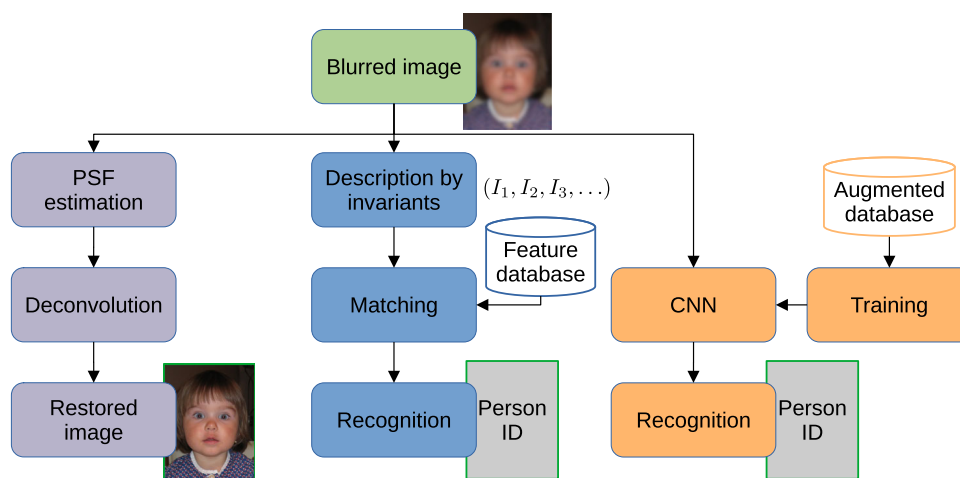
et al., 2000a, 2002; Flusser et al., 2000, 2003; Candocia, 2004; Ojansivu & Heikkilä, 2007; 2008b; 2008a; Makaremi & Ahmadi, 2010, 2012; Galigekere & Swamy, 2006) and in numerous application-oriented papers (Bentoutou et al., 2005; Liu et al., 2011c; Hu et al., 2007; Bentoutou et al., 2002; Bentoutou & Taleb, 2005b, a; Mahdian & Saic, 2007; Ahonen et al., 2008; Zhang et al., 2000b; Yap & Raveendran, 2004).

By a similar heuristic approach, various invariants to circularly symmetric blur (Chen et al., 2011; Dai et al., 2013; Flusser & Zitová, 2004; Ji & Zhu, 2009; Liu et al., 2011b, a; Zhu et al., 2010), linear motion blur (Flusser et al., 1996b; Guang-Sheng & Peng, 2012; Guan et al., 2005; Peng & Jun, 2011; Stern et al., 2002; Wang et al., 2007; Zhong et al., 2013), and Gaussian blur (Liu & Zhang, 2005; Xiao et al., 2012; Zhang et al., 2013) have been proposed.

Significant progress of the theory of blur invariants was made by Flusser et al. (2015), where the invariants to arbitrary N-fold rotation symmetric blur were proposed. In that paper, a derivation based on a mathematical theory rather than on heuristics was presented for the first time. The invariants were constructed by means of projection of the blurred image onto the subspace of the PSFs. A similar result achieved by another technique was later published by Pedone et al. (2013).

Some authors came up with the (computationally very expensive) concept of *blur-invariant metric*, which is similar to blur invariants but allows a blur-robust comparison of two images without constructing the invariants explicitly (Gopalan et al., 2012; Lébl et al., 2019, 2021; Vageeswaran et al., 2013; Zhang et al., 2013).

The main limitation of all the above mentioned methods is their restriction to a given single class of blurs. In other words, the authors first defined the blur type they were considering and then they derived the invariants based on the specific properties of the blur. In this paper, we approach the problem the other way round. Regardless of the particular blur, we find a general formula for blur invariants. Then, for any type of

admissible PSF, this general formula immediately provides specific invariants. This is the major original contribution of this paper that differentiates the proposed theory from the previous ones.

2.2 CNN-Based Recognition of Blurred Images

Traditional CNNs use pixel-wise representation of images, which is significantly changed if the image has been blurred. So, a fast drop of performance if the network, trained on clear images only, is fed with blurred test images, has been reported in several papers. To achieve a reasonable recognition rate, the training set has to be extended using a representative set of degradations of the training images. This process is called *data augmentation*. Common data augmentation methods for image recognition have been designed manually and the best augmentation strategies are dataset-specific (Krizhevsky et al., 2017). To learn the best augmentation strategy was recently proposed in Cubuk et al. (2019), which searches the space of various augmentation operations and finds a policy yielding the highest validation accuracy. Large-scale data augmentation is, however, time and memory consuming and for big databases it might be not feasible even on current super-computers.

Only few papers have studied quantitatively the impact of blur on the network recognition performance. Vasiljevic et al. (2016) first confirmed experimentally that introduction of even a moderate blur hurts the performance of networks trained on clear images and proposed a fine-tuning of the network to overcome this. A similar study was published by Zhou et al. (2017). Dodge and Karam (2016) compared the impact of four types of image degradations—Gaussian blur, contrast deficiency, JPEG artifacts and noise on the CNN-based recognition. They showed that blur is the most significant factor. Pei et al. (2021) studied the influence of various degradations (local occlusion, lens distortion, aliasing, haze, noise, blur) in detail. Their paper again confirmed

that the CNN performance in recognition of blurred images is very low if the training set contains clear images only. Most recently, Lébl et al. (2023) studied the influence of data augmentation with respect to various blur types and sizes and the relationship between the number of augmented samples and the recognition rate. They concluded that even for small training set a relatively large augmentation is necessary to keep the success rate sufficient but the number of augmentation samples can be slightly reduced by choosing a good augmentation strategy.

3 Blur Invariants

In this Section, we show how blur invariants can be constructed by means of suitable projectors. Let us start with introducing some basic terms.

Definition 1 By an *image function* (or *image*) we understand any real function $f \in L_1\left(\mathbb{R}^d\right) \cap L_2\left(\mathbb{R}^d\right)$ with a compact support.[1] The set of all image functions is denoted as $\mathcal{I}$.

On the one hand, this definition of $\mathcal{I}$ is not very restrictive because the image space is "large" enough (more precisely, $\mathcal{I}$ is dense both in L_1 and L_2). On the other hand, it ensures the mathematical correctness of some basic operations. In particular, for any $f, g \in \mathcal{I}$ their convolution exists and $f * g \in \mathcal{I}$ (this follows from the Young's inequality) and their Fourier transforms $\mathcal{F}(f), \mathcal{F}(g)$ exist as well (but may lie outside $\mathcal{I}$). For the convenience, we assume the Dirac δ-function to be an element of $\mathcal{I}$.[2]

Definition 2 A linear operator $P : \mathcal{I} \to \mathcal{I}$ is called a *projection operator* (or *projector* for short) if it is idempotent, i.e. $P^2 = P$.

Let us denote $\mathcal{S} = P(\mathcal{I})$. Then $\mathcal{S}$ is a linear subspace of $\mathcal{I}$ and $\mathcal{I}$ can be expressed as a direct sum $\mathcal{I} = \mathcal{S} \oplus \mathcal{A}$, where $\mathcal{A}$ is called the complement of $\mathcal{S}$. Any $f \in \mathcal{I}$ can be unambiguously written as a sum $f = Pf + f_A$, where Pf is a projection of f onto $\mathcal{S}$ and $f_A \in \mathcal{A}$ is simply defined as $f_A = f - Pf$. For any $f \in \mathcal{I}$ we have $Pf_A = 0$ and, consequently, $\mathcal{S} \cap \mathcal{A} = \{0\}$. If $f \in \mathcal{S}$ then $f = Pf$ and vice versa.

Projector P is called *orthogonal* (OG), if the respective subspaces $\mathcal{S}$ and $\mathcal{A}$ are mutually orthogonal.

Let $\mathcal{S}$ be, from now on, the set of blurring functions (PSFs), with respect to which we want to design the invariants. Any meaningful $\mathcal{S}$ must contain at least one non-zero function

Fig. 2 Partitioning of the image space. Projection operator P decomposes image f into its projection Pf onto $\mathcal{S}$ and its complement f_A, which is a projection onto $\mathcal{A}$. The ellipsoids depict blur-equivalent classes. Blur-invariant information is contained in the primordial images f_r and g_r (see the text)

and must be closed under convolution. In other words, for any $h_1, h_2 \in \mathcal{S}$ must be $h_1 * h_2 \in \mathcal{S}$. This is the basic assumption without which the question of invariance does not make sense. If $\mathcal{S}$ was not closed under convolution, then any potential invariant would be in fact invariant w.r.t. convolution with functions from the "convolution closure" of $\mathcal{S}$, which is the smallest superset of $\mathcal{S}$ closed to convolution.

Under the closure assumption, $(\mathcal{S}, *)$ forms a commutative *semi-group* (it is not a group because the existence of inverse elements is not guaranteed). Hence, the convolution may be understood as a *semi-group action* of $\mathcal{S}$ on $\mathcal{I}$. The convolution defines the following equivalence relation on $\mathcal{I}$: $f \sim g$ if and only if there exist $h_1, h_2 \in \mathcal{S}$ such that $h_1 * f = h_2 * g$. Thanks to the closure property of $\mathcal{S}$ and to the commutativity of convolution, this relation is transitive, while symmetry and reflexivity are obvious. This relation factorizes $\mathcal{I}$ into classes of *blur-equivalent images*. In particular, all elements of $\mathcal{S}$ are blur-equivalent. The image space partitioning and the action of the projector are visualized in Fig. 2.

Now we are ready to formulate the following *General theorem of blur invariants* (GTBI), which performs the main contribution of the paper and a significant difference from all previous work on this field.

Theorem 3 (GTBI) *Let $\mathcal{S}$ be a linear subspace of $\mathcal{I}$, which is closed under convolution and correlation. Let P be an orthogonal projector of $\mathcal{I}$ onto $\mathcal{S}$. Then*

$$I(f)(\mathbf{u}) \equiv \frac{\mathcal{F}(f)(\mathbf{u})}{\mathcal{F}(Pf)(\mathbf{u})} \tag{2}$$

is an invariant w.r.t. a convolution with arbitrary $h \in \mathcal{S}$ at all frequencies $\mathbf{u}$ where $I(f)$ is well defined.

---

[1] Symbol $L_p\left(\mathbb{R}^d\right)$ denotes a space of all functions of d real variables such that $\int |f|^p < \infty$.

[2] From mathematical point of view, this is formally incorrect since $\delta \notin L_1 \cap L_2$. We could correctly include δ by means of theory of distributions but this would be superfluous for the purpose of this paper.
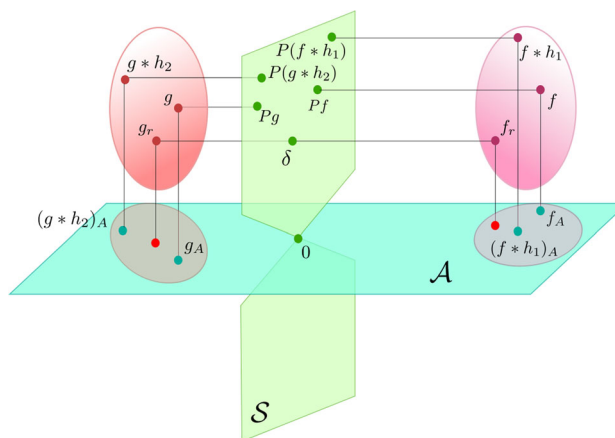
Proof Let us assume that P is "distributive" over a convolution with functions from $\mathcal{S}$, which means $P(f * h) = Pf * h$ for arbitrary f and any $h \in \mathcal{S}$. Then the proof is trivial, we just employ the basic properties of Fourier transform:

$$
\begin{aligned}
I(f * h) &\equiv \frac{\mathcal{F}(f * h)}{\mathcal{F}(P(f * h))} = \frac{\mathcal{F}(f) \cdot \mathcal{F}(h)}{\mathcal{F}(Pf * h)} \\
&= \frac{\mathcal{F}(f) \cdot \mathcal{F}(h)}{\mathcal{F}(Pf) \cdot \mathcal{F}(h)} = I(f).
\end{aligned}
\tag{3}
$$

The "distributive property" of P is equivalent to the constraint that the complement $\mathcal{A}$ is closed w.r.t. convolution with functions from $\mathcal{S}$. This follows from

$$
\begin{aligned}
P(f * h) &= P((Pf + f_A) * h) = P(Pf * h + f_A * h) \\
&= Pf * h + P(f_A * h).
\end{aligned}
\tag{4}
$$

Now let us show that this constraint is implied by the orthogonality of P regardless of its particular form.

Since $\mathcal{S} \perp \mathcal{A}$ and Fourier transform on $L_1 \cap L_2$ preserves the scalar product (this property is known as Plancherel theorem), then $\mathcal{F}(\mathcal{S}) \perp \mathcal{F}(\mathcal{A})$. Let us consider arbitrary functions $a \in \mathcal{A}$ and $h_1, h_2 \in \mathcal{S}$. Using the Plancherel theorem, the convolution theorem, and the correlation theorem (the correlation $\circledast$ of two functions is just a convolution with a flipped function), we have

$$
\begin{aligned}
\langle a * h_1, h_2 \rangle &= \langle \mathcal{F}(a * h_1), \mathcal{F}(h_2) \rangle = \int A \cdot H_1 \cdot H_2^* \\
&= \langle A, H_1^* \cdot H_2 \rangle = \langle a, h_2 \circledast h_1 \rangle = 0.
\end{aligned}
\tag{5}
$$

The last equality follows from the closure of $\mathcal{S}$ w.r.t. correlation. Hence, $\mathcal{A}$ has been proven to be closed w.r.t. convolution with functions from $\mathcal{S}$, which completes the entire proof. □

The invariant $I(f)$ is not defined if $Pf = 0$, which means this Theorem cannot be applied if $f \in \mathcal{A}$. In all other cases, $I(f)$ is well defined almost everywhere. Since $\mathcal{S}$ contains compactly-supported functions only, $\mathcal{F}(Pf)(\mathbf{u})$ cannot vanish on any open set and therefore the set of frequencies, where $I(f)$ is not defined, has a zero measure.[3] In addition to the blur invariance, I is also invariant w.r.t. correlation with functions from $\mathcal{S}$. This is a "side-product" of the assumptions imposed on $\mathcal{S}$. The proof of that is the same as before, with only the operations convolution and correlation swapped.

Under the assumptions of the GTBI, $\mathcal{S}$ itself is always an equivalence class and $\delta \in \mathcal{S}$ (to see this, note that for arbitrary $h \in \mathcal{S}$ we have $h = \delta * h = P(\delta * h) = P\delta * h$ which leads to $P\delta = \delta$).

The GTBI is a very strong theorem because it constructs the blur invariants in a unified form regardless of the particular class of the blurring PSFs and regardless of the image dimension d. The only thing we have to do in a particular situation is to find, for a given subspace $\mathcal{S}$ of the admissible PSFs, an orthogonal projector P. This is mostly much easier job than to construct blur invariants "from scratch" for any $\mathcal{S}$. This is the most important distinction from our previous paper (Flusser et al., 2015), where the invariants were constructed specifically for N-fold symmetric blur without a possibility of generalization.

Before we proceed further, let us show that the assumptions laid on $\mathcal{S}$ and P cannot be skipped.

As a counterexample, let us consider a 1D case where $\mathcal{S}$ is a set of even functions. Let P be defined such that $Pf(x) = f(|x|)$. So, P is a kind of "mirroring" of f and actually, it is a linear (but not orthogonal) projector on $\mathcal{S}$. In this case, $\mathcal{A}$ is a set of functions that vanish for $x \geq 0$. Clearly, $\mathcal{A}$ is not closed to convolution with even functions and functional I defined in GTBI is not an invariant.

Let us consider another example, again in 1D. Let $\mathcal{S}$ be a set of functions that vanish for any $x < 0$. $\mathcal{S}$ is a linear subspace closed to convolution but it is not closed to correlation. Let us define operator P as follows: $Pf(x) = f(x)$ if $x \geq 0$ and $Pf(x) = 0$ if $x < 0$. Obviously, P is a linear orthogonal projector onto $\mathcal{S}$. However, $\mathcal{A}$ is again not closed to convolution with functions from $\mathcal{S}$ and GTBI does not hold. These two simple examples show, that the assumptions of convolution and correlation closure of $\mathcal{S}$ and orthogonality of P cannot be generally relaxed (although GTBI may stay valid in some cases even if these assumptions are violated, see Sect. 5).

The property of blur invariance does not say anything about the ability of the invariant to distinguish two different images. In an ideal case, the invariant should be able to distinguish *any* two images belonging to distinct blur-equivalence classes (images sharing the same equivalence class of course cannot be distinguished due to the invariance). Such invariants are called *complete*. The following *completeness theorem* shows that I is a complete invariant within its definition area.

Theorem 4 (Completeness theorem) *Let I be the invariant defined by GTBI and let $f, g \in \mathcal{I} \backslash \mathcal{A}$. Then $I(f) = I(g)$ almost everywhere if and only if $f \sim g$.*

Proof The proof of the backward implication follows immediately from the blur invariance of I. To prove the forward implication, we set $h_1 = Pg$ and $h_2 = Pf$. Then it holds $f * h_1 = g * h_2$, which means $f \sim g$ due to the definition of the equivalence class. □

To summarize, I cannot distinguish functions belonging to the same equivalence class due to the invariance and functions from $\mathcal{A}$ since they do not lie in its definition area. All

---

[3] This may not be true if $\mathcal{I}$ and $\mathcal{S}$ contained functions of unlimited support. Then $\mathcal{F}(Pf)(\mathbf{u})$ might vanish on a nonzero-measure set, which would decrease the discrimination power of I.

other functions are fully distinguishable. Note that the completeness may be violated on other image spaces, for instance, on a space of functions with unlimited support where we find such f and g that $I(f) = I(g)$ at all frequencies where both $I(f)$ and $I(g)$ are well defined but f and g belong to different equivalence classes.

Understanding what properties of f are reflected by $I(f)$ is important both for theoretical considerations as well as for practical application of the invariant. $I(f)$ is a ratio of two Fourier transforms. As such, it may be interpreted as deconvolution of f with the kernel Pf. This "deconvolution" eliminates the part of f belonging to $\mathcal{S}$ (more precisely, it transfers Pf to δ-function) and effectively acts on the f_A only:

$$I(f) = \frac{\mathcal{F}(f)}{\mathcal{F}(Pf)} = \frac{\mathcal{F}(Pf) + \mathcal{F}(f_A)}{\mathcal{F}(Pf)} = 1 + \frac{\mathcal{F}(f_A)}{\mathcal{F}(Pf)}.$$

$I(f)$ can be viewed as a Fourier transform of so-called *primordial image* f_r. Even if the primordial image itself may not exist (the existence of $\mathcal{F}^{-1}(I(f))$ is not guaranteed in $\mathcal{I}$), it is a useful concept that helps to understand how the blur invariants work. The primordial image is unique for each equivalence class, it is the "most deconvolved" representative of the class. Two images f and g share the same equivalence class if and only if $f_r = g_r$. For instance, the primordial image of all elements of $\mathcal{S}$ is δ-function.

Any element of the equivalence class can be reached from the primordial image through a convolution. Any features, which describe the primordial image, are unique blur-invariant descriptors of the entire equivalence class. At the same time, the primordial image can also be viewed as a kind of normalization. It plays the role of a canonical form of f, obtained as the result of the "maximally possible" deconvolution of f (see Fig. 3 for schematic illustration).

As the last topic in this section, we briefly analyze the robustness of $I(f)$ to noise. Let us assume an additive zero-mean white noise, so we have $g = f * h + n$ and, consequently, $Pg = h * Pf + Pn$. As we will see in Sect. 5, all meaningful projection operators contain summation/integration over a certain set (often large) of pixels, which makes Pn to converge to the mean value of n, which is zero. So, we have

$$I(g) \equiv \frac{\mathcal{F}(g)}{\mathcal{F}(Pg)} = \frac{\mathcal{F}(f) \cdot \mathcal{F}(h) + \mathcal{F}(n)}{\mathcal{F}(Pf) \cdot \mathcal{F}(h)}$$
$$= I(f) + \frac{\mathcal{F}(n)}{\mathcal{F}(Pf) \cdot \mathcal{F}(h)}. \quad (6)$$

Considering the magnitude of the second term, note that $|\mathcal{F}(n)(\mathbf{u})| = \sigma$ because the noise is white. Hence, at least at low frequencies where $\mathcal{F}(Pf) \cdot \mathcal{F}(h)$ dominates, this term is close to zero and I exhibits a robust behavior as $I(g) \doteq I(f)$.
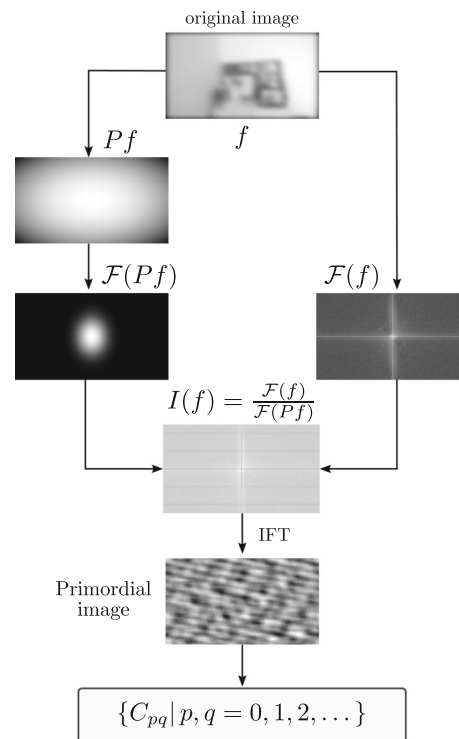

Fig. 3 The concept of the primordial image: The blurred image is projected onto $\mathcal{S}$ and this projection is used to "deconvolve" the input image in the Fourier domain. Blur-invariant primordial image is obtained as a seeming Fourier inversion of $I(f)$. Its moments are blur invariant and can be calculated directly from f

However, this may be violated at high frequencies where $\mathcal{F}(Pf) \cdot \mathcal{F}(h)$ is often low.

4 Invariants and Moments

The blur invariants defined in the frequency domain by GTBI may suffer from several drawbacks when we use them in practical object recognition tasks. Since $I(f)$ is a ratio, we possibly divide by very small numbers which requires careful numerical treatment. Moreover, if the input image is noisy, the high-frequency components of $I(f)$ may be significantly corrupted. This can be overcome by suppressing them by a low-pass filter, but this procedure introduces a user-defined parameter (the cut-off frequency) which should be set up with respect to the particular noise level. That is why we prefer to work directly in the image domain. Some heuristically discovered image-domain blur invariants were already published in the early papers (Flusser et al., 1995, 1996a; Flusser & Suk, 1998). Here we present a general theory, which originates from the GTBI.

A straightforward solution might be to calculate an inverse Fourier transform of $I(f)$, which leads to obtaining the primordial image f_r and to characterize f_r by some popular

descriptors such as moments. This would, however, be time-consuming and also problematic from the numerical point of view. We would not only have to calculate the projection Pf, two forward and one inverse Fourier transforms, but even worse, the result may not lie in $\mathcal{I}$. In this Section, we show how to substantially shorten and simplify this process. We show, that the moments of the primordial image can be calculated directly from the input blurred image, without an explicit construction of Pf and $I(f)$. Since f_r is a blur invariant, each its moment must be a blur invariant, too. This direct construction of blur invariants in the image domain, again without specifying particular $\mathcal{S}$ and P, is the major theoretical result of the paper and performs a very useful tool for practical image recognition.

Image moments can be understood as "projections" of the image onto an arbitrary polynomial basis, which is formalized by the following definition.

Definition 5 Let $\mathcal{B} = \{\pi_\mathbf{k}(\mathbf{x})\}$ be a set of d-variate polynomials. Then the integral

$$M_\mathbf{p}^{(f)} = \int \pi_\mathbf{p}(\mathbf{x}) f(\mathbf{x}) \, d\mathbf{x} \tag{7}$$

is called *moment* of function f with respect to set $\mathcal{B}$. The non-negative integer $|\mathbf{p}|$, where $\mathbf{p}$ is a d-dimensional multi-index, is the *order* of the moment.

Clearly, for any $f \in \mathcal{I}$ its moments w.r.t. arbitrary $\mathcal{B}$ exist and are finite. Moments are widely used image descriptors. Depending on $\mathcal{B}$, we recognize various types of moments. If $\pi_\mathbf{k}(\mathbf{x}) = \mathbf{x}^\mathbf{k}$ we speak about *geometric* moments. If $d = 2$ and $\pi_{pq}(x, y) = (x + iy)^p (x - iy)^q$, we obtain *complex* moments. If the polynomials $\pi_\mathbf{k}(\mathbf{x})$ are orthogonal, we get *orthogonal* (OG) moments. Legendre, Zernike, Chebyshev, and Fourier–Mellin moments are the most common examples. For the theory of moments and their application in image analysis we refer to Flusser et al. (2016b).

Various bases have been employed to construct moment invariants (Flusser et al., 2009). There is no significant difference among them since between any two polynomial bases there exists a transition matrix. In other words, from the theoretical point of view, all polynomial bases and all respective moments carry the same information, provide the same recognition power and generate equivalent invariants. However, working with some basis might be in a particular situation easier than with the others, and also numerical properties and stability of the moments may differ from each other. Here we choose to work with a basis that *separates* the moments of Pf and f_A, although equivalent invariants could be derived in any basis at the expense of the complexity of respective formulas.

Let $\mathcal{S}$ and P fulfill the assumptions of GTBI. Considering the decomposition $f = Pf + f_A$, we have for the moments

$$M_\mathbf{p}^{(f)} = M_\mathbf{p}^{(Pf)} + M_\mathbf{p}^{(f_A)}. \tag{8}$$

We say that $\mathcal{B}$ separates the moments if there exist a non-empty set of multi-indices D such that it holds for any $f \in \mathcal{I}$

$$M_\mathbf{p}^{(Pf)} = M_\mathbf{p}^{(f)} \tag{9}$$

if $\mathbf{p} \in D$ and

$$M_\mathbf{p}^{(Pf)} = 0 \tag{10}$$

if $\mathbf{p} \notin D$. In other words, this condition says that the moments are either preserved or vanish under the action of P. If fulfilled, the condition also says that the value of $M_\mathbf{p}^{(f_A)}$ is complementary to $M_\mathbf{p}^{(Pf)}$.

A sufficient condition for $\mathcal{B}$ to separate the moments is that $\pi_\mathbf{p} \in \mathcal{S}$ if $\mathbf{p} \in D$ and $\pi_\mathbf{p} \in \mathcal{A}$ otherwise. Since $\mathcal{S}$ and $\mathcal{A}$ are assumed to be mutually orthogonal, the separability of such $\mathcal{B}$ is obvious. This has nothing to do with a (non)orthogonality of $\mathcal{B}$ itself, as we show in the following simple 1D example. Let $\mathcal{S}$ be a set of even functions and $\mathcal{A}$ be a set of odd functions. Let $\pi_p(x) = x^p$. If we take $D = \{p = 2k | k \geq 0\}$, we obtain the moment-separating polynomials.

For the given $\mathcal{S}$ and projector P, the existence of a basis that separates the moments is not guaranteed, although in most cases of practical interest we can find some. If it does not exist, the moment blur invariants still can be derived. It is sufficient if the moments $M_\mathbf{p}^{(Pf)}$ can be expressed in terms of $M_\mathbf{p}^{(f)}$ if $\mathbf{p} \in D$ and some functions of $M_\mathbf{p}^{(f)}$ equal zero for $\mathbf{p} \notin D$. This makes the derivation more laborious and the formulas more complicated but does not make a principle difference. Anyway, to keep things simple, we try for any particular $\mathcal{S}$ to find such $\mathcal{B}$ that provides the moment separability.

To get the link between $I(f)$ and the moments $M_\mathbf{p}^{(f)}$, we recall that Taylor expansion of Fourier transform is

$$\mathcal{F}(f)(\mathbf{u}) = \sum_\mathbf{p} \frac{(-2\pi i)^{|\mathbf{p}|}}{\mathbf{p}!} m_\mathbf{p}^{(f)} \mathbf{u}^\mathbf{p} \tag{11}$$

where $m_\mathbf{p}$ is a geometric moment. In the sequel, we assume that the power basis $\pi_\mathbf{p}(\mathbf{x}) = \mathbf{x}^\mathbf{p}$ separates the moments. If it was not the case, one would substitute into (11) any separating basis through the polynomial transition relation.

The GTBI can be rewritten as

$$\mathcal{F}(Pf)(\mathbf{u}) \cdot I(f)(\mathbf{u}) = \mathcal{F}(f)(\mathbf{u}). \tag{12}$$

All these three Fourier transforms can be expanded similarly to (11) into absolutely convergent Taylor series. Thanks to

the moment separability, we can for any $\mathbf{p} \in D$ simply write $m_{\mathbf{p}}^{(Pf)} = m_{\mathbf{p}}^{(f)} = m_{\mathbf{p}}$. So, we have

$$\sum_{\mathbf{p} \in D} \frac{(-2\pi i)^{|\mathbf{p}|}}{\mathbf{p}!} m_{\mathbf{p}}^{(f)} \mathbf{u}^{\mathbf{p}} \cdot \sum_{\mathbf{p}} \frac{(-2\pi i)^{|\mathbf{p}|}}{\mathbf{p}!} C_{\mathbf{p}} \mathbf{u}^{\mathbf{p}}$$
$$= \sum_{\mathbf{p}} \frac{(-2\pi i)^{|\mathbf{p}|}}{\mathbf{p}!} m_{\mathbf{p}}^{(f)} \mathbf{u}^{\mathbf{p}}, \tag{13}$$

where $C_{\mathbf{p}}$ can be understood as the moments of the primordial image f_r. Comparing the coefficients of the same powers of $\mathbf{u}$ we obtain, for any $\mathbf{p}$

$$\sum_{\mathbf{k} \in D}^{\mathbf{p}} \frac{(-2\pi i)^{|\mathbf{k}|}}{\mathbf{k}!} \frac{(-2\pi i)^{|\mathbf{p}-\mathbf{k}|}}{(\mathbf{p}-\mathbf{k})!} m_{\mathbf{k}} C_{\mathbf{p}-\mathbf{k}} = \frac{(-2\pi i)^{|\mathbf{p}|}}{\mathbf{p}!} m_{\mathbf{p}}, \tag{14}$$

which can be read as

$$\sum_{\mathbf{k} \in D}^{\mathbf{p}} \binom{\mathbf{p}}{\mathbf{k}} m_{\mathbf{k}} C_{\mathbf{p}-\mathbf{k}} = m_{\mathbf{p}}. \tag{15}$$

The summation goes over those $\mathbf{k} \in D$ for which $0 \le k_i \le p_i$, $i = 1, \ldots, d$. Note that always $\mathbf{0} \in D$. (To see that, it is sufficient to find an image whose zero-order moment is preserved under the projection. Such an example is δ-function, because $P(\delta) = \delta$, as we already showed.)

After isolating $C_{\mathbf{p}}$ on the left-hand side we obtain the final recurrence

$$m_{\mathbf{0}} C_{\mathbf{p}} = m_{\mathbf{p}} - \sum_{\substack{\mathbf{k} \in D \\ \mathbf{k} \neq \mathbf{0}}}^{\mathbf{p}} \binom{\mathbf{p}}{\mathbf{k}} m_{\mathbf{k}} C_{\mathbf{p}-\mathbf{k}}. \tag{16}$$

This recurrent formula is a general definition of blur invariants in the image domain (provided that $m_{\mathbf{0}} \neq 0$).[4] Since $I(f)$ has been proven to be invariant to blur belonging to $\mathcal{S}$, all coefficients $C_{\mathbf{p}}$ must also be blur invariants. The beauty of Eq. (16) lies in the fact that we can calculate the invariants from the moments of f, without constructing the primordial image explicitly either in frequency or in the spatial domain.

Some of the invariants $C_{\mathbf{p}}$ are trivial for any f and useless for recognition. We always have $C_{\mathbf{0}} = 1$ and some other invariants may be constrained depending on the index set D. If for arbitrary $\mathbf{p}, \mathbf{k} \in D$ also $(\mathbf{p} - \mathbf{k}) \in D$, then $C_{\mathbf{p}} = 0$ for any $\mathbf{p} \in D$ as can be deduced from Eq. (16) via induction. This commonly happens in many particular cases of practical interest and then only the invariants with $\mathbf{p} \notin D$

---

[4] If $m_{\mathbf{0}} = 0$, then $C_{\mathbf{p}}$ is not defined. We find the first non-zero moment $m_{\mathbf{n}}$, $\mathbf{n} \in D$, and derive an analogous recurrence for $C_{\mathbf{p}-\mathbf{n}}$.

should be used. In addition to that, some invariants may vanish depending on f. In particular, if $f \in \mathcal{S}$, then $C_{\mathbf{p}} = 0$ for *any* $\mathbf{p} \neq \mathbf{0}$.

5 Blur Examples

In this section, we show the blur invariants provided by the GTBI for several concrete choices of $\mathcal{S}$ and P with a particular focus on those of practical importance in image recognition. Some of them are equivalent to the invariants already published in earlier papers; in such cases, we show the link between them. Some other invariants are published here for the first time.

5.1 Trivial Cases

The formally simplest case ever is $\mathcal{S} = \mathcal{I}$ and $Pf = f$. Although this choice fulfills the assumptions of GTBI, it is not of practical importance because the entire image space forms a single equivalence class, and any two images are blur equivalent. Actually, GTBI yields $I(f) = 1$ for any f.

An opposite extreme is to choose $\mathcal{S} = \{a\delta | a \in \mathbb{R}\}$. This "blur" is in fact only a contrast stretching. If we set $Pf = (\int f) \cdot \delta$, P is not orthogonal but still $P(f * h) = Pf * h$ and GTBI can be applied provided that $\int f \neq 0$. We obtain $I(f) = \mathcal{F}(f)/\int f$, which leads to a contrast-normalized primordial image $f_r = f/\int f$.

Another rather trivial case is $\mathcal{S} = \{h | \int h = 1\}$. This is the set of all brightness-preserving blurs without any additional constraints. We may construct $Pf = f/\int f$, which actually is a projector; however it is neither linear nor orthogonal. Since $P(f * h) = Pf * h$, we can still apply GTBI, which yields a single-valued blur invariant $I(f) = \int f$, that corresponds to the primordial image $f_r = (\int f) \cdot \delta$.

5.2 Symmetric Blur in 1D

In 1D, the only blur space $\mathcal{S}$, which can be defined generically and is of practical interest, is the space of all even functions. 1D symmetric blur invariants were firstly described in Flusser and Suk (1997) and later adapted to wavelet domain by Makaremi and Ahmadi (2010). Kautsky and Flusser (2011) rigorously investigated these invariants and showed how to construct them in terms of arbitrary moments. Galigekere and Swamy (2006) studied the blur invariants of 2D images in the Radon domain, which inherently led to 1D blur invariants.

If we consider the projector

$$Pf(x) = (f(x) + f(-x))/2 \tag{17}$$

then $\mathcal{A}$ is a space of odd functions, P is orthogonal and GTBI can be applied directly. As for the moment expansion, the

simplest solution is to use the standard monomials $\pi_p(x) = x^p$, which separate the geometric moments for D being the set of even non-negative indices.

5.3 Centrosymmetric Blur in 2D

Invariants w.r.t. centrosymmetric blur in 2D have attracted the attention of the majority of authors who have been involved in studying blur invariants. The number of papers on this kind of blur exceeds significantly the number of all other papers on this field. This is basically for two reasons—such kind of blur appears often in practice and the invariants are easy to find heuristically, without the knowledge of the state-of-the-art theory of projection operators.

A natural way of defining P is

$$Pf(x, y) = (f(x, y) + f(-x, -y))/2. \qquad (18)$$

Then standard geometric moments are separated at $D = \{(p, q) | (p + q)\ \text{even}\}$ and Eq. (16) leads to moment expansion that appeared in some earlier papers such as in Flusser and Suk (1998) and others cited in Sect. 2.1.

This approach can be extended into 3D, where the definition of centrosymmetry is analogous. Existing 3D blur invariants (Flusser et al., 2000, 2003) are just special cases of Eq. (16).

5.4 Radially Symmetric Blur

Radially (circularly) symmetric PSF's satisfying $h(r, \phi) = h(r)$ appear in imaging namely as an out-of-focus blur on a circular aperture (see Fig. 4a for an example). The projector P_∞ is defined as

$$(P_\infty f)(r) = \frac{1}{2\pi} \int\limits_0^{2\pi} f(r, \phi)\, d\phi. \qquad (19)$$

The standard power basis does not separate the moments. This is why various radial moments have been used to ensure the separation. Basis $\mathcal{B}$ consists of circular harmonics-like functions of the form $\pi(r, \phi) = R_{pq}(r)e^{i\chi(p,q)\phi}$, where $R_{pq}(r)$ is a radial polynomial and $\chi(p, q)$ is a simple function of the indices. There are several choices of $\mathcal{B}$, which separate the respective moments and yield blur invariants (the index set D depends on the particular $\mathcal{B}$). Some of them were introduced even without the use of projection operators. They mostly employed Zernike moments (Chen et al., 2011; Dai et al., 2013; Ji & Zhu, 2009; Zhu et al., 2010), Fourier–Mellin moments (Liu et al., 2011b) and complex moments (Flusser & Zitová, 2004).
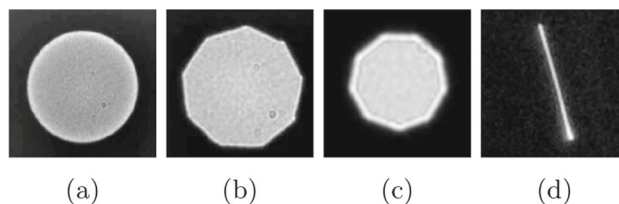

(a)   (b)   (c)   (d)

Fig. 4 Blurring PSFs obtained as photographs of a bright point. Out-of-focus blur on a circular aperture (**a**), on polygonal apertures (**b**, **c**), and directional blur (**d**). Aperture **a** has radial symmetry, **b** has 9-fold rotation symmetry and **c** exhibits 9-fold dihedral symmetry

5.5 N-Fold Symmetric Blur

N-fold rotationally symmetric blur performs one of the most interesting cases, both from theoretical and practical points of view. This kind of blur appears as an out-of-focus blur on a polygonal aperture. Most cameras have an aperture the size of which is controlled by physical diaphragm blades, which leads to polygonal or close-to-polygonal aperture shapes if the diaphragm is not fully open (see Fig. 4b, c).

The blur space is defined as

$$\mathcal{S}_N = \{h | h(r, \theta) = h(r, \theta + 2\pi/N)\}. \qquad (20)$$

$\mathcal{S}_N$ is a vector space closed under convolution and correlation. We can construct projector P_N as

$$(P_N f)(r, \theta) = \frac{1}{N} \sum_{j=1}^N f(r, \theta + \alpha_j), \qquad (21)$$

where $\alpha_j = 2\pi j/N$. Since P_N is an orthogonal projector, GTBI can be immediately applied. Complex moments are separated with

$$D = \{(p, q) | (p - q)/N\ \text{is integer}\} \qquad (22)$$

which allows to get particular blur invariants from Eq. (16).

Invariants to N-fold symmetric blur were originally studied in Flusser et al. (2015), where the idea of projection operators appeared for the first time. Their application for registration of blurred images was reported in Pedone et al. (2013).

5.6 Dihedral Blur

The N-fold symmetry, discussed in the previous subsection, may be coupled with the axial symmetry. In such a case, the number of the axes equals N and we speak about the N-fold dihedral symmetry. Many out-of-focus blur PSFs are actually dihedral, particularly if the diaphragm blades are straight (see Fig. 4c).

The blur space $\mathcal{D}_N$ is a subset of $\mathcal{S}_N$ given as

$$\mathcal{D}_N = \{h \in \mathcal{S}_N \,|\, \exists \alpha \text{ such that } h(x, y) = h^\alpha(x, y)\}, \qquad (23)$$

where $\alpha \in \langle 0, \pi/2 \rangle$ is the angle between the symmetry axis a and the x-axis and $h^\alpha(x, y)$ denotes function $h(x, y)$ flipped over a. However, the set $\mathcal{D}_N$ is *not* closed under convolution if we allow various axis directions. Only if we fix the symmetry axis orientation to a constant angle α, we get the closure property. Then we can define the projection operator Q_N^α as

$$Q_N^\alpha f = P_N(f + f^\alpha)/2 \qquad (24)$$

and GTBI can be applied.

Dihedral blur invariants were firstly studied in Boldyš and Flusser (2013). Their major limitation comes from the fact that the orientation of the symmetry axis must be apriori known (and the same for all images entering the classifier). This is far from being realistic and the only possibility is to estimate α from the blurred image itself (Pedone et al., 2015).

5.7 Directional Blur

Directional blur (sometimes called linear motion blur) is a 2D blur of a 1D nature that acts in a constant direction only. Directional blur may be caused by camera shake, scene vibrations, and camera or scene motion. The velocity of the motion may vary during the acquisition, but this model assumes the motion along the line. We do not consider a general motion blur along an arbitrary curve in this paper.[5]

The respective PSF has the form (for the sake of simplicity, we start with the horizontal direction)

$$h(x, y) = h_1(x)\delta(y), \qquad (25)$$

where $h_1(x)$ is an arbitrary 1D image function. The space $\mathcal{S}$ is defined as a set of all functions of the form (25). When considering a constant direction only, $\mathcal{S}$ is closed under 2D convolution and correlation. The projection operator P is defined as

$$Pf(x, y) = \delta(y) \int f(x, y)\, \mathrm{d}y. \qquad (26)$$

P is not orthogonal but geometric moments are separated with $D = \{(p, q)\,|\, q = 0\}$ and Eq. (16) yields the directional blur invariants in terms of geometric moments.

If the blur direction under a constant angle β is known, the projector $P^\beta f$ is defined analogously to (26) by means a line integral along a line which is perpendicular to the blur

direction (see Fig. 4d for an example of a real directional PSF).

The idea of invariants to linear motion blur appeared for the first time in Flusser et al. (1996b) and in a similar form in Stern et al. (2002), without any connection to the projection operator. Zhong used the motion blur invariants for recognition of reflections on a waved water surface (Zhong et al., 2013). Peng et al. used them for weed recognition from a camera moving quickly above the field (Peng & Jun, 2011) and for classification of wood slices on a moving conveyor belt (Guang-Sheng & Peng, 2012) (these applications were later enhanced by Flusser et al. 2013, 2014). Other applications can be found in Guan et al. (2005) and Wang et al. (2007). The necessity of knowing the blur direction beforehand is, however, an obstacle to the wider usage of these invariants.

5.8 Gaussian Blur

Gaussian blur appears whenever the image has been acquired through a turbulent medium. It is also introduced into the images as the sensor blur due to the finite size of the sampling pulse and may be sometimes applied intentionally as a part of denoising.

Since Gaussian function has an unlimited support, we have to extend our current definition of $\mathcal{I}$ by including functions of exponential decay. We define the set $\mathcal{S}$ as

$$\mathcal{S} = \{aG_\Sigma \,|\, a > 0,\ \Sigma \text{ positive definite}\}, \qquad (27)$$

where Σ is the covariance matrix which controls the shape of the Gaussian G_Σ.

$\mathcal{S}$ is closed under convolution but it is not a vector space. We define Pf to be such element of $\mathcal{S}$ which has the same integral and covariance matrix as the image f itself. Clearly, $P^2 = P$ but P is neither linear nor orthogonal. Although the assumptions of GTBI are violated, the Theorem still holds thanks to $P(f * h) = Pf * h$. The moment expansion analogous to Eq. (16) can be obtained when employing the parametric shape of the blurring function. Thanks to this, we express all moments of order higher than two as functions of the low-order ones, which substantially increases the number of non-trivial invariants.

Several heuristically found Gaussian blur moment invariants appeared in Liu and Zhang (2005), Xiao et al. (2012) and Zhang et al. (2013). Invariants based on projection operators were proposed originally in Flusser et al. (2016a) for circular Gaussians and in Kostková et al. (2020) for blurs with a non-diagonal covariance matrix.

---

[5] Imposing no restrictions on the blur trajectory would lead to a very broad blur space, where only trivial invariants exist.

6 Combined Invariants

An important property of the blur invariants Eq. (16) is that they can be easily made invariant also to arbitrary linear transformation. This is crucial in practice, where the image to be recognized is often degraded by an unknown spatial deformation, which originates from non-perpendicular view and other geometric distortions. CNNs are not able to handle this problem efficiently, they again rely on a massive augmentation of the training set by many samples with various spatial deformations.

Let us assume a general linear transformation

$$\mathbf{x}' = A\mathbf{x},$$

where A is a $d \times d$ regular matrix. The compound transformation is then given as

$$g(\mathbf{x}) = (f * h)(A\mathbf{x}). \tag{28}$$

This can be rewritten as

$$g(\mathbf{x}) = \|A\|(f' * h')(\mathbf{x}) , \tag{29}$$

where $\|A\|$ is the absolute value of the determinant of $A, f'(\mathbf{x}) = f(A\mathbf{x})$, and $h'(\mathbf{x}) = h(A\mathbf{x})$, so the blur and the linear coordinate transform are commutative in this sense.

Let us assume that $\mathcal{S}$ is closed under linear transformation, which means $h' \in \mathcal{S}$ for arbitrary A and any $h \in \mathcal{S}$. Now we formulate the theorem which tells in a constructive manner how to generate combined invariants in arbitrary dimension d, if some geometric invariants w.r.t. linear transformation are available (Fig. 5).

Theorem 6 (Substitution theorem) *Let $f \in \mathcal{I}$, $f'(\mathbf{x}) = f(A\mathbf{x})$, $I(f)$ be the blur invariant constructed according to the GTBI and $C_\mathbf{p}$ be blur moment invariants defined by Eq. (16). Let the blur space $\mathcal{S}$ be closed w.r.t. linear coordinate transformation. Let $J(m_\mathbf{0}, \ldots, m_\mathbf{p})$ be an absolute invariant of image moments w.r.t. A, i.e. $J(m'_\mathbf{0}, \ldots, m'_\mathbf{p}) = J(m_\mathbf{0}, \ldots, m_\mathbf{p})$. Then $J(C_\mathbf{0}, \ldots, C_\mathbf{p})$ is a* relative *invariant w.r.t. both A and the blur as*

$$J(C'_\mathbf{0}, \ldots, C'_\mathbf{p}) = \|A\|^{-w} J(C_\mathbf{0}, \ldots, C_\mathbf{p}),$$

where w is the weight[6] *of invariant $J(m_\mathbf{0}, \ldots, m_\mathbf{p})$.*

See Appendix for the proof of Theorem 6.

---

[6] The term *weight* of an invariant has been commonly used in the theory of algebraic invariants, see for instance (Suk & Flusser, 2011) for the definition. For any given invariant, its weight is known and follows from the way how the invariant has been constructed.
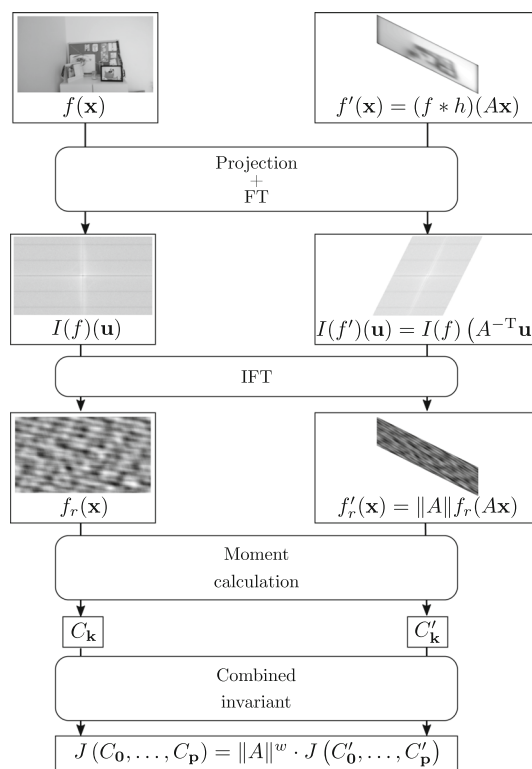

Fig. 5 Visual explanation of the idea of the Substitution Theorem. In practice, invariant $J(C_\mathbf{0}, \ldots, C_\mathbf{p})$ is computed directly from image f without the intermediate steps

Substitution theorem says that combined invariants can be simply obtained by replacing the moments in the formula for geometric invariants by moment-based blur invariants (16). The strength of the Theorem lies in the fact that it holds for any blur set $\mathcal{S}$, general linear transformation, arbitrary image dimension and even for arbitrary polynomial basis of the moments. The Theorem benefits from relatively rich literature on geometric moment invariants, that can be taken as the invariant J. If A is a 2D rotation/scaling, we refer to rotation invariants from complex moments (Abu-Mostafa & Psaltis, 1984; Flusser, 2000). For 3D rotation/scaling, the use of traditional moment invariants (Cyganski & Orr, 1988; Sadjadi & Hall, 1980) or those based on spherical harmonics (Lo & Don, 1989; Suk et al., 2015) is recommended depending of $\mathcal{S}$. In case of general linear/affine transformation, various affine moment invariants are available in 2D (Hickman, 2012; Reiss, 1991; Suk & Flusser, 2011) as well as in 3D (Xu & Li, 2006).

7 Experimental Evaluation

In this section, we show the performance of the proposed invariants in the recognition of blurred facial photographs, in template matching within a blurred scene and in two common image processing problems—multichannel deconvolution

and multifocus fusion—where we use the proposed invariants for registration of blurred frames. The first experiment was performed on simulated data, which makes possible to evaluate the results quantitatively, while the other three experiments show the performance on real images and blurs. A comparison to CNN and another "handcrafted" blur-invariant technique is also presented.

7.1 Face Recognition

In this experiment, we used 38 facial images of distinct persons from the YaleB dataset (Georghiades et al., 2001) (frontal views only). Each class was represented by a single image resized to 256×256 pixels and normalized to brightness. As the test images, we used synthetically blurred and noisy instances of the database images starting from mild (5×5 blur, SNR = 50 dB) to heavy (125×125 blur, SNR = 5 dB) distortions. We used four types of centrosymmetric blur (circular, random, linear motion, Gaussian) and Gaussian white noise in these simulations (see Fig. 6 for some examples). In each setting, we generated 10 instances of each database image.

The faces were classified by four different methods—blur invariants, CNN trained on clear images only, CNN trained on images augmented with blur, and the Gopalan's invariant (Gopalan et al., 2012). As blur invariants, we used the particular version of $I(f)$ from Theorem 6 with operator P defined in Sect. 5.3. As the CNN, we used a pre-trained ResNet18 (He et al., 2016) initially trained on the ImageNet dataset (Deng et al., 2009). Data augmentation was done by adding 100 differently blurred and noisy instances to the training set such that the blur was of the same size as that of the test images. The Gopalan's invariant belongs to "handcrafted" features and measure the "distance" between two images in a way that should be insensitive to blur. Unlike the proposed invariants, the Gopalan's method requires the knowledge of the blur support size, which is no problem in simulated experiments.

The recognition results are summarized in Table 1. The performance of the proposed invariants is excellent except for the last two settings, where the blur caused an extreme smoothing and significant boundary effect (but still the performance over 90% is very good). Figure 7 shows examples of a very heavy blur that was handled correctly by the proposed invariants. The CNN trained only on clear images fails for mid-size and large blurs, which corresponds to the results of earlier studies. However, if we augment the training data extensively with blurred images, the performance is close to 100% but the training time was about four hours compared to few seconds required by the invariants. In this scenario, introducing new images/persons to the database requires additional lengthy training of CNNs. The performance of the Gopalan's method decreases as the blur increases because

Fig. 6 Sample faces used in the experiment. From top to bottom: no blur, motion, Gaussian, uniform and random blur; from left to right: blur size 5, 7, 9, 11, 13, and 15 pixels

Table 1 The recognition rate [%] for different degradations achieved by the proposed invariants (In), CNN trained on clear images, CNN with augmentation by blurred images, and the Gopalan's method (Gopalan et al., 2012) (G)

Degradation		Method			
SNR	Blur	In	CNN	A-CNN	G
50	Circular 5×5	100	100	100	100
50	Circular 10×10	100	90	100	98
50	Circular 15×15	100	35	100	76
50	Circular 125×125	100	–	100	40
5	Circular 125×125	99.9	–	99.8	5
5	Random 125×125	99.8	–	99.9	5
5	Motion 125×125	92	–	99.6	3
5	Gaussian 125×125	91	–	99.5	3

this method is blur-invariant only approximately. Its computing complexity is less than that of the augmented CNN but much higher than that of the proposed invariants and the CNN without augmentation.

7.2 Template Matching

Localization of sharp templates in a blurred scene is a common task in many application areas such as in landmark-based image registration and in stereo matching. In this experiment, we show how the blur invariants can be used for this purpose.

We took two pictures of the same indoor scene—the first one was sharp while the other one was intentionally taken with wrong focus. In the sharp image, we selected 21 square
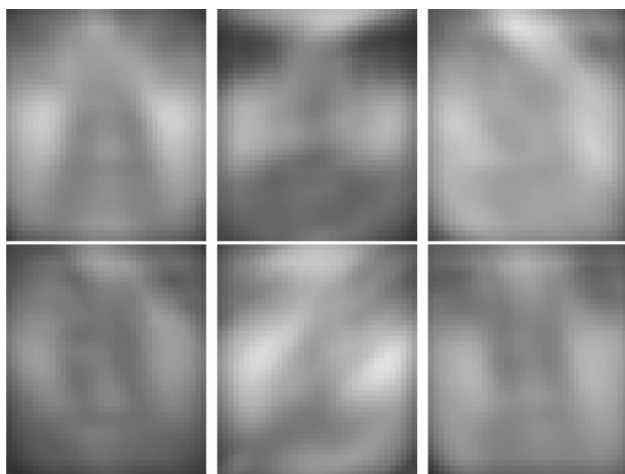
Fig. 7 Extreme cases recognized correctly by the blur invariants but misclassified both by CNN and the Gopalan's method

templates (see Fig. 8a) and the goal was to find these templates in the blurred scene. Since the out-of-focus blur has approximately a circular shape, we used the blur invariants w.r.t. radially symmetric blur (see Sect. 5.4). Since the templates are relatively small, we used the invariants defined directly in the image domain by means of moments (16). The matching was performed by searching over the whole scene, without using any prior information about the template position. The matching criterion was the minimum distance in the space of blur invariants. Nine templates were localized with an error less than or equal to 10 pixels, eight templates with an error 11–20 pixels, three templates with an error 21–30 pixels, and one template with an error greater than 30 pixels (see Fig. 8b). In the sense of a target error, each template was localized in a position that is less than half of the template size from the ground truth.

The localization error is caused by the fact that the blurred template is not exactly a convolution of the ground truth template and the PSF. We observe a strong boundary effect as pixels outside the template influence pixels inside the template. This interaction is, of course, beyond the assumed convolution model. In the case of a large PSF, it influences the matching. If the distance matrix has a flat minimum, then a small disturbance of the invariants due to the boundary effect may result in an inaccurate match.

For comparison, we performed the same task using plain moments instead of the invariants while keeping the number and order of the features the same. Results are unacceptable, most of the templates were matched in totally wrong positions (see Fig. 8c). This clearly shows that introducing blur-invariant features brings a significant improvement.
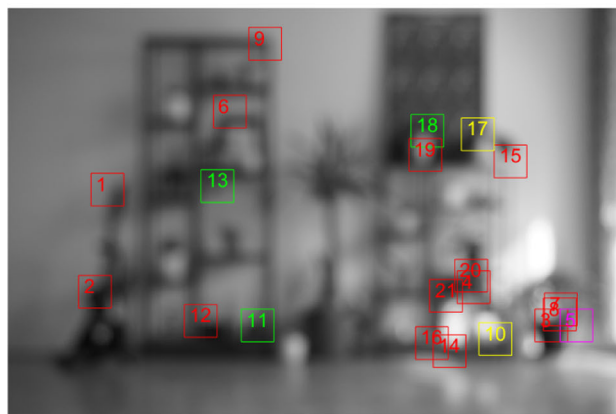
7.3 Multichannel Deconvolution

Multichannel blind deconvolution (MBD) is a process where two or more differently blurred images of the same scene are

Fig. 8 Template matching experiment. The original sharp scene with the selected templates (**a**), the matched templates in the defocused scene using blur invariants (**b**) and the same using plain moments (**c**). The template color encodes the localization error (green: 0–10 pixels, yellow: 11–20, violet: 21–30, red: > 30)

given as an input and a single de-blurred image is obtained as an output (Campisi & Egiazarian, 2007). The restoration is blind, so no parametric form of the PSF's is required. Comparing to single-channel deconvolution, it is more stable

Fig. 9 Multichannel deconvolution. The original frames blurred by a camera shake (left and middle). Note the shift and rotation misalignment between them, that was registered by blur-invariant phase correlation. The result of MBD (Kotera et al., 2017) applied on the registered frames (right)
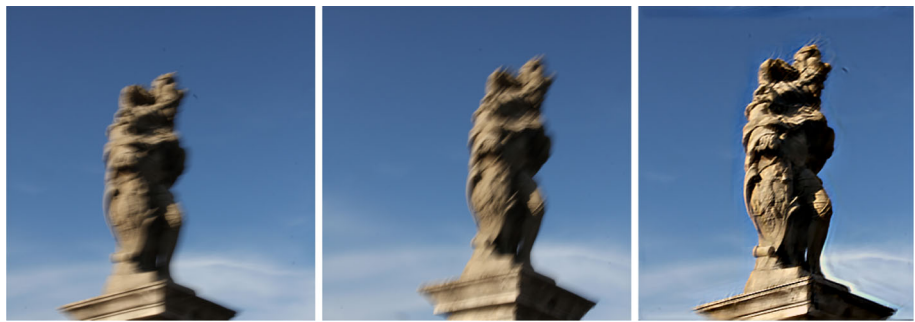

Fig. 10 Multifocus fusion. The input frames focused on the foreground (left) and on the background (middle). The frames were registered by blur-invariant phase correlation and fused by the method from Zhang et al. (2020) (right)

and usually produces much better results. However, the crucial requirement is that the input frames must be registered before entering the deconvolution procedure. The registration accuracy up to several pixels is sufficient because advanced MBD algorithms are able to compensate for a small misalignment (Šroubek & Flusser, 2005). Since the input frames are blurred, most of the common registration techniques designed originally for sharp images (Zitová & Flusser, 2003) fail.

For the registration of blurred frames, the proposed invariants can be used. In Fig. 9 (left and middle), we see two input images of a statue blurred by camera shake. Since the camera was handheld and there was a few-second interval between the acquisitions, the images differ from each other not only by the particular blur but also by a shift and a small rotation. To register them, we use "blur-invariant phase correlation" method. It is an efficient landmark-free technique inspired by traditional phase correlation (de Castro & Morandi, 1987). Our method uses directly the blur invariants $I(f)$ and $I(g)$ (instead of whitened Fourier spectrum $F/|F|$ and $G/|G|$ used in the phase correlation) to find the correlation peak. Since we do not have much prior information about the blurs, we use operator P_2 from Section V.E to design the invariants, because it is less specific than the others and should work for many blurs. Switching between Cartesian and polar domains, the method can register both shift and rotation.

In this real-data example we do not have any ground truth so we cannot explicitly measure the registration accuracy. However, it is documented by a good performance of the subsequent MBD algorithm. The registered frames were used as an input of the MBD proposed in Kotera et al. (2017). The

result can be seen in Fig. 9 right. We acknowledge a sharp image with very little artifacts, which proves a sufficient registration accuracy (and of course a good performance of the MBD algorithm itself).

7.4 Multifocus Fusion

Multifocus image fusion (MIF) is a well-known technique of combining two or more images of the same 3D scene, that were taken by a camera with a shallow depth of field (Blum & Liu, 2006). Typically, one frame is focused to the foreground while the other one to the background (see Fig. 10 for an example). The fusion algorithms basically decide locally in which frame this part of the scene is best focused and generate the fused image by stitching the selected parts together without performing any deconvolution. Obviously, an accurate registration of the inputs is a key requirement.

The registration problem is here even more challenging than in the previous experiment, because the convolution model holds only on the foreground or background and the required accuracy is higher them in the MBD case.

The input frames and the fused product are shown in Fig. 10. As in the previous experiment, we applied the blur-invariant phase correlation. Since there was just a shift between the frames, the entire procedure run in the Cartesian coordinates. We assumed a circular out-of-focus blur, so we used the operator P_∞. After the registration, the fusion itself was performed by the method proposed in Zhang et al. (2020). High visual quality of the fused product with almost no artifacts proves the accuracy of the registration.

8 Conclusion

8.1 Paper Summary

In this paper, we presented the general theory of invariants with respect to blur. The main original contribution of the paper lies in Theorems 3 and 6. We showed that all previously published examples of blur invariants are just particular cases of a unified theory, which can be formulated by means of projection operators without a limitation to a single blur type. The central role is played by the *orthogonality* of the projection operator, which is sufficient to possess the invariant properties. The connection between orthogonality of projections and blur invariance had not been known before. This discovery significantly contributes to the understanding of blur invariants. Using the proposed theory, we derived blur invariants to new unexplored blur types, which would be difficult to construct otherwise.

8.2 Evaluation of the Experiments

The experiments demonstrated a very good performance of the proposed invariants in the recognition of blurred objects and in blurred frames registration. Blur invariants exist in equivalent forms in Fourier domain where they are expressed directly by the projection operator and in the image domain where they use moment expansion. Both domains can be used in experiments and our choice mostly depends on the image size (for large images, Fourier invariants are more efficient and vice versa).

8.3 Comparison to CNN: What is Better?

The comparison to deep-learning methods, represented here by the ResNet CNN, performs another interesting outcome of the paper. We showed that in our experiments blur invariants significantly outperform the CNN without an augmentation in terms of recognition power and the CNN with a data augmentation in terms of computing complexity. This might, however, not be true in general for any kind of data. Considering memory consumption, the invariants are significantly more efficient even if no data augmentation was applied. Image representation by the invariants is highly compressive. We usually work with the feature vector of length less than 20 while CNN works with a complete pixel-wise representation of the images.

Blur invariants perform well if the intra-class variations can be mathematically modelled in a simple way (here the model was the convolution, possibly combined with a spatial transformation) even if the number of classes is high. Then the object classes are exactly the same as orbits of respective group action. This typically happens in the case of single-sample classes (for instance a single person defines one class and query images are various blurred and transformed photographs of the person). To reach a comparable recognition rate, CNNs require a massive augmentation over a wide range of blurs, which makes the training extremely time-consuming.

On the other hand, the proposed invariants can hardly be used for classification into broad generic classes such as "person", "car", "animal", "tree", etc. The intra-class variability is so rich that cannot be parameterized. The invariants do not have the ability to analyze the image content and they are not "continuous", which means that two visually similar objects (two dogs or two cars for instance) might have very different invariant values. These scenarios can be well resolved by deep learning, however, there is still the necessity of a large-scale augmentation of the training set with blur if blurred images are expected on the input of the system.

8.4 Where to go Next?

The proposed invariants and CNNs with augmentation should be understood as *complementary* rather than *competitive* approaches, each of them dominating in distinct situations (two extremes are mentioned above). Practical scenarios are often somewhere "in between". An optimal solution could be found as a weighted fusion of handcrafted and learned features where the weights are automatically determined during the learning process depending on the data. The fusion itself may be done in various ways. We can concatenate learned and handcrafted features into a single feature vector which is sent to SVM classifier. Alternatively, we can run two parallel networks of different architectures, which we fuse before the decision level. Yet another approach is to run traditional CNN and calculate the handcrafted features on the learned feature maps. Regardless of the particular fusion rule, this general idea can be viewed as a kind of augmentation, not applied on the training set as usually but on the feature space representing the images, where the pixel level is extended by a high-level handcrafted features.

It is highly probable that using learned and handcrafted blur-invariant features together will require developing of new network architectures. The research on this field is at a very initial stage and we envisage its dynamic development in the near future.

Declarations

Appendix

Proof of Theorem 6.

The proof follows the scheme shown in Fig. 5. Let us first investigate the relation between $I(f)$ and $I(f')$. By definition, we have

$$I(f')(\mathbf{u}) = \frac{\mathcal{F}(f')(\mathbf{u})}{\mathcal{F}(Pf')(\mathbf{u})}.$$

Now we show that $Pf'(\mathbf{x}) = (Pf)(A\mathbf{x})$. Since

$$f'(\mathbf{x}) \equiv f(A\mathbf{x}) = (Pf)(A\mathbf{x}) + f_A(A\mathbf{x})$$

and $(Pf)(A\mathbf{x}) \in \mathcal{S}$ thanks to the transformation closure, it is sufficient to show that $f_A(A\mathbf{x}) \in \mathcal{A}$. This follows from orthogonality of $\mathcal{S}$ and $\mathcal{A}$. Picking an arbitrary $h \in \mathcal{S}$, we have

$$\langle h(\mathbf{x}), f_A(A\mathbf{x}) \rangle = \int h(\mathbf{x}) f_A(A\mathbf{x}) \, \mathrm{d}\mathbf{x}$$
$$= \left\| A^{-1} \right\| \int h\left(A^{-1}\mathbf{x} \right) f_A(\mathbf{x}) \, \mathrm{d}\mathbf{x} = 0.$$

Since the decomposition is unique, it really holds $Pf'(\mathbf{x}) = Pf(A\mathbf{x})$. Recalling that the Fourier transform of arbitrary f is under A transformed as $\mathcal{F}(f')(\mathbf{u}) = \left\| A^{-1} \right\| \mathcal{F}(f)\left(A^{-\mathrm{T}}\mathbf{u} \right)$, where $A^{-\mathrm{T}}$ denotes $\left(A^{-1} \right)^{\mathrm{T}}$, we obtain

$$I(f')(\mathbf{u}) = \frac{\mathcal{F}(f)\left(A^{-\mathrm{T}}\mathbf{u} \right)}{\mathcal{F}(Pf)\left(A^{-\mathrm{T}}\mathbf{u} \right)} = I(f)\left(A^{-\mathrm{T}}\mathbf{u} \right).$$

After applying (formally) the inverse Fourier transform, we get the relation between the primordial image f_r of f and f_r' of f'

$$f_r'(\mathbf{x}) = \|A\| f_r(A\mathbf{x}).$$

This relation tells us that the primordial image is transformed by *the same* coordinate transformation as the original image, up to multiplicative factor $\|A\|$.

Let us recall that the blur invariants (16) are in fact moments of the primordial image. Let us denote those of $f_r(\mathbf{x})$ as $C_\mathbf{k}$, those of $f_r'(\mathbf{x})$ as $C_\mathbf{k}'$, and those of $f_r(A\mathbf{x})$ as $\tilde{C}_\mathbf{k}$. The relation between $C_\mathbf{k}'$ and $\tilde{C}_\mathbf{k}$ is

$$C_\mathbf{k}' = \int \pi_\mathbf{k}(\mathbf{x}) f_r'(\mathbf{x}) \, \mathrm{d}\mathbf{x}$$
$$= \int \|A\| \pi_\mathbf{k}(\mathbf{x}) f_r(A\mathbf{x}) \, \mathrm{d}\mathbf{x} = \|A\| \tilde{C}_\mathbf{k}.$$

Since J is an invariant w.r.t. A, it must hold $J(C_\mathbf{0}, \ldots, C_\mathbf{p}) = J(\tilde{C}_\mathbf{0}, \ldots, \tilde{C}_\mathbf{p})$. Consequently,

$$J(C_\mathbf{0}', \ldots, C_\mathbf{p}') = J(\|A\|\tilde{C}_\mathbf{0}, \ldots, \|A\|\tilde{C}_\mathbf{p})$$
$$= \|A\|^{-w} J(C_\mathbf{0}, \ldots, C_\mathbf{p}).$$

The last equality follows from the theory of affine moment invariants (Suk & Flusser, 2011), where it was shown that any absolute invariant J must have a form of a finite sum, where all terms are products of K moments (K is called the *degree* of the invariant) divided by $(K + w)$-th power of $m_\mathbf{0}$. This completes the proof.

References

Abu-Mostafa, Y. S., & Psaltis, D. (1984). Recognitive aspects of moment invariants. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 6*(6), 698–706.

Ahonen, T., Rahtu, E., Ojansivu, V., & Heikkila, J. (2008). Recognition of blurred faces using local phase quantization. In *19th International conference on pattern recognition, ICPR'08* (pp. 1–4). IEEE.

Bentoutou, Y., & Taleb, N. (2005a). A 3-D space-time motion detection for an invariant image registration approach in digital subtraction angiography. *Computer Vision and Image Understanding, 97*, 30–50.

Bentoutou, Y., & Taleb, N. (2005b). Automatic extraction of control points for digital subtraction angiography image enhancement. *IEEE Transactions on Nuclear Science, 52*(1), 238–246.

Bentoutou, Y., Taleb, N., Chikr El Mezouar, M., Taleb, M., & Jetto, L. (2002). An invariant approach for image registration in digital subtraction angiography. *Pattern Recognition, 35*(12), 2853–2865.

Bentoutou, Y., Taleb, N., Kpalma, K., & Ronsin, J. (2005). An automatic image registration for applications in remote sensing. *IEEE Transactions on Geoscience and Remote Sensing, 43*(9), 2127–2137.

Blum, R. S., & Liu, Z. (2006). *Multi-sensor image fusion and its applications*. CRC.

Boldyš, J., & Flusser, J. (2013). Invariants to symmetrical convolution with application to dihedral kernel symmetry. In A. Petrosino (Ed.), *Proceedings of the 17th international conference on image analysis and processing ICIAP'13, Lecture notes in computer science* (Vol. 8157, pp. 369–378), part II. Springer.

Campisi, P., & Egiazarian, K. (2007). *Blind image deconvolution: Theory and applications*. CRC.

Candocia, F. M. (2004). Moment relations and blur invariant conditions for finite-extent signals in one, two and *n*-dimensions. *Pattern Recognition Letters, 25*, 437–447.

Chen, B., Shu, H., Zhang, H., Coatrieux, G., Luo, L., & Coatrieux, J. L. (2011). Combined invariants to similarity transformation and to blur using orthogonal Zernike moments. *IEEE Transactions on Image Processing, 20*(2), 345–360.

Cubuk, E. D., Zoph, B., Mané, D., Vasudevan, V., & Le, Q. V. (2019). Autoaugment: Learning augmentation strategies from data. In *2019 IEEE/CVF conference on computer vision and pattern recognition (CVPR)* (pp. 113–123).

Cyganski, D., & Orr, J. A. (1988). Object recognition and orientation determination by tensor methods. In T. S. Huang (Ed.), *Advances in computer vision and image processing* (pp. 101–144). JAI Press.

Dai, X., Zhang, H., Shu, H., & Luo, L. (2010a). Image recognition by combined invariants of Legendre moment. In *Proceedings of the IEEE international conference on information and automation ICIA'10*, Harbin, China (pp. 1793–1798).

Dai, X., Zhang, H., Shu, H., Luo, L., & Liu, T. (2010b). Blurred image registration by combined invariant of Legendre moment and Harris-Laplace detector. In *Proceedings of the fourth pacific-rim symposium on image and video technology PSIVT'10* (pp. 300–305). IEEE.

Dai, X., Liu, T., Shu, H., & Luo, L. (2013). Pseudo-Zernike moment invariants to blur degradation and their use in image recognition. In J. Yang, F. Fang, & C.Sun (Eds.), *Intelligent science and intelligent data engineering IScIDE'12*, Lecture notes in computer science (Vol. 7751., pp. 90–97). Springer.

de Castro, E., & Morandi, C. (1987). Registration of translated and rotated images using finite Fourier transform. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 9*(5), 700–703.

Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. In *CVPR09*.

Dodge, S., & Karam, L. (2016). Understanding how image quality affects deep neural networks. In *2016 Eighth international conference on quality of multimedia experience (QoMEX)* (pp. 1–6). IEEE.

Flusser, J., & Zitová, B. (2004). Invariants to convolution with circularly symmetric PSF. In *Proceedings of the 17th international conference on pattern recognition ICPR'04* (pp. 11–14). IEEE Computer Society.

Flusser, J. (2000). On the independence of rotation moment invariants. *Pattern Recognition, 33*(9), 1405–1410.

Flusser, J., Boldyš, J., & Zitová, B. (2000). Invariants to convolution in arbitrary dimensions. *Journal of Mathematical Imaging and Vision, 13*(2), 101–113.

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., Farokhi, S., Höschl, C., IV., Suk, T., Zitova, B., & Pedone, M. (2016a). Recognition of images degraded by Gaussian blur. *IEEE Transactions on Image Processing, 25*(2), 790–806.

Flusser, J., & Suk, T. (1997). Classification of degraded signals by the method of invariants. *Signal Processing, 60*(2), 243–249.

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

Flusser, J., Suk, T., & Saic, S. (1995). Image features invariant with respect to blur. *Pattern Recognition, 28*(11), 1723–1732.

Flusser, J., Suk, T., & Saic, S. (1996a). Recognition of blurred images by the method of moments. *IEEE Transactions on Image Processing, 5*(3), 533–538.

Flusser, J., Suk, T., & Saic, S. (1996b). Recognition of images degraded by linear motion blur without restoration. *Computing Supplement, 11*, 37–51.

Flusser, J., Suk, T., & Zitová, B. (2009). *Moments and moment invariants in pattern recognition*. Wiley.

Flusser, J., Suk, T., & Zitová, B. (2013). On the recognition of wood slices by means of blur invariants. *Sensors and Actuators A: Physical, 198*, 113–118.

Flusser, J., Suk, T., & Zitová, B. (2014). Comments on 'Weed recognition using image blur information' by Peng, Z. & Jun, C., Biosystems Engineering 110 (2), p. 198–205. *Biosystems Engineering, 126*, 104–108.

Flusser, J., Suk, T., & Zitová, B. (2016b). *2D and 3D image analysis by moments*. Wiley.

Galigekere, R. R., & Swamy, M. N. S. (2006). Moment patterns in the Radon space: Invariance to blur. *Optical Engineering, 45*(7), 077003.

Georghiades, A., Belhumeur, P., & Kriegman, D. (2001). From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 23*(6), 643–660.

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.

Guang-Sheng, C., & Peng, Z. (2012). Dynamic wood slice recognition using image blur information. *Sensors and Actuators A: Physical, 176*, 27–33.

Guan, B., Wang, S., & Wang, G. (2005). A biologically inspired method for estimating 2D high-speed translational motion. *Pattern Recognition Letters, 26*, 2450–2462.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *The IEEE conference on computer vision and pattern recognition (CVPR)*.

Hickman, M. S. (2012). Geometric moments and their invariants. *Journal of Mathematical Imaging and Vision, 44*(3), 223–235.

Hilbert, D. (1993). *Theory of algebraic invariants*. Cambridge University Press.

Hu, S. X., Xiong, Y. M., Liao, M. Z., & Chen, W. F. (2007). Accurate point matching based on combined moment invariants and their new statistical metric. In *Proceedings of the international conference on wavelet analysis and pattern recognition ICWAPR'07* (pp. 376–381). IEEE Computer Society.

Ji, H., & Zhu, H. (2009). Degraded image analysis using Zernike moment invariants. In *Proceedings of the international conference on acoustics, speech and signal processing ICASSP'09* (pp. 1941–1944).

Kautsky, J., & Flusser, J. (2011). Blur invariants constructed from arbitrary moments. *IEEE Transactions on Image Processing, 20*(12), 3606–3611.

Kostková, J., Flusser, J., Lébl, M., & Pedone, M. (2020). Handling Gaussian blur without deconvolution. *Pattern Recognition, 103*, 107264.

Kotera, J., Šmídl, V., & Šroubek, F. (2017). Blind deconvolution with model discrepancies. *IEEE Transactions on Image Processing, 26*(5), 2533–2544.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM, 60*(6), 84–90.

Kundur, D., & Hatzinakos, D. (1996). Blind image deconvolution. *IEEE Signal Processing Magazine, 13*(3), 43–64.

Lébl, M., Šroubek, F., & Flusser, J. (2023). Impact of image blur on classification and augmentation of deep convolutional networks. In *2023 Scandinavian conference on image analysis*. Lecture notes in computer science (Vol. 13886, pp. 108–117). Springer.

Lébl, M., Šroubek, F., Kautsky, J., & Flusser, J. (2019). Blur invariant template matching using projection onto convex sets. In *Proc. 18th Int'l. Conf. Computer analysis of images and patterns CAIP'19* Lecture notes in computer science (Vol. 13886, pp. 108–117). Springer.

Lébl, M., Šroubek, F., & Flusser, J. (2021). Blur-invariant similarity measurement of images. *IEEE Transactions on Pattern Analysis & Machine Intelligence, 43*(08), 2882–2884.

Liu, Q., Zhu, H., & Li, Q. (2011a). Image recognition by combined affine and blur Tchebichef moment invariants. In *Proceedings of 4th international conference on image and signal processing (CISP)* (pp. 1517–1521).

Liu, Q., Zhu, H., & Li, Q. (2011b). Object recognition by combined invariants of orthogonal Fourier-Mellin moments. In *Proceedings of 8th international conference on information, communications and signal processing ICICS'11* (pp. 1–5). IEEE.

Liu, Z., An, J., & Li, L. (2011). A two-stage registration algorithm for oil spill aerial image by invariants-based similarity and improved ICP. *International Journal of Remote Sensing, 32*(13), 3649–3664.

Liu, J., & Zhang, T. (2005). Recognition of the blurred image by complex moment invariants. *Pattern Recognition Letters, 26*(8), 1128–1138.

Lo, C. H., & Don, H. S. (1989). 3-D moment forms: Their construction and application to object identification and positioning. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 11*(10), 1053–1064.

Mahdian, B., & Saic, S. (2007). Detection of copy-move forgery using a method based on blur moment invariants. *Forensic Science International, 171*(2–3), 180–189.

Makaremi, I., & Ahmadi, M. (2010). Blur invariants: A novel representation in the wavelet domain. *Pattern Recognition, 43*, 3950–3957.

Makaremi, I., & Ahmadi, M. (2012). Wavelet domain blur invariants for image analysis. *IEEE Transactions on Image Processing, 21*(3), 996–1006.

Ojansivu, V., & Heikkilä, J. (2008a). Blur insensitive texture classification using local phase quantization. In A. Elmoataz, O. Lezoray, & F. Nouboud (Eds.), *Image and signal processing ICISP'08*, (Vol. 5099, pp. 236–243). Heidelberg, Berlin.

Ojansivu, V., & Heikkilä, J. (2008b) A method for blur and affine invariant object recognition using phase-only bispectrum. In *The international conference on image analysis and recognition ICIAR'08. LNCS* (Vol. 5112, pp. 527–536). Springer.

Ojansivu, V., & Heikkilä, J. (2007). Image registration using blur-invariant phase correlation. *IEEE Signal Processing Letters, 14*(7), 449–452.

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.

Pedone, M., Flusser, J., & Heikkilä, J. (2015). Registration of images with *N*-fold dihedral blur. *IEEE Transactions on Image Processing, 24*(3), 1036–1045.

Pei, Y., Huang, Y., Zou, Q., Zhang, X., & Wang, S. (2021). Effects of image degradation and degradation removal to CNN-based image classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 43*(4), 1239–1253.

Peng, Z., & Jun, C. (2011). Weed recognition using image blur information. *Biosystems Engineering, 110*(2), 198–205.

Rajagopalan, A. N., & Chellappa, R. (2014). *Motion deblurring: Algorithms and systems*. Cambridge University Press.

Reiss, T. H. (1991). The revised fundamental theorem of moment invariants. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 13*(8), 830–834.

Sadjadi, F. A., & Hall, E. L. (1980). Three dimensional moment invariants. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 2*(2), 127–136.

Šroubek, F., & Flusser, J. (2005). Multichannel blind deconvolution of spatially misaligned images. *IEEE Transactions on Image Processing, 14*(7), 874–883.

Stern, A., Kruchakov, I., Yoavi, E., & Kopeika, N. S. (2002). Recognition of motion-blured images by use of the method of moments. *Applied Optics, 41*, 2164–2172.

Suk, T., & Flusser, J. (2011). Affine moment invariants generated by graph method. *Pattern Recognition, 44*(9), 2047–2056.

Suk, T., Flusser, J., & Boldyš, J. (2015). 3D rotation invariants by complex moments. *Pattern Recognition, 48*(11), 3516–3526.

Vageeswaran, P., Mitra, K., & Chellappa, R. (2013). Blur and illumination robust face recognition via set-theoretic characterization. *IEEE Transactions on Image Processing, 22*(4), 1362–1372.

Vasiljevic, I., Chakrabarti, A., & Shakhnarovich, G. (2016). Examining the impact of blur on recognition by convolutional networks. arXiv preprint arXiv:1611.05760

Wang, S., Guan, B., Wang, G., & Li, Q. (2007). Measurement of sinusoidal vibration from motion blurred images. *Pattern Recognition Letters, 28*, 1029–1040.

Wee, C. Y., & Paramesran, R. (2007). Derivation of blur-invariant features using orthogonal Legendre moments. *IET Computer Vision, 1*(2), 66–77.

Xiao, B., Ma, J. F., & Cui, J. T. (2012). Combined blur, translation, scale and rotation invariant image recognition by Radon and pseudo-Fourier–Mellin transforms. *Pattern Recognition, 45*, 314–321.

Xu, D., & Li, H. (2006). 3-D affine moment invariants generated by geometric primitives. In *Proceedings of the 18th international conference on pattern recognition ICPR'06* (pp. 544–547). IEEE Computer Society.

Yap, P. T., & Raveendran, P. (2004). Image focus measure based on Chebyshev moments. *IEE Proceedings of the Vision, Image and Signal Processing, 151*(2), 128–136.

Zhang, Z., Klassen, E., & Srivastava, A. (2013). Gaussian blurring-invariant comparison of signals and images. *IEEE Transactions on Image Processing, 22*(8), 3145–3157.

Zhang, Y., Liu, Y., Sun, P., Yan, H., Zhao, X., & Zhang, L. (2020). IFCNN: A general image fusion framework based on convolutional neural network. *Information Fusion, 54*, 99–118.

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.

Zhang, Y., Wen, C., & Zhang, Y. (2000a). Estimation of motion parameters from blurred images. *Pattern Recognition Letters, 21*(5), 425–433.

Zhang, Y., Wen, C., Zhang, Y., & Soh, Y. C. (2002). Determination of blur and affine combined invariants by normalization. *Pattern Recognition, 35*(1), 211–221.

Zhang, Y., Zhang, Y., & Wen, C. (2000b). A new focus measure method using moments. *Image and Vision Computing, 18*(12), 959–965.

Zhong, S., Liu, Y., Liu, Y., & Li, C. S. (2013). Water reflection recognition based on motion blur invariant moments in curvelet space. *IEEE Transactions on Image Processing, 22*(11), 4301–4313.

Zhou, Y., Song, S., & Cheung, N. M. (2017). On classification of distorted images with deep convolutional neural networks. *2017 IEEE International conference on acoustics, speech and signal processing (ICASSP)* (pp. 1213–1217). IEEE.

Zhu, H., Liu, M., Ji, H., & Li, Y. (2010). Combined invariants to blur and rotation using Zernike moment descriptors. *Pattern Analysis and Applications, 3*(13), 309–319.

Zitová, B., & Flusser, J. (2003). Image registration methods: A survey. *Image and Vision Computing, 21*(11), 977–1000.

Zuo, X., Dai, X., & Luo, L. (2010). M-SIFT: A new descriptor based on Legendre moments and SIFT. In *Proceedings of the 3rd international conference on machine vision ICMV'10* (pp. 183–186).