On the Null-Space of the Shape-Color Moment Invariants

Jitka Kostková(✉) and Jan Flusser

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

Abstract. In this paper, we extend the theory of the combined Shape-Color Affine Moment Invariants (SCAMIs) for recognition of color images, proposed originally by Gong et al. in [3]. Since in the real pictures the shape deformation is always accompanied by the color deformation, it is not sufficient to use the shape invariant or color invariant descriptors only and the use of combined invariants is needed. However, the SCAMIs are not able to recognize images, the color channels of which are linearly dependent or highly correlated. This situation is not rare in practice and is of particular importance in hyper-spectral image analysis, where the spectral bands are highly correlated. We analyze why the SCAMIs fail in such situations, correct the theory and propose a solution to overcome such drawback. Unlike the SCAMIs, the new invariants have in the null-space the constant-zero images only, which leads to a better discrimination power, as demonstrated also on various pictures.

Keywords: Color image recognition · Affine transformation · Channel mixing · Moment invariants · Null-space

1 Introduction

Image descriptors, which are invariant with respect to certain group of spatial and/or color transformations, have been a topic of much research in image analysis. In the recent paper [3], Gong et al. proposed invariants of 2D color images, represented by their RGB channels, with respect to affine transformation of the spatial coordinates and simultaneously to affine transformation of the RGB space. Their invariants are based on classical theory of moments (see, for instance, [1,2] for a survey). The paper [3] was on one hand inspired by the work of Suk and Flusser [7], who proposed a general framework based on graph theory, allowing a systematic construction of affine moment invariants (AMIs) of graylevel images. Suk and Flusser extended their theory to color images, too, by incorporating a between-channel bond and introducing so-called joint affine

This work was supported by the Czech Science Foundation (Grant No. GA18-07247S), by the *Praemium Academiae*, and by the Grant Agency of the Czech Technical University (Grant No. SGS18/188/OHK4/3T/14).

M. Vento and G. Percannella (Eds.): CAIP 2019, LNCS 11678, pp. 402–408, 2019.
https://doi.org/10.1007/978-3-030-29888-3_32

invariants [6]. On the other hand, another motivation came from the work of Mindru et al. [4,5], where also the color changes (called "photometric changes" in their work) of RGB images were considered along with spatial transformations. Gong et al. put these two approaches together but also added a significant novelty. While in [7] and [6] only a spatial affine transform is considered and in [5] the main part is devoted to affine transformation of the RGB space with no transformation of the coordinates, Gong et al. [3] considered for the first time these two transformation groups acting jointly (actually, such joint model was mentioned already in [5] but was considered to be too complicated and was left without any invariants being derived.) Gong et al. succeeded in derivation of *shape-color affine moment invariants* (SCAMIs). However, they missed an important issue. Their SCAMIs do not form a complete set of invariants and their recognition power is limited. There exists a relatively large class of images, totally different from each other visually, which all lie in the joint null-space of the SCAMIs and cannot be distinguished by means of them.

In this paper, we analyze why this happens, show what images are in the SCAMIs null-space, and propose an extension of the Gong's theory which makes the invariants complete. The null-space of the new invariants contains just a constant zero image. Simple experiments demonstrate the advantages of the extended SCAMIs (ESCAMIs).

2 Recalling the SCAMIs

To show where is the weak point of the paper [3] and to explain how it can be corrected, we have to start with a very brief summary of the main idea of [3].

Gong et al. assumed the following transformation model

$$\mathbf{c}'(\mathbf{x}) = M\mathbf{c}(A\mathbf{x}) \tag{1}$$

where $\mathbf{x} = (x, y)^T$, $\mathbf{c}(\mathbf{x}) = (R(\mathbf{x}), G(\mathbf{x}), B(\mathbf{x}))^T$ is a color RGB image, A is a regular 2×2 matrix of affine spatial transformation, and M is a regular 3×3 matrix describing a linear color mixing. Gong et al. claimed this model describes the change of viewpoint and illumination in outdoor scenes (for indoor scenes, M is supposed to have a diagonal form, which brings us back to the work by Mindru et al. [5]). For the model (1), it is reasonable to try to construct invariants, because this transformation forms a group which is known to be a necessary condition for the existence of invariants [2].

The problem of finding invariants w.r.t. A was treated correctly in [3] (the authors followed the theory from [7]) and no modification is necessary. The null-space problem is caused solely by the way how the invariance w.r.t. M is achieved. To focus on that, we will from now on assume that $A = I$. This simplifies the reasoning but all our conclusions stay valid for arbitrary regular matrix A.

Consider arbitrary three points $\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k$ and construct the determinant

$$V(i,j,k) \equiv |C_{ijk}| = \begin{vmatrix} R(\mathbf{x}_i) & R(\mathbf{x}_j) & R(\mathbf{x}_k) \\ G(\mathbf{x}_i) & G(\mathbf{x}_j) & G(\mathbf{x}_k) \\ B(\mathbf{x}_i) & B(\mathbf{x}_j) & B(\mathbf{x}_k) \end{vmatrix}. \tag{2}$$

In the color space, $V(i,j,k)$ can be understood as a volume of the pyramid with vertices in the three points $(R(\mathbf{x}_i), G(\mathbf{x}_i), B(\mathbf{x}_i))$, $(R(\mathbf{x}_j), G(\mathbf{x}_j), B(\mathbf{x}_j))$, and $(R(\mathbf{x}_k), G(\mathbf{x}_k), B(\mathbf{x}_k))$, respectively, and in the origin $(0,0,0)$. $V(i,j,k)$ is a relative invariant w.r.t. M. If $\mathbf{c}'(\mathbf{x}) = M\mathbf{c}(\mathbf{x})$, then $V'(i,j,k) = |M|V(i,j,k)$.

If we use $n > 3$ points $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, we can construct m point triplets (m could be arbitrary high because the same triplet can be used many times) and multiply the corresponding determinants as

$$E(n,m) = V(1,2,3)V(g,h,i)\cdots V(r,s,n). \tag{3}$$

$E(n,m)$ is also a relative invariant w.r.t. M since $E'(n,m) = |M|^m E(n,m)$.

Gong proposed to multiply $E(n,m)$ by so-called *shape primitives* that eliminate the influence of A. Since here we assume $A = I$, the shape primitives are just powers of spatial coordinates. Finally, the product is integrated n-times over the entire image plane

$$E_{\mathbf{p},\mathbf{q}} = \int\int\cdots\int x_1^{p_1} y_1^{q_1} \cdots x_n^{p_n} y_n^{q_n} E(n,m)\, \mathbf{dx}_1 \cdots \mathbf{dx}_n. \tag{4}$$

Doing so, they obtained relative invariant $E'_{\mathbf{p},\mathbf{q}} = |M|^m E_{\mathbf{p},\mathbf{q}}$ which depends on the number of chosen points and on the selected triplets but does not depend on particular positions of the points $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$. It is worth noting that $E_{\mathbf{p},\mathbf{q}}$ can be expressed in terms of *generalized color moments* μ_{pq} which are defined as

$$\mu_{pq}^{abc} = \int\int x^p y^q R^a(x,y) G^b(x,y) B^c(x,y)\, \mathrm{d}x\, \mathrm{d}y. \tag{5}$$

That is why we can consider $E_{\mathbf{p},\mathbf{q}}$ to be a member of a family of *moment invariants.*

Absolute invariants are then obtained by an obvious normalization as a ratio of two appropriate relative invariants. At this point, the authors of [3] considered the theory complete.

3 Where Is the Problem?

Let us imagine what happens if matrix C_{ijk} is singular for any triplet (i,j,k) used in Eq. (3). (Note that the rank of C_{ijk} is in no way linked with the rank of M, which is assumed to be regular, and that the rank of C_{ijk} does not change if the image undergoes transformation (1).) In that case, $V(i,j,k) = 0$ for arbitrary (i,j,k). Consequently, any $E(n,m)$ vanishes and also $E_{\mathbf{p},\mathbf{q}} = 0$ regardless of $n, \mathbf{p}, \mathbf{q}$ and of the choice of the triplets.

Let us explain what the singularity of C_{ijk} means for the image $\mathbf{c}(\mathbf{x})$. It happens if and only if its R,G,B channels are linearly dependent. Hence, the null-space of the SCAMIs is formed by all images with linearly dependent channels. This conclusion is not violated even if we admit an arbitrary regular A, as is clearly apparent from Eq. (26) of [3]. In other words, images with linearly dependent channels cannot be distinguished by the SCAMIs, although they may be visually quite different (see Fig. 1 for some examples) and easy-to-distinguish by some other methods. Since the numerical experiments in [3] apparently did not included such images, this phenomenon has remained unrevealed.

4 The Proposed Solution

Fortunately, there is an easy and elegant solution to the problem. We simply disregard one color channel (for instance B) and work with R and G only. Instead of the triplets, we use only pairs of points. Instead of $V(i,j,k)$ we have

$$V(i,j) \equiv |C_{ij}| = \begin{vmatrix} R(\mathbf{x}_i) & R(\mathbf{x}_j) \\ G(\mathbf{x}_i) & G(\mathbf{x}_j) \end{vmatrix}, \tag{6}$$

which is again a relative invariant. We proceed to the construction of $E(n,m)$ and $E_{\mathbf{p},\mathbf{q}}$ similarly to the previous case.

If the rank of C_{ijk} equals two, then C_{ij} is regular (at least for some point configurations, which is sufficient for $E_{\mathbf{p},\mathbf{q}}$ to be non-zero). In the joint null-space of these invariants, we find all images whose rank of C_{ijk} equals one (see Fig. 2 for some examples). These images are still not distinguishable. To distinguish among them, we have to create yet another set of invariants.

It is sufficient to consider just a single color channel (the other two are its linear transformations). There is in fact no true "channel mixing", matrix M (even if it is not diagonal) only makes a contrast stretching of the channels. We can treat the selected channel as a graylevel image. We take graylevel affine moment invariants from [7] and normalize them to contrast stretching. This completes our solution. In the joint null-space of the extended invariants we can find just a constant-zero image.

When designing an image recognition system, we of course need not to check for each image the rank of C_{ijk} for all possible point triplets. That would be extremely time consuming. We can start with the calculation of SCAMIs. If at least one of them is non-zero, we do not calculate any other invariants and use the SCAMIs directly. If all of them are zero (in practice, the term "all of them" means up to our maximum order and the term "they are zero" means they are within a user-defined interval around zero), then we remove one color band and calculate the invariants from two other color bands. If they all are still zero, we take a single channel and calculate the graylevel AMIs. This extended set, which we call ESCAMIs, can discriminate any two images (modulo the action of the transformation group given by Eq. (1)).

5 Discussion

The null-space problem is a general problem of any features designed for object recognition. The images laying in the null-space are in principle not distinguishable by these features because all features are zero. In general, the smaller null-space the better recognition power. In an ideal case, the null-space should contain only images equivalent (modulo the considered transformation group) to the zero image. The reasons why an image falls into the null-space depend on the nature of the features. We showed that in the case of SCAMIs, the necessary and sufficient condition is a linear dependence among the color channels.

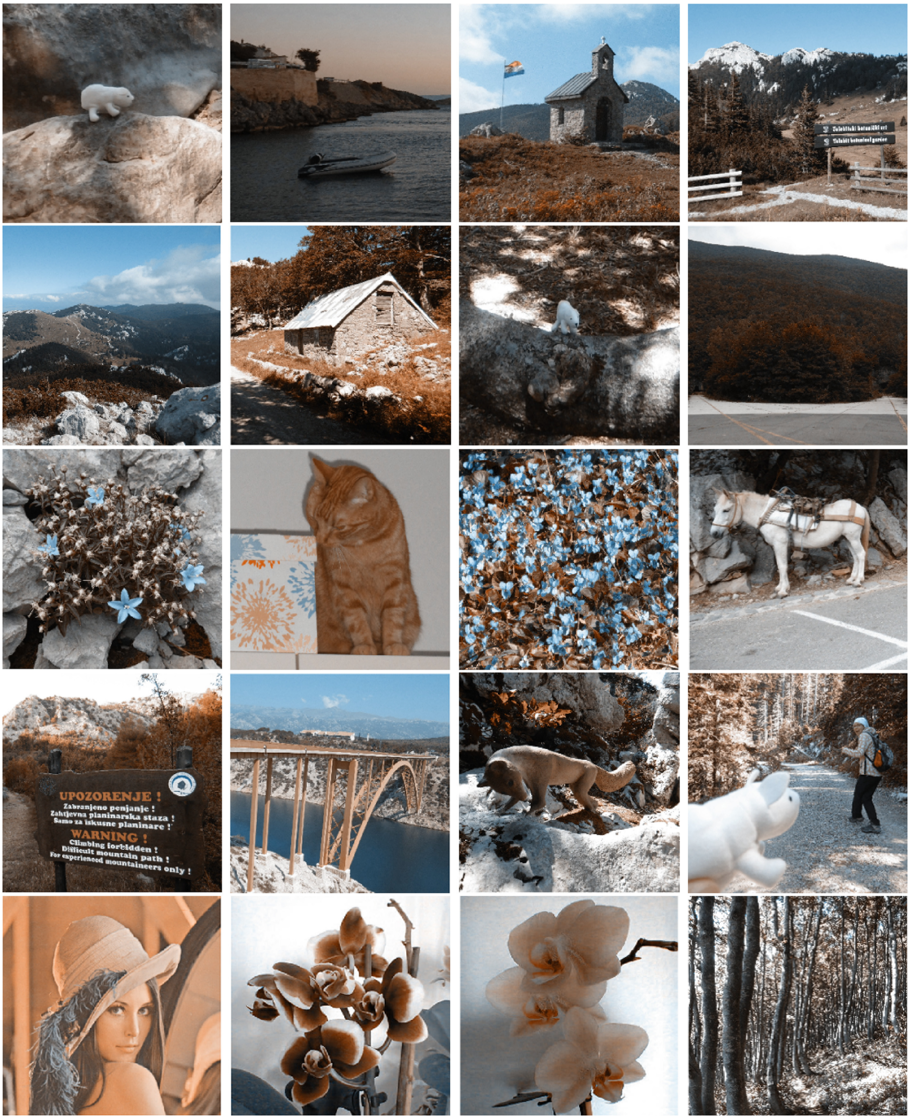

Fig. 1. Examples of images that are not distinguishable by the SCAMIs from [3]. The rank of C_{ijk} equals two in all cases. The images are clearly distinguishable by the proposed ESCAMIs using two channels. This figure should be viewed in colors.

One might feel that such images are rare in practice but it is not the case. Color bands of many real images are not exactly dependent but are highly correlated, which leads to "almost vanishing" SCAMIs. The SCAMI values are not zero but they are comparable to noise. This effect becomes more apparent as the number of color bands increases. All the theory can be easily modified for hyper-

Fig. 2. Examples of images that are neither distinguishable by the original SCAMIs nor by ESCAMIs from Eq. (6), but are clearly distinguishable by the grayscale AMIs [7] applied to a single channel. The rank of C_{ijk} equals one in all cases. This figure should be viewed in colors.

spectral images with arbitrary number of bands (current hyperspectral sensors use to have several hundreds of them). Those bands are very narrow and close to each other in terms of the wavelength. Hence, they are highly correlated and matrix C is quite often close to be singular. Without the extension presented in this paper, application of SCAMIs on hyperspectral images would be useless.

Another class of images, where the SCAMIs vanish completely, are some artificially colored pictures, that often contain linearly dependent channels.

6 Conclusion

In this paper, we extended the theory of the SCAMI invariants for recognition of color images, presented originally in [3]. The proposed solution not only corrects the theory from [3] but it is also important for practical image classification.

Our extension can handle a special case of images with linearly dependent or highly correlated channels, for which the original method fails. The efficiency of the proposed extension is in the fact that it not only improves the recognition power but also in its fast calculation. It can be applied directly after the original SCAMIs have been calculated and have vanished. The new method does not require any preliminary testing or pre-classification of the images.

References

1. Flusser, J., Suk, T., Zitová, B.: Moments and Moment Invariants in Pattern Recognition. Wiley, Chichester (2009)
2. Flusser, J., Suk, T., Zitová, B.: 2D and 3D Image Analysis by Moments. Wiley, Chichester (2016)
3. Gong, M., Hao, Y., Mo, H., Li, H.: Naturally combined shape-color moment invariants under affine transformations. Comput. Vis. Image Underst. **162**(Suppl. C), 46–56 (2017). https://doi.org/10.1016/j.cviu.2017.07.003
4. Mindru, F., Moons, T., Gool, L.V.: Color-based moment invariants for viewpoint and illumination independent recognition of planar color patterns. In: Singh, S. (ed.) International Conference on Advances in Pattern Recognition ICAPR 1999, pp. 113–122. Springer, London (1999). https://doi.org/10.1007/978-1-4471-0833-7_12
5. Mindru, F., Tuytelaars, T., Gool, L.V., Moons, T.: Moment invariants for recognition under changing viewpoint and illumination. Comput. Vis. Image Underst. **94**(1–3), 3–27 (2004)
6. Suk, T., Flusser, J.: Affine moment invariants of color images. In: Jiang, X., Petkov, N. (eds.) CAIP 2009. LNCS, vol. 5702, pp. 334–341. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-03767-2_41
7. Suk, T., Flusser, J.: Affine moment invariants generated by graph method. Pattern Recogn. **44**(9), 2047–2056 (2011)