Signal Processing 143 (2018) 282–291

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Rotation invariants from Gaussian-Hermite moments of color images

Bo Yang[a], Tomáš Suk[b], Jan Flusser[b,*], Zhongke Shi[a], Xiaofeng Chen[a]

[a] School of Automation, Northwestern Polytechnical University, 127 West Youyi Road, 710 072 Xi'an Shaanxi, PR China
[b] Institute of Information Theory and Automation of the CAS, Pod Vodárenskou věží 4, 18 208 Prague 8, Czech Republic

ARTICLE INFO

Article history:
Received 27 February 2017
Revised 15 August 2017
Accepted 31 August 2017
Available online 14 September 2017

Keywords:
Color images
Object recognition
Rotation invariants
Gaussian–Hermite moments
Joint invariants

ABSTRACT

The topic of the paper is recognition of objects and patterns in color images regardless of their position, orientation, and scale. Gaussian–Hermite moment invariants designed especially for color images are introduced in this paper. We extend the existing invariants for graylevel images and show that in the case of color images there exist additional independent invariants, which can be constructed as joint invariants from cross-channel moments and/or from new non-trivial low-order moments. The experiments on real data confirmed that the new invariants improve the recognition rate.

1. Introduction

Invariant-based object recognition has been a goal of much recent research. This requirement appears quite often in real recognition problems, where the patterns that we would like to detect in the images are different from the dictionary (training) patterns. While the training patterns are usually stored under ideal conditions and in somehow normalized positions, the actual patterns may have been rotated, translated or scaled. That is why the invariance of the features with respect to the assumed intra-class variabilities is very important. The invariance is, however, not the only requirement imposed on good features. The other one, which is of the same importance, is the ability to discriminate patterns belonging to different classes.

Designing a class of suitable features to represent 2D objects has been a topic of thousands of papers and several monographs (see, for instance, [1] for a survey and other references). Among various types of features which have been proposed, *moment invariants* play an important role thanks to their global information representation, to their ability to cope with many pattern deformations, and to their easy and stable numerical implementations. They have been regarded as a kind of robust and powerful descriptors for object representation and recognition. Many successful practical applications have been reported, both in recognition of 2D as well as 3D structures (see [1] or [2] or for the history and the state of the art of moment invariants).

Invariants to translation, rotation and scaling (TRS) of *graylevel* (i.e. single-channel) images were introduced as early as in the 60's [3] and improved, modified and further developed many times [4–10]. Many authors have approved that using orthogonal moments leads to better numerical stability and improves the recognition power. In 2D, there exist two families of orthogonal (OG) polynomials, which differ from one another by the area of orthogonality – polynomials orthogonal on a disk and polynomials orthogonal on a square/rectangle. The former group is inherently suitable for constructing rotation invariants, because these moments change under rotation in a simple way and the rotation parameter can be eliminated easily. This was noted for example by Teague [11], Khotanzad and Hong [5], and Wallin and Kubler [12] who used Zernike moments, and by other authors who employed pseudo-Zernike moments [13], Fourier–Mellin moments [14–16], Jacobi–Fourier moments [17], and Chebyshev–Fourier moments [18]. The negative aspect of using moments OG on a disk is that they require mapping of the image into the disk, which is equivalent to image scaling and polar transformation. This operation leads to a precision loss due to the image resampling and also increases the computation time. That is why some authors prefer to use the moments OG on a square/rectangle, such as Legendre moments [7,19,20], Chebyshev moments [21–23], Hermite and Gaussian–Hermite moments [24,25], Krawtchouk moments [26], and Gegenbauer moments [27,28]. However, construction of rotation invariants from these moments is generally very difficult. Hermite (and modified Hermite) moments are the only exception. They offer a possibility of an easy and efficient design of rotation and scaling invariants because Hermite polynomials are trans-

* Corresponding author.
E-mail address: flusser@utia.cas.cz (J. Flusser).

http://dx.doi.org/10.1016/j.sigpro.2017.08.027
0165-1684/

formed under rotation exactly in the same way as do the monomials $x^p y^q$, as was discovered by Yang et al. [10,29]. Hence, the theory which had been developed for geometric moments was adapted to design rotation invariants of Hermite and Gaussian–Hermite (GH) moments of theoretically arbitrary orders and reasonable numerical stability [30,31]. Scaling invariance of the GH moments was achieved as well [32]. Thanks to that, the GH invariants have been one of the latest state-of-the-art techniques in object recognition. They have been used in a variety of practical applications such as in the traffic management and surveillance [33], in license plate recognition [34], in biometric identification for fingerprint recognition [35], iris recognition [36], and for infrared face recognition [37].

So far, most existing moment invariants including the GHIs have been designed for single-channel images. This has been justified by the fact that in many practical applications moment invariants are applied on segmented binary objects to capture their shape rather than their texture. However, in some cases we may want to apply invariants directly on color images. Color images can be viewed as three graylevel images corresponding to R, G, and B channels. In principle, single-channel invariants can be applied to each channel, but such an approach is only suboptimal. There exists a strong link among R, G, and B channels, which can be employed. If the color image has been rotated, the rotation angle is the same for all three channels. The same is true for translation and scaling. Considering that, we can derive so-called *joint invariants*, which contain moments of different channels. These joint invariants do not have any counterpart in graylevel image analysis and increase the number of the independent invariants, which consequently may increase the discrimination power. Looking for an independent and complete invariant set, composed of both single-channel and joint invariants, is the main goal of the paper.

The rest of the paper is organized as follows. Section 2 gives a brief introduction to 2D GHIs. The GHIs for color images are introduced in Section 3. Section 4 demonstrates experimentally their performance in image classification and template matching. Section 5 concludes the paper.

2. Gaussian–Hermite moments and invariants

In this section, we recall shortly the basic terms used further in the paper. For more details we refer to [10,29,30,38].

Hermite polynomial of degree p is defined as

$$H_p(x) = (-1)^p \exp(x^2) \frac{\mathrm{d}^p}{\mathrm{d}x^p} \exp(-x^2). \tag{1}$$

Hermite polynomials are orthogonal on $(-\infty, \infty)$ with respect to weighting function

$$w(x) = e^{-x^2}, \tag{2}$$

which yields

$$\int_{-\infty}^{\infty} w(x) H_n(x) H_m(x) \mathrm{d}x = n! 2^n \sqrt{\pi} \delta_{nm} \tag{3}$$

where δ_{nm} is the Kronecker symbol. Note, that Hermite polynomials are not orthonormal without a further normalization.

The amplitudes of Hermite polynomials grow very fast as $|x|$ increases. That is why it is common to work with weighted (modulated) Gaussian–Hermite (GH) polynomials

$$\hat{H}_p(x;\sigma) = H_p\left(\frac{x}{\sigma}\right) \exp\left(-\frac{x^2}{2\sigma^2}\right), \tag{4}$$

where σ is a user-defined scale parameter which controls the attenuation of the polynomials. Gaussian–Hermite moments of single-channel image $f(x, y)$ are then defined as

$$\hat{\eta}_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \hat{H}_p(x - x_c;\sigma) \hat{H}_q(y - y_c;\sigma) f(x, y) \mathrm{d}x \mathrm{d}y, \tag{5}$$

where (x_c, y_c) is the image centroid. Still, these moments grow very fast with the order. To keep them in a reasonable range, we work with *normalized GH moments*

$$\eta_{pq} = \frac{1}{\sigma\sqrt{\pi 2^{p+q} ((p+q)!)^{1/2} \Gamma((p+q)/2+1)}} \hat{\eta}_{pq}. \tag{6}$$

This normalization does not influence the design of the invariants, it only insures better numerical stability of the moments. Other normalization coefficients of a similar form may be used as well (see [1] for more details).

Rotation invariants of GHMs for graylevel images were designed based on a quantity d_{pq}, which is actually a linear combination of GHMs of the same order

$$d_{pq} = \sum_{k=0}^{p} \sum_{j=0}^{q} \binom{p}{k} \binom{q}{j} (-1)^{q-j} i^{p+q-k-j} \eta_{k+j,p+q-k-j}, \tag{7}$$

where i is imaginary unit. Note that the value of d_{pq} is generally a complex number for which $d_{pq} = d^*_{qp}$.

The key property of d_{pq} is its simple transformation under rotation. If the image has been rotated by angle α, d_{pq} preserves its magnitude while its phase is shifted as

$$d'_{pq} = d_{pq} \exp(-i(p-q)\alpha). \tag{8}$$

The above relation offers an infinite number of rotation invariants which are constructed as products of various d_{pq}'s such that the overall phase shift is cancelled. A complete and independent system of rotation invariants can be obtained as

$$\Phi_{pq} = d_{pq} d_{q_0 p_0}^{p-q}, \quad \text{with} \quad p \geq q, \quad p_0 - q_0 = 1, \tag{9}$$

where again $\Phi_{pq} = \Phi^*_{qp}$ (see [30] for a detailed derivation and properties of these invariants).

3. Invariants for color images

The chief problem of designing moment invariants for color images is how to incorporate the link, which exists among R, G, and B channels, into the invariant set. As we already pointed out, by applying traditional single-channel invariants individually to R, G, and B, we miss other independent invariants, which can be constructed from the cross-moments of two or three channels and which can contribute to the discrimination power.

There have been very few papers on moment invariants of color images [39–41]. One may discover two different approaches to this problem in the literature. Suk and Flusser [39] proposed *joint affine invariants* for color images, but they were constructed from geometric moments only and suffer with numerical instability. They cannot be easily re-formulated in terms of GH or other orthogonal moments. Another approach was proposed by Mindru et al. [40], who used geometric moments of integer powers of R, G, and B channels to create the invariants. However, this approach is numerically unstable because of high dynamic range of the moments and may lead to redundant invariant set, if the exponents of the channel image functions have not been selected carefully. Guo et al. [41] used three imaginary components of quaternions for description of colors, but quaternion formalism seems to be useless here because we face the rotation in the coordinate plane only while the intensity values in the color space do not rotate.

In this paper, we chose the first mentioned principle and we construct the joint invariants from GH moments of different channels. First, we do so for a translation and rotation (TR) only, then we resolve the general case of a *similarity transformation*, which includes translation, rotation, and scaling (TRS).

3.1. Invariance to translation

Translation invariance is traditionally provided by shifting the coordinate origin to the centroid of the object. The centroid (x_c, y_c)

is computed from the zero and first-order geometric moments m_{00}, m_{10}, m_{01} as $x_c = m_{10}/m_{00}$ and $y_c = m_{01}/m_{00}$ (see [1] or [2] for the definition and basic properties of geometric moments). The color image has three centroids of individual color channels and also a joint centroid (we keep the notation (x_c, y_c) for simplicity), defined as

$$\begin{aligned} m_{00} &= m_{00}^{(R)} + m_{00}^{(G)} + m_{00}^{(B)} \\ x_c &= (m_{10}^{(R)} + m_{10}^{(G)} + m_{10}^{(B)})/m_{00} \\ y_c &= (m_{01}^{(R)} + m_{01}^{(G)} + m_{01}^{(B)})/m_{00}. \end{aligned} \tag{10}$$

The symbols with a superscript refer to individual channels, the symbols without a superscript refer to joint quantities. Now, if we shift the coordinate origin into the joint centroid, the first-order central moments of individual channels are generally non-zero (they express the deviation of the channel centroids from the joint centroid) and we can use them for constructing invariants. This is one of the differences from the single-channel case, where the first-order central moments are zero by definition.

3.2. The number of independent invariants

How many invariants can we construct from the moments up to the order r? It is well known that in general the number n_i of independent invariants created from n_m independent measurements (i.e. moments in our case) is

$$n_i = n_m - n_p, \tag{11}$$

where n_p is the number of independent transformation parameters that should be eliminated. In the case of TR, $n_p = 3$ (two translations and a rotation), in the case of TRS we have $n_p = 4$ (two translations, rotation, and scaling). The number of moments of orders from 0 to r of a single channel equals

$$n_m = 1 + 2 + 3 + \ldots + (r+1) = \frac{(r+1)(r+2)}{2}. \tag{12}$$

In a multichannel case, we can create $n_c n_i$ single-channel invariants (n_c is the number of the channels; $n_c = 3$ for RGB color images). However, since the transformation of all channels is the same, we have to eliminate only n_p parameters, the same number as in the single-channel case. That means there must be $n_c n_m - n_p$ independent invariants in total. It implies there exist $(n_c - 1)n_p$ additional independent invariants (which means six in the case of TR and eight for TRS transformations, respectively), which should be added to the set of all single-channel invariants. Adding such invariants may improve the recognition power without increasing the order of the moments involved. In the next section, we discuss two possibilities how to design these "additional" invariants.

3.3. New independent invariants for color images

In the previous section, we explained why there must exist additional invariants for multichannel images and we also showed that their number is a simple function of the number of the transformation parameters and the number of the image channels. Now we show how to actually construct them.

There are basically two approaches. One way is to employ the zeroth- and the first-order moments, which have not been involved in the single-channel invariants (we recall that for single-channel images the first-order central moments always vanish, but for color images we consider the central moments with respect to the joint centroid of all channels which makes the first-order central moments nontrivial). The other way is to construct *joint invariants*, which contain moments of more than one channel. Both ways could be even combined together.

In the next two sections, we demonstrate both approaches in the case of TR and TRS transformations, respectively.

3.4. Invariants to translation and rotation

The easiest way to designing a complete system is to start with accepting all the single-channel invariants from Eq. (9): $\Phi_{pq}^{(R)}$, $\Phi_{pq}^{(G)}$, and $\Phi_{pq}^{(B)}$ for $p + q \geq 2$. Three additional independent invariants could be the zero-order ones $\Phi_{00}^{(R)}$, $\Phi_{00}^{(G)}$, and $\Phi_{00}^{(B)}$, which actually equal to the zero-order moments and are inherently invariant to rotation. The other three additional invariants could be the first-order single-channel invariants $\Phi_{01}^{(C)}$, $C \in \{R, G, B\}$, as was mentioned in the previous section. We denote this set of invariants as SRGB (Single-channel Red, Green, and Blue).

Now let us incorporate the joint invariants. In SRGB, we replace three invariants $\Phi_{01}^{(C)}$ by three joint invariants

$$\begin{aligned} \Phi_{01}^{(J1)} &= d_{10}^{(R)} d_{01}^{(G)} \\ \Phi_{01}^{(J2)} &= d_{10}^{(G)} d_{01}^{(B)} \\ \Phi_{01}^{(J3)} &= d_{10}^{(B)} d_{01}^{(R)}. \end{aligned} \tag{13}$$

Note that they are actually independent of the rest of the SRGB set. We call this set, which contains three joint invariants, JSRGB (Joint and Single-channel Red, Green, and Blue).

Yet another possibility of a creation of the complete and independent set is to use the joint invariants whenever it is possible. Joint invariants $\Phi(p, q)^{(JC)}$ are defined as

$$\left.\begin{aligned} \Phi_{pq}^{(JR)} &= d_{pq}^{(R)} \\ \Phi_{pq}^{(JG)} &= d_{pq}^{(G)} \\ \Phi_{pq}^{(JB)} &= d_{pq}^{(B)} \end{aligned}\right\} p = q, \quad \left.\begin{aligned} \Phi_{pq}^{(JR)} &= d_{qp}^{(R)} d_{pq}^{(G)} \\ \Phi_{pq}^{(JG)} &= d_{qp}^{(G)} d_{pq}^{(B)} \\ \Phi_{pq}^{(JB)} &= d_{qp}^{(B)} d_{pq}^{(R)} \end{aligned}\right\} p < q,$$
$$\left.\begin{aligned} \Phi_{pq}^{(JR)} &= \mathcal{R}e(d_{pq}^{(R)} (d_{12}^{(G)})^{p-q}) \\ \Phi_{pq}^{(JG)} &= \mathcal{R}e(d_{pq}^{(G)} (d_{12}^{(B)})^{p-q}) \\ \Phi_{pq}^{(JB)} &= \mathcal{R}e(d_{pq}^{(B)} (d_{12}^{(R)})^{p-q}) \end{aligned}\right\} p > q. \tag{14}$$

This invariant set is called JRGB (Joint Red, Green, and Blue).

3.5. Invariants to translation, rotation, and scaling

There are several distinctions between the TRS and TR transformations, which influence the way how the invariant sets are created. The first one comes from the behavior of the GH moments under scaling. Geometric moments, as well as all other common moments, are only multiplied by certain power of the image integral if the image has been scaled. Since the GH moments are weighted with a Gaussian (see (4)), the parameter σ of this Gaussian must be also modified to achieve scale-invariant version of GH moments. This is not a trivial task which has been resolved by means of a *variable modulation*, see [32]. Roughly speaking, that method sets σ such that it is not constant but depends on the zero-order moment of the image. One has to use this variable modulation whenever scaling invariance of the GH moments is required.

As we already explained, we look for eight additional TRS invariants to be appended to the set of the single-channel invariants. The simplest way is to take the SRGB set from the previous section and incorporate two zero-order ratios

$$\Phi_{00}^{(RG)} = d_{00}^{(R)}/d_{00}^{(G)},$$

$$\Phi_{00}^{(BG)} = d_{00}^{(B)}/d_{00}^{(G)}.$$

Alternatively, we may replace three invariants $\Phi_{10}^{(C)}$, $C \in \{R, G, B\}$ in the SRGB set with three joint invariants and add other three independent joint invariants. We used the low-order ones: $\Phi_{01}^{(J1)}$, $\Phi_{01}^{(J2)}$, and $\Phi_{01}^{(J3)}$ as in the TR model and three new joint invariants

$$\Phi_{10}^{(J4)} = \mathcal{R}e((d_{10}^{(R)})^2 d_{02}^{(G)})$$

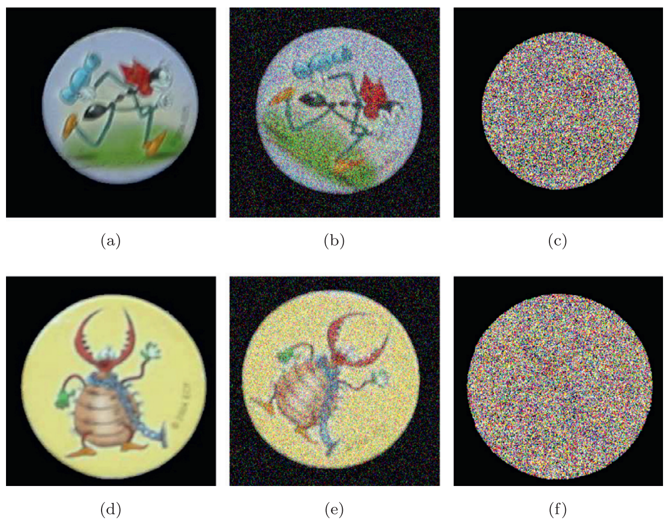


Fig. 1. Two "pick a pair" cards used in the first experiment: (a) – (c) "Ferdy the Ant", (d) – (f) "Ant lion". (b) and (e) are rotated cards with Gaussian noise of STD = 0.2 in the entire image, and (c) and (f) are rotated cards with Gaussian noise of STD = 2 added only to the object.

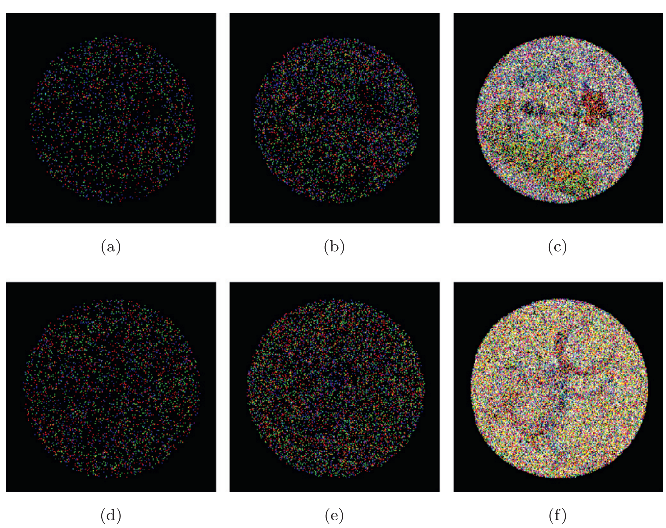


Fig. 2. Two cards corrupted by Poisson noise: (a) and (d) with intensity scaling factor $p = 0.05$, (b) and (e) with $p = 0.1$, and (c) and (f) with $p = 1$.

$$\Phi_{10}^{(J5)} = \mathcal{R}e((d_{10}^{(G)})^2 d_{02}^{(B)})$$
$$\Phi_{10}^{(J6)} = \mathcal{R}e((d_{10}^{(B)})^2 d_{02}^{(R)}). \qquad (15)$$

In this way, we constructed the set JSRGB for TRS transformation.

As in the case of TR, we may create a set JRGB of joint invariants. It consists of the invariants $\Phi_{00}^{(RG)}$, $\Phi_{00}^{(BG)}$, further $\Phi_{10}^{(J1)}$, $\Phi_{10}^{(J2)}$, $\Phi_{10}^{(J3)}$ from (13), and also $\Phi_{10}^{(J4)}$, $\Phi_{10}^{(J5)}$, and $\Phi_{10}^{(J6)}$ from (15). The set is completed by $\Phi(p, q)^{(JC)}$ from (14) for $p + q \geq 2$.

Summarizing, we have SRGB, JSRGB and JRGB sets of the invariants for each of the transformations TR and TRS. In the next section, we test numerically their performance in real object recognition tasks.

4. Numerical experiments

In this section we design experiments to show the behavior of the GHIs for color images. The first and the third experiments assume the TRS model; the second one, where the scaling was not present, assumes TR transformation. Two experiments show classification of simple color images irrespectively on their orientation/size, one experiment shows the performance of the GHI's in rotation-invariant template matching. In all experiments, whenever the invariants Φ_{pq} are used, we set $p_0 = 2$ and $q_0 = 1$. The goal of these experiments is not only to show if (and how well) the GHI's perform but namely to test and evaluate different sets of the GH invariants. The experiments do not compare GHI's to other invariants such as Zernike or geometric because such studies were published several times and clearly proved the advantageous properties of the GHI's, especially the numerical stability of higher orders (see for instance [30–32]).

It is worth mentioning that for stable numerical implementation of Hermite polynomials (as well as of all other orthogonal polynomials) one should use the three-term recurrent formulas (see [38] or [1]) instead of the definition (1).

Table 1
The number of errors out of 77 trials in the experiment with "Ferdy the Ant" cards with Gaussian background noise.

Order	Gray	RGB	SRGB	JSRGB	JRGB
3	63	29	4	3	6
4	54	17	6	8	3
5	57	27	12	13	12
6	50	18	12	12	6
7	46	16	11	11	6
8	51	15	15	15	7

4.1. Card recognition

In this experiment, we used photographs of twelve round cards from the "pick a pair" game. Each card was captured eight times by a camera that was rotated approximately by 45° between the snaps, so the image rotation is real. We did not introduce differences in the scale intentionally, but still, since the camera was hand-held, the size of the images is not precisely the same. This is why we used the TRS invariants in this experiment.

The aim of the experiment was to classify these test images. One image of each card was used as a representative of the class, the others were recognized by a nearest-neighbor classifier.

When we used the card images in the original quality, no classification error occurred whatever invariant set was used. To make the task more challenging, we introduced computer-generated zero-mean Gaussian noise of standard deviation 0.2 to all the images (the range of brightness of the images was from 0 to 1). The noise in different channels was independent and was applied on the entire image including the background, see Fig. 1(b) and (e) for two noisy cards. The other original cards used in the experiment are shown in Fig. 4.

First, we converted the images into graylevel versions and computed $\Phi(p, q)$ for $p + q \geq 2$. The number of errors is shown in Table 1 in the column "gray". Then we took the same invariants computed from individual color channels (column "RGB"). We compared them with the feature sets SRGB, JSRGB, and JRGB, which were proposed in Section 3.5. We ran all the experiments for different orders of the invariants; the first column of the table shows the maximum order used.

As we can see from Table 1, all color invariants perform much better than the graylevel ones. The best performing are JRGB, while SRGB and JSRGB are only slightly worse. The performance varies with the order, but overall there is no significant difference. Relatively high misclassification rates were caused namely by the background noise. The background noise is also responsible for poor performance when high-order invariants SRGB and JSRGB were used, because higher-order moments are more vulnerable to noise.

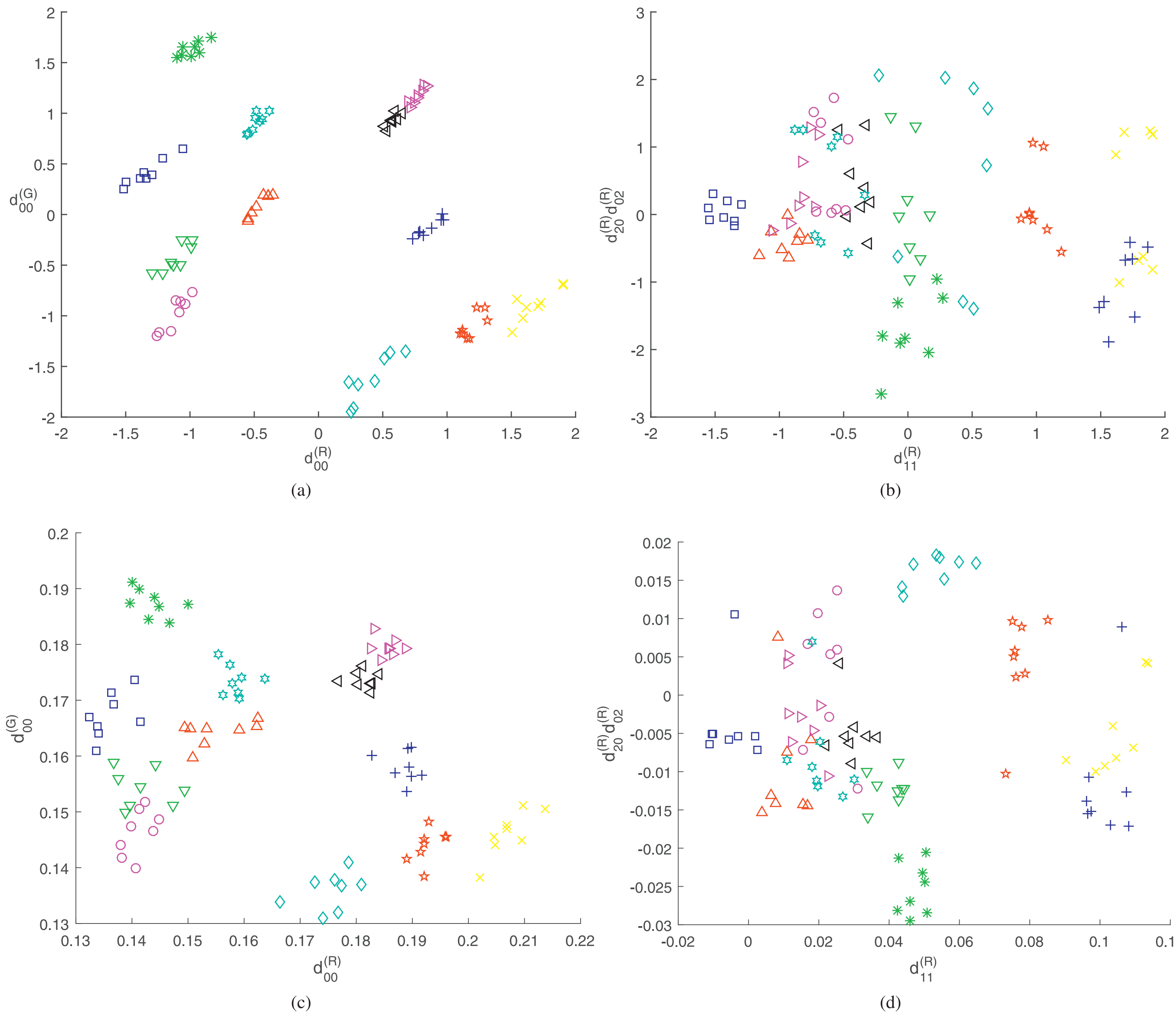


Fig. 3. The card positions in the space of two invariants (a) $\Phi_{00}^{(R)}$ and $\Phi_{00}^{(G)}$ (background noise, STD = 0.2), (b) $\Phi_{11}^{(R)}$ and $\Phi_{02}^{(R)}$ (background noise, STD = 0.2), (c) $\Phi_{00}^{(R)}$ and $\Phi_{00}^{(G)}$ (noise in the image only, STD = 2), and (d) $\Phi_{11}^{(R)}$ and $\Phi_{02}^{(R)}$ (noise in the image only, STD = 2). Zero-order invariants are affected by noise less than the second-order ones, especially in case of background noise. Legend: ▽ – Ferdy the Ant 1, △ – Ferdy the Ant 2, □ – Ladybird, ◇ – Poke the Bug, ▷ – Ant-lion 1, ◁ – Ant-lion 2, × – Mole cricket, + – Snail, ☆ – Butterfly, ✳ – Cricket, ✡ – Bumblebee, ○ – Heteropter.

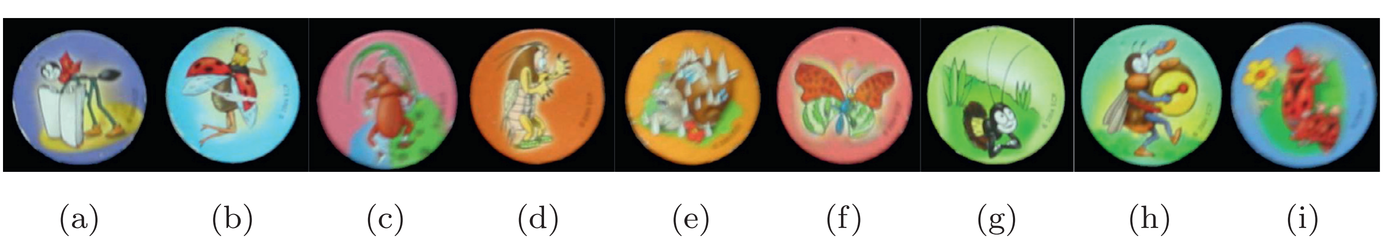


Fig. 4. The other cards used in the experiment: (a) Ferdy the Ant 2, (b) Ladybird, (c) Poke the Bug, (d) Mole cricket, (e) Snail, (f) Butterfly, (g) Cricket, (h) Bumblebee, (i) Heteropter.

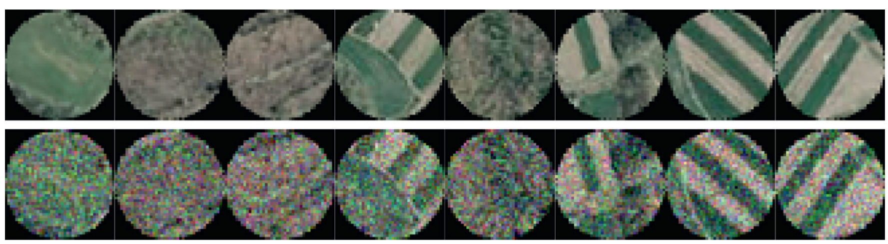
The robustness to noise of joint invariants from JRGB is higher, since the noise is independent in different channels.

We repeated this experiment with the same setting but with the noise added only to the card pictures themselves, not to the black background. In this case, the standard deviation of the noise was two, which is ten times more than before, see Fig. 1(c) and (f). The results are in Table 2. If we use the newly proposed features, we reach correct results even if the noise is so heavy that people cannot recognize the cards visually.

In Fig. 3 we can see that the zero-order invariants create much more compact clusters than the second-order ones, especially in the case of background noise. It illustrates the conclusion that the

(a) (b)

Fig. 5. The aerial image of the size of 257×257 pixels. In (b), the white circles show some extracted templates.



(a) (b) (c) (d) (e) (f) (g) (h)

Fig. 6. Sample extracted and rotated templates without noise (top row) and noisy (bottom row). The diameter of the template is 41 pixels. Only 8 templates out of 100 are shown.

Table 2
The number of errors out of 77 trials in the experiment with "Ferdy the Ant" cards with Gaussian noise in the card picture only.

Order	Gray	RGB	SRGB	JSRGB	JRGB
3	49	6	0	0	0
4	40	1	0	0	0
5	45	4	0	0	0
6	40	3	0	0	0

Table 3
The number of errors out of 77 trials in the experiment with "Ferdy the Ant" cards with Poisson noise, $p = 0.05$.

Order	Gray	RGB	SRGB	JSRGB	JRGB
3	63	25	2	1	4
4	58	17	2	2	3
5	54	24	4	3	4
6	51	15	4	4	5

Table 4
The number of errors out of 77 trials in the experiment with "Ferdy the Ant" cards with Poisson noise, $p = 0.1$.

Order	Gray	RGB	SRGB	JSRGB	JRGB
3	54	14	3	1	0
4	46	3	0	0	0
5	47	5	0	0	0
6	42	4	0	0	0

sets of invariants that include zero-order moments, can improve reliability and noise robustness of recognition of color images.

To test robustness to other kinds of noise, we repeated the experiment again with Poisson noise. In digital imaging, Poisson noise occurs as shot noise in photon counting on the chip, where it is associated with the particle nature of light. It can be modelled by an array of random variables with Poisson distribution $P(\lambda)$. Poisson distribution has a single parameter λ, which determines both mean value and variance. For $\lambda > 10$, Poisson distribution is close to normal distribution $N(\lambda, \sqrt{\lambda})$. Poisson noise is neither additive nor signal-independent. To simulate it, we replace each image pixel by randomly generated sample from $P(h)$, where h is the pixel intensity. Hence, Poisson noise does not have any independent parameter that could be changed analogously to the variance of Gaussian noise. However, we can control the SNR by scaling of the image intensities before the noise has been applied and inverse scaling afterwards, which leads to the noisy image with the intensities $P(ph)/p$ where p is the scaling factor. The noise was applied on the cards only, not on the background. The results are summarized in Table 3, where $p = 0.05$ was used to obtain heavy noise and in Table 4, where the image was scaled with $p = 0.1$ to get moderate noise (see Fig. 2 for examples of noisy images). Even if the noise perceived as very heavy, the overall number of misclassifications is low. Comparing different sets among themselves clearly shows that the color invariants (all three sets) are also in this case significantly better than the single-channel RGB invariants and graylevel invariants.

4.2. Template matching

We tested the invariants in a template matching experiment, which is one of the most frequent applications of the invariants in practice. We downloaded an aerial image from the website

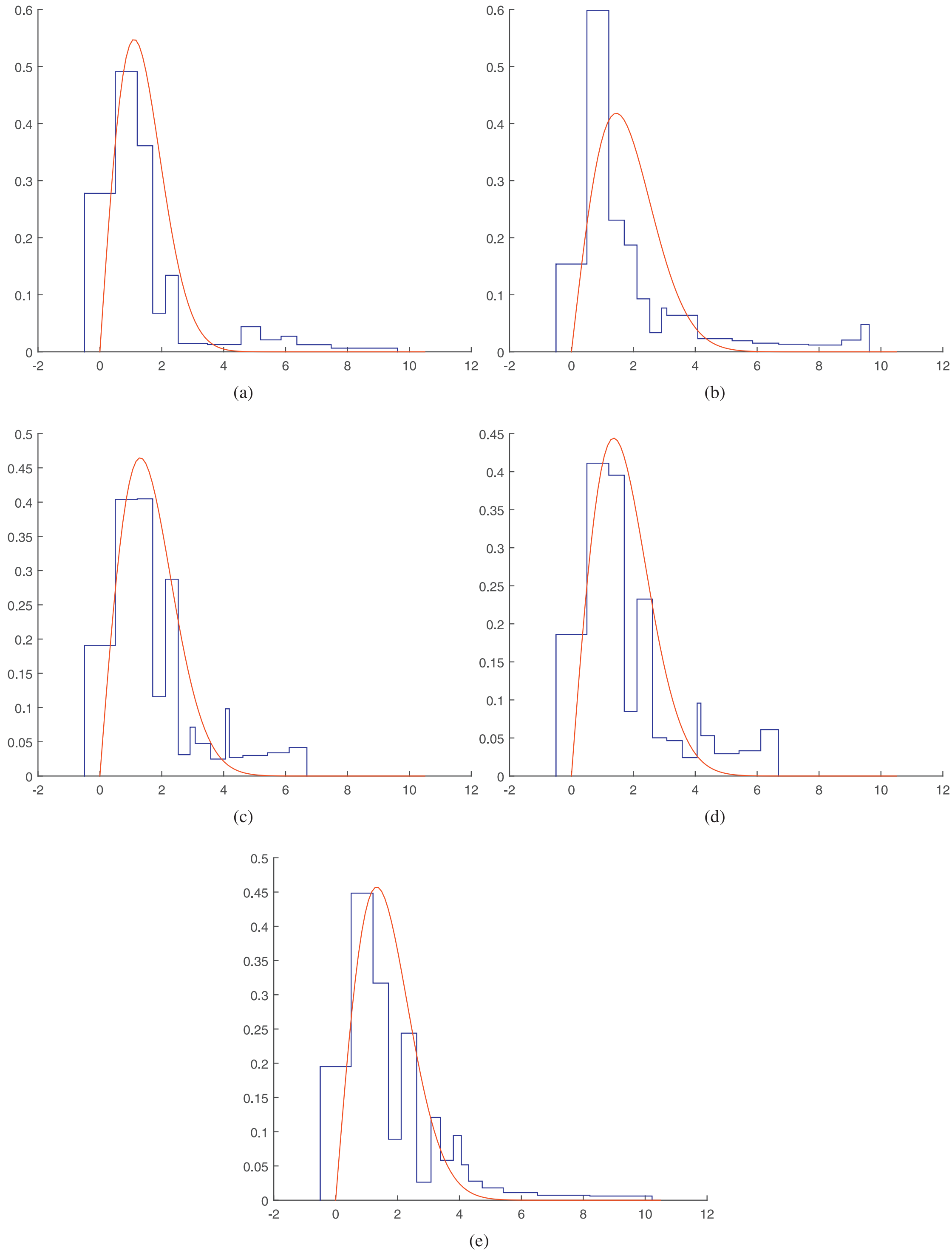


Fig. 7. The histogram (bar graph) and the fitted Rayleigh distribution (the curve) of the small error magnitudes for the invariant sets: (a) gray, (b) RGB, (c) SRGB, (d) JSRGB, and (e) JRGB.

[42] and randomly selected 100 circular templates (see Fig. 5(a) for the original scene and Fig. 5(b) for some template positions). The templates were rotated by a random angle with a uniform distribution between 0° and 360° and corrupted by additive Gaussian white noise of standard deviation 0.125, which resulted in SNR of the templates being from −5.7 dB to 3.1 dB. A few selected templates before and after adding the noise are depicted in Fig. 6. The aim of this experiment is to locate the noisy templates in the original image.

We employed the TR invariants from Section 3.4 up to the fourth order. The matching position was determined as the minimum ℓ_2 distance in the space of the invariants (simple full search through the entire image was applied, we did not use any iterative hierarchical approach).

We encountered two different kinds of errors – "small" errors up to 10 pixels, which can tell as how accurate the matching is, and "big" errors that mean a total mismatch. The particular values of these big errors does not say anything meaningful about

Fig. 8. Segmented national flags. These images were used as the class representatives. (a) Italy, (b) Mali, (c) Chad, (d) France, (e) Guinea, (f) Ireland, (g) Romania, (h) Algeria, (i) Turkey, (j) Pakistan, (k) Mauritania, (l) Tunisia, (m) Finland, (n) Denmark, (o) Sweden, (p) Iceland, (q) Norway, (r) Zimbabwe, (s) São Tomé and Príncipe, (t) Cuba, (u) Bahamas, (v) Jordan, (w) Philippines, (x) Sierra Leone, (y) Russia, (z) Yemen, (aa) Austria, (ab) Bulgaria, (ac) Hungary, (ad) Netherlands, (ae) Tuvalu, (af) Fiji, (ag) New Zealand, (ah) Australia, (ai) Uzbekistan, (aj) Tajikistan, (ak) Slovenia, (al) Croatia, (am) Azerbaijan, (an) Argentina, (ao) Honduras, (ap) Ethiopia, (aq) Afghanistan, (ar) Andorra, (as) Cameroon, (at) Guatemala, (au) Senegal, (av) Somalia, (aw) Morocco, (ax) Vietnam, (ay) China, (az) Albania, (ba) Lesotho, (bb) Slovakia, (bc) India, (bd) Venezuela, (be) Syria, (bf) Bolivia, (bg) Ghana, (bh) Niger, (bi) Egypt, (bj) Czechia, (bk) South Africa, and (bl) Grenada.

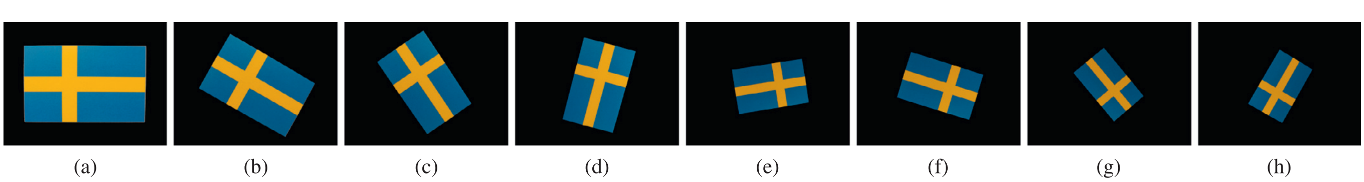
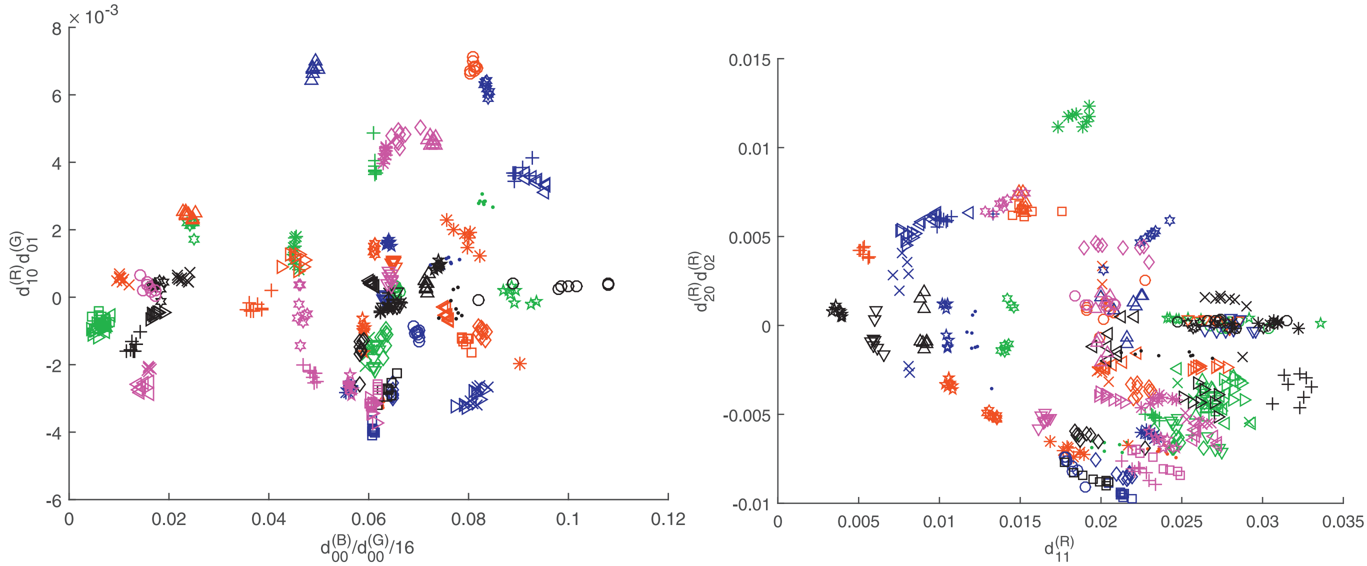


Fig. 9. Eight snaps of the Swedish flag – an example of the test images.



(a) (b)

Fig. 10. Flag positions in the space of two invariants (a) $\Phi_{00}^{(SB)}$ and $\Phi_{01}^{(J1)}$ and (b) $\Phi_{11}^{(R)}$ and $\Phi_{02}^{(R)}$. Legend: ▽ Italy, △ Mali, □ Chad, ◇ France, ▷ Guinea, × Ireland, ◁ Romania, + Algeria, ☆ Turkey, ✳ Pakistan, ✡ Mauritania, ○ Tunisia, • Finland, ▽ Denmark, △ Sweden, □ Iceland, ◇ Norway, ▷ Zimbabwe, × São Tomé and Príncipe, ◁ Cuba, + Bahamas, ☆ Jordan, ✳ Philippines, ✡ Sierra Leone, ○ Russia, • Yemen, ▽ Austria, △ Bulgaria, □ Hungary, ◇ Netherlands, ▷ Tuvalu, × Fiji, ◁ New Zealand, + Australia, ☆ Uzbekistan, ✳ Tajikistan, ✡ Slovenia, ○ Croatia, • Azerbaijan, ▽ Argentina, △ Honduras, □ Ethiopia, ◇ Afghanistan, ▷ Andorra, × Cameroon, ◁ Guatemala, + Senegal, ☆ Somalia, ✳ Morocco, ✡ Vietnam, ○ China, • Albania, ▽ Lesotho, △ Slovakia, □ India, ◇ Venezuela, ▷ Syria, × Bolivia, ◁ Ghana, + Niger, ☆ Egypt, ✳ Czechia, ✡ South Africa, and ○ Grenada.

Table 5
The error statistics in the template matching experiment. NBE – the number of "big" errors, MME – the mean magnitude of "small" errors in pixels.

	Gray	RGB	SRGB	JSRGB	JRGB
NBE	28	22	16	14	18
MME	1.39	1.82	1.64	1.71	1.66

Table 6
Flag recognition – the number of errors out of 512 trials.

Order	Gray	RGB	SRGB	JSRGB	JRGB
3	223	87	29	29	26
4	147	63	11	11	15
5	110	41	9	9	14
6	116	43	10	10	13

the accuracy because the error of 50 pixels is of the same practical significance as the error of 200 pixels for example. To evaluate the "big" errors, we only counted their number without considering their actual value. The "small" errors are caused by random measurement errors in both coordinates. Assuming they are independent and normally distributed in x and y, their magnitudes underlay the Rayleigh distribution

$$R(x;\sigma) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}. \quad (16)$$

The mean value of the magnitude of these "small" errors, along with the number of "big" errors, provide a good measure of the matching accuracy. The results are summarized in Table 5.

The strength of the color invariants is in the less number of big errors. Apparently, the color invariants are more stable to noise than gray and RGB invariants. On the other hand, the invariants JRGB, JSRGB, and SRGB are comparable. The distribution of "small" errors is similar for all the methods (see Fig. 7; the red curve is the fitted Rayleigh probability density function), but note that for gray and RGB some "small" errors turned to the "big" ones.

4.3. Flag recognition

This experiment is similar to that one with the "pick a pair" cards, but here we used pictures of various national flags. We chose 64 flags such that there were groups of flags visually very similar to one another, see Fig. 8. We intentionally changed not only the orientation of the camera, but also the distance of the camera from the flag to obtain snaps of different scales. Each flag was photographed eight times with different rotation and scale, see Fig. 9 for an example.

First, we tested traditional TRS invariants for graylevel images computed from snaps converted to gray levels (see "gray" in the table). Then, we computed the same features from each color channel separately (RGB). Finally, we experimented with three sets of invariants proposed for color images under TRS transformation: SRGB, JSRGB, and JRGB, see Section 3.5. The results are summarized in Table 6, the graphs of selected invariants are in Fig. 10.

It is not surprising that the invariants applied to graylevels only yielded poor results. Many flags differs from each other namely by colors, while their texture (stripes, stars, etc.) are similar. Whenever the color information was used, the number of misclassifications decreased. The distribution of misclassifications is not uniform. The following six flags are responsible for significantly more errors than the others: the flags of Chad (Fig. 8(c)) and Romania (Fig. 8(g)) differ from one another only by the precise tint of the blue color; the flags of Mali (Fig. 8(b)) and Guinea (Fig. 8(e)) differ from one another only by a left-right flip so they cannot be distinguished by rotation invariants; and the flags of New Zealand (Fig. 8(ag)) and Australia (Fig. 8(ah)) differ only slightly by the color and the number of the stars.

The proposed invariants significantly reduced the numbers of errors, while there is no difference between the single-channel invariants (SRGB) and the single-channel invariants with low-order joint ones (JSRGB). The joint invariants (JRGB) were slightly worse in this case.

5. Conclusions

Gaussian–Hermite rotation invariants for color images are proposed in this paper. This new kind of invariants is based on the existing GHIs for graylevel images, to which new invariants were added. The experiments with image classification and template matching demonstrated their superior performance comparing to graylevel and concatenated single-channel invariants.

To complete the set of TR and TRS color invariants, we studied two approaches which independent invariants should be added to a union of single-channel invariants applied on RGB. First, we added low-order moments which had not been used before and then we created joint invariants. Both approaches allowed us to construct complete sets, so in this sense they are theoretically equivalent. The experiments confirmed there is no significant difference between them in terms of their recognition power. It is an interesting conclusion because in case of color affine invariants studied earlier in [39], the use of joint invariants is inevitable.

Acknowledgment

This project has been financially supported by the National Natural Science Foundation of China (Grant No. 61502389), the Fundamental Research Funds for the Central Universities (Grant No. 3102015ZY047) as well as the Natural Science Basis Research Plan in Shaanxi Province of China (Grant No. 2015JM1013). The research has been supported also by the Czech Science Foundation (GAČR) under the grant No. GA15-16928S.

References

[1] J. Flusser, T. Suk, B. Zitová, 2D and 3D Image Analysis by Moments, Wiley, Chichester, U.K., 2016.
[2] J. Flusser, T. Suk, B. Zitová, Moments and Moment Invariants in Pattern Recognition, Wiley, Chichester, U.K., 2009.
[3] M.K. Hu, Visual pattern recognition by moment invariants, IRE Trans. Inf. Theory 8 (1962) 179–187.
[4] J. Flusser, On the independence of rotation moment invariants, Pattern Recognit. 33 (2000) 1405–1410.
[5] A. Khotanzad, Y.H. Hong, Invariant image recognition by zernike moments, IEEE Trans. Pattern Anal. Mach.Intell. 12 (1990) 489–497.
[6] C.W. Chong, P. Raveendran, R. Mukundan, Translation and scale invariants of Legendre moments, Pattern Recognit. 37 (2004) 119–129.
[7] K.M. Hosny, New set of rotationally Legendre moment invariants, Int. J. Electr., Comput.Syst. Eng. 4 (2010) 176–180.
[8] H.Q. Zhu, H.Z. Shu, T. Xia, L.M. Luo, J.L. Coatrieux, Translation and scale invariants of Tchebichef moments, Pattern Recognit. 40 (2007) 2530–2542.
[9] B. Xiao, Image analysis by Bessel–Fourier moments, Pattern Recognit. 43 (2010) 2620–2629.
[10] B. Yang, G.X. Li, H.L. Zhang, M. Dai, Rotation and translation invariants of Gaussian–Hermite moments, Pattern Recognit. Lett. 32 (2011) 1283–1298.
[11] A. Khotanzad, Y.H. Hong, Image analysis via the general theory of moments, J. Opt. Soc. Am. 70 (1980) 920–930.
[12] A. Wallin, O. Kubler, Complete sets of complex Zernike moment and the role of the pseudo invariants, IEEE Transa. Pattern Anal. Mach. Intell. 17 (1995) 1106–1110.
[13] S.O. Belkasim, M. Shridhar, M. Ahmadi, Pattern recognition with moment invariants: a comparative study and new results, Pattern Recognit. 24 (12) (1991) 1117–1138.
[14] Y. Sheng, H.H. Arsenault, Experiments on pattern recognition using invariant Fourier-Mellin descriptors, J. Opt. Soc Am. A, Opt. image sci. 3 (6) (1986) 771–776.
[15] Y. Sheng, J. Duvernoy, Circular-Fourier-radial-Mellin transform descriptors for pattern recognition, J. Opt. Soc. Am. A 3 (6) (1986) 885–888.
[16] Y. Sheng, L. Shen, Orthogonal Fourier–Mellin moments for invariant pattern recognition, J. Opt. Soc. Am. A 11 (1994) 1748–1757.
[17] Z. Ping, H. Ren, J. Zou, Y. Sheng, W. Bo, Generic orthogonal moments: Jacobi-Fourier moments for invariant image description, Pattern Recognit. 40 (4) (2007) 1245–1254.
[18] Z. Ping, J. Wu, Y. Sheng, Image description with Chebyshev-Fourier moments, J. Opt. Soc. Am. A 19 (9) (2002) 1748–1754.
[19] C.-H. Teh, R.T. Chin, On image analysis by the method of moments, IEEE Trans. Pattern Anal. Mach.Intell. 10 (4) (1988) 496–513.
[20] M. Pawlak, Image analysis by moments: reconstruction and computational aspects, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław (2006). Poland.
[21] R. Mukundan, S.H. Ong, P.A. Lee, Image analysis by Tchebichef moments, IEEE Trans. Image Process. 10 (9) (2001) 1357–1364.
[22] R. Mukundan, Some computational aspects of discrete orthonormal moments, IEEE Trans. Image Process. 13 (8) (2004) 1055–1059.
[23] J. Flusser, J. Kautsky, F. Šroubek, Implicit moment invariants, International Journal of Computer Vision 86 (1) (2010) 72–86.
[24] J. Shen, Orthogonal Gaussian-Hermite moments for image characterization, in: Proc. SPIE on Intelligent Robots and Computer Vision, Pittsburgh, USA, 1997, pp. 224–233.
[25] J. Shen, W. Shen, D.F. Shen, On geometric and orthogonal moments, Int. J. Pattern Recognit.Artif. Intell. 14 (2000) 875–894.
[26] P. Yap, R. Paramesran, S. Ong, Image analysis by Krawtchouk moments, IEEE Trans. Image Process. 12 (2003) 1367–1377.
[27] A. Chiang, S. Liao, Q. Lu, M. Pawlak, W. Kinsner, Gegenbauer moment-based applications for Chinese character recognition, in: A. Sebak, K. Ferens (Eds.), Proceedings of the Canadian Conference on Electrical Computer Engineering CCECE'02,IEEE, 2, 2002, pp. 908–911.
[28] S.X. Liao, A. Chiang, Q. Lu, M. Pawlak, Chinese character recognition via Gegenbauer moments, in: Proceedings of the 16th International Conference on Pattern Recognition ICPR'02, 3, IEEE Computer Society, 2002, pp. 485–488.
[29] B. Yang, M. Dai, Image analysis by Gaussian – Hermite moments, Signal Process., 91, (10) 2011, pp. 2290–2303.
[30] B. Yang, J. Flusser, T. Suk, Design of high-order rotation invariants from Gaussian-Hermite moments, Signal Process. 113 (1) (2015) 61–67.
[31] B. Yang, J. Flusser, T. Suk, 3D rotation invariants of Gaussian–Hermite moments, Pattern Recognit. Lett. 54 (1) (2015) 18–26.
[32] B. Yang, J. Kostková, J. Flusser, T. Suk, Scale invariants from Gaussian–Hermite moments, Signal Process. 132 (2017) 77–84.
[33] Y. Wu, J. Shen, M. Dai, Traffic object detections and its action analysis, Pattern Recognit. Lett. 26 (2005) 1963–1984.
[34] X. Ma, R. Pan, L. Wang, License plate character recognition based on Gaussian-Hermite moments, in: IEEE Second Workshop on Education Technology and Computer Science ETCS'10, Shanghai, China, 2010, pp. 11–14.
[35] L. Wang, M. Dai, Application of a new type of singular points in fingerprint classification, Pattern Recognit. Lett. 28 (2007) 1640–1650.
[36] L. Ma, T. Tan, Y. Wang, Local intensity variation analysis for iris recognition, Pattern Recognit. 37 (2004) 1287–1298.
[37] S. Farokhi, U. Sheikh, J. Flusser, B. Yang, Near infrared face recognition using Zernike moments and Hermite kernels, Inf. Sci. 316 (2015) 234–245.
[38] T.S. Chihara, An Introduction to Orthogonal Polynomials, Gordon and Breach, New York, USA, 1978.
[39] T. Suk, J. Flusser, Affine moment invariants of color images, in: Computer Analysis of Images and Patterns CAIP 2009, Münster, Germany, 2009, pp. 334–341.
[40] F. Mindru, T. Tuytelaars, L.V. Gool, T. Moons, Moment invariants for recognition under changing viewpoint and illumination, Comput. Vision Image Understanding 94 (2004) 3–27.
[41] L. Guo, M. Dai, M. Zhu, Quaternion moment and its invariants for color object classification, Inf. Sci. 273 (2014) 132–143.
[42] C. Orange, Aerial image sample, 2015. http://www.orangecountync.gov/departments/deapr/soil_and_water_conservation_district/aerial_photography.php.