Moment invariants for recognition under changing viewpoint and illumination

Florica Mindru, a Tinne Tuytelaars, a,* Luc Van Gool, a,b and Theo Moons c

a Katholieke Universiteit Leuven, ESAT–PSI, Leuven, Belgium
b Swiss Federal Institute of Technology, ETH–BIWI, Zürich, Switzerland
c Katholieke Universiteit Brussel, Brussel, Belgium

Received 1 December 2002; accepted 29 October 2003

Abstract

Generalised color moments combine shape and color information and put them on an equal footing. Rational expressions of such moments can be designed, that are invariant under both geometric deformations and photometric changes. These generalised color moment invariants are effective features for recognition under changing viewpoint and illumination. The paper gives a systematic overview of such moment invariants for several combinations of deformations and photometric changes. Their validity and potential is corroborated through a series of experiments. Both the cases of indoor and outdoor images are considered, as illumination changes tend to differ between these circumstances. Although the generalised color moment invariants are extracted from planar surface patches, it is argued that invariant neighbourhoods offer a concept through which they can also be used to deal with 3D objects and scenes.

© 2003 Elsevier Inc. All rights reserved.

Keywords: Color; Moment invariants; Recognition; Viewpoint changes; Illumination changes

1. Introduction

This paper deals with the problem of viewpoint and illumination independent recognition of planar colored patterns, like labels, logos, or pictograms. By their nature,
the information of such objects is typically not contained in their outline or frame, but in the intensity content within.

When objects are viewed under different angles and different lighting conditions, their image displays photometric and geometric changes. This means that the image colors are different, and geometric deformations like scaling, rotation, and skewing have to be taken into account. A variety of approaches exist to the problem of identifying the presence of the same object under such photometric and/or geometric changes. One way of proceeding is to estimate the transformations and compensate for their effects. An alternative is deriving invariant features, that is deriving features that do not change under a given set of transformations. The main advantage of using invariants is that they eliminate expensive parameter estimation steps like camera and light source calibration or color constancy algorithms, as well as the need for normalization steps against the transformations involved.

Much research has been put into invariants for planar shapes under geometric deformations and especially into invariants for the shapes' contours [25,26]. For the patterns considered here, the pictorial content usually is too complicated to robustly extract object contours. Also color information has proven very useful, e.g., [4,12,32]. Color histograms often serve as a basis for the illumination independent characterization of the color distribution of the pattern [10,11,14,29]. Color histograms, however, do not exploit the spatial layout of the colors. Hence, vital information may be lost. A good way of including part of this information is to use moments, as described next.

The invariant features presented in this paper are based on generalized color moments. These are a generalization of the traditional moments: they combine powers of the pixel coordinates and the intensities in the different color bands within the same integral. These moments are introduced more formally in Section 3.1. They characterize the shape and the color distribution of the pattern in a uniform manner. By combining such moments, one can obtain moment invariants if the whole pattern undergoes the same transformation and remains completely visible. From a practical point of view, another advantage of using color moments instead of traditional moments is that a larger set of such robust generalized color moments can be extracted, which leads to lower-order and hence more stable invariants. Moments need a closed bounding contour for their computation. But this is relatively easy to provide for the patterns we focus on here, like signs, labels or billboards, as they normally have simple, predefined shapes, such as parallelograms and ellipses. Also, automatic methods of delineating local regions of interest exist, which allow the moment invariants to be used as descriptors for 3D scenes.

Achieving viewpoint and illumination invariance means dealing with a combination of geometric and photometric changes of the patterns. We investigate alternative choices for the geometric and photometric transformations. Invariance is achieved for affine geometric transformations, and perspective transformations are dealt with by normalization. For the photometric transformations we consider several types of linear transformations as models for indoor and outdoor changes.

A systematic classification of generalized color moment invariants is provided. These invariant functions are rational expressions of the generalized color moments.
They are invariant under the combined selected geometric and photometric transformations. The invariants have been obtained through Lie group methods as described in [39]. The combinations of geometric and photometric models are compared in terms of the discriminant power of their invariants and the resulting classification performance.

The structure of the paper is as follows. Section 2 deals with the types of geometric and photometric transformations a planar surface typically undergoes when viewed under different viewpoints and illumination. Section 3 derives the moment invariants corresponding to the considered geometric and photometric transformations. Section 4 discusses the outcome of several recognition experiments based on these invariants, and an extension of their use to 3D scenes. Finally, Section 5 concludes the paper.

2. Geometric and photometric transformations of planar patterns

Two images of a plane taken from different viewpoints are related by a projectivity. In the most general case, the geometric deformations to be considered are projective transformations. When the camera is relatively far from the viewed object, however, the geometric deformations of the pattern can be simplified to affine transformations

\[
\begin{pmatrix}
  x' \\
  y'
\end{pmatrix} = \begin{pmatrix}
  a_{11} & a_{12} \\
  a_{21} & a_{22}
\end{pmatrix} \begin{pmatrix}
  x \\
  y
\end{pmatrix} + \begin{pmatrix}
  b_1 \\
  b_2
\end{pmatrix} = A \begin{pmatrix}
  x \\
  y
\end{pmatrix} + b,
\]

with \( |A| = a_{11}a_{22} - a_{12}a_{21} \neq 0 \).

A model of the photometric transformations describes the way in which the intensities in the red, green, and blue bands \((R, G, B)\) transform between images. These changes are influenced by the scene illumination, the reflective characteristics of the objects, and the camera sensors. Due to the complexity of the problem, physics-based theoretical models for the resulting photometric transformations are difficult to derive for general cases. Modeling the photometric transformations is therefore often performed from a phenomenological point of view. Model fitting on real images is a useful step of verification, as one needs to know how far from reality the assumed models are. Our work focuses on planar matte surfaces, with light sources far from the objects. This implies that the geometry of light reflection is more or less the same for all points. We can therefore consider that all pixels on the surface undergo the same photometric transformation, which is an important assumption when using moments as measurements on images. For this type of surfaces and viewing conditions it is generally agreed that linear models, like in Eqs. (2)–(4), represent a good fit to the photometric transformations. The following notation is used: a color pixel \( \mathbf{p} = (R, G, B)^T \) is transformed into the corresponding color pixel in the second image \( \mathbf{p'} = (R', G', B')^T \).

‘Type D’: diagonal

\[
\begin{pmatrix}
  R' \\
  G' \\
  B'
\end{pmatrix} = \begin{pmatrix}
  s_R & 0 & 0 \\
  0 & s_G & 0 \\
  0 & 0 & s_B
\end{pmatrix} \begin{pmatrix}
  R \\
  G \\
  B
\end{pmatrix}.
\]
The literature is not unanimous about the type of transformations that best fits real photometric changes for different types of scenes and illumination conditions. On the one hand, good results have been reported based on the rather simple diagonal photometric model of Eq. (2) [6,7,12,17], and this model is often used for indoor images, because it seems to provide the best quality–complexity ratio. On the other hand, some experiments suggested the need for more complicated linear transformations, like the affine model in Eq. (4) [5,30,33], especially in the case of outdoor images.

Gros [13] presents statistical model selection tests for a set of real indoor images viewed under a series of internal changes of light (i.e., different intensity or color of the emitted light). His conclusions, based on confidence intervals for the model parameters, indicate that the SO model (Eq. (3)) is a good compromise between complexity and accuracy. In [23] a series of model selection tests are performed for the case of outdoor imagery consisting of several views of billboards taken under different viewing angles and different illumination (natural light). Possible candidates for a transformation model on \((R, G, B)\) color space were investigated and different approaches for the model selection problem were considered. The paper concludes that the affine model (Eq. (4)) is statistically the best explaining model for the photometric changes in these outdoor images.

Given such dependencies on the particular problem at hand, we propose invariants for each of the above mentioned types of photometric transformations, i.e., Eqs. (2)–(4).

Four types of combinations of photometric and geometric changes are considered. Two sets of moment invariants deal with a combination of affine geometric and linear photometric transformations, and two sets contain photometric invariants combined with normalization against geometric (affine or perspective) deformations of the pattern. Section 3.3 gives a more precise description of these cases and a systematic classification of the corresponding moment invariants.

3. Moment based invariant features

A whole strand of research has focussed on moment invariants under different types of geometric and/or photometric changes. A number of contributions are
directly related to our work. The pioneering investigation of moment invariants in pattern recognition is due to Hu [15], where a set of moment invariants for the similarity transformation group (i.e., translation, scaling, and rotation) were developed using the theory of algebraic invariants. Maitra [18] and Abo-Zaid et al. [1] discussed variations of Hu’s metric and geometric moment invariants [15] that are also invariant under global scaling of the intensity. Another direction of research has concentrated on deriving moment invariants under affine geometric transformations [8,27,28].

A series of publications extend the affine moment invariants presented by Flusser and Suk [8]. Among the latest results there are the work of Flusser and Zitova [9] with moment-based features invariant to rotations and changes in contrast (i.e., scaling of intensities), combined with invariance to image convolution with a centrosymmetric point-spread function (PSF) and the invariants to blur (convolution with a centrosymmetric PSF) and affine geometric transformations of Flusser and Suk [31]. The moments used in these papers are complex moments.

Wang and Healey [40] present a method for recognizing planar matte color texture independent of linear illumination changes, and geometric transformations of type rotation and scale. The features are based on Zernike moments of multispectral correlation functions. Scale invariance is obtained by normalizing the correlation functions by an estimated scale parameter. Illumination intensity effects are removed also by normalization. The experimental results presented in [40] show good performance, but the method involves a rather high computational complexity and a series of parameter estimations and normalizations. Also, the types of geometric transformations handled by these invariant features do not cover the entire set of affine transformations.

Actually our work could be considered a generalization of the work of Reiss [27] and of Van Gool et al. [38]. In [27] Reiss presents 10 functions of central intensity moments up to the fourth order for greyvalue intensity patterns, which are invariant under affine geometric transformations of the image. Photometric changes are dealt with by normalization against both intensity scaling and offset. In [22] an evaluation of the recognition performance obtained with these invariants is presented. When applied to the greylevel version of a set of color outdoor images, a rather weak performance is reported, which most probably is caused by the high order of the moments involved in the invariant functions.

The approach which comes closest to what is reported here is the work of Van Gool et al. presented in [38]. The geometric/photometric invariants in [38] involve shape and intensity moments up to the second order of greyvalue intensity patterns. The invariants are systematically classified according to the highest order of the moments involved. For each case, geometric invariants (affine geometric transformations), photometric invariants, as well as combined geometric/photometric invariants are given. The photometric changes involve either intensity scaling and offset, or only scaling. Some of these invariants require an affine invariant area subdivision of the pattern (which makes them computationally more demanding). An improvement of the affine invariant area subdivision of the pattern is the solution introduced by Mindru et al. [21]. That approach considers a pixelwise subdivision of the pattern, based on separate moments for the pixels darker and lighter than
the average intensity (histogram based subdivision). This method has the advantage that it provides an affine invariant subdivision which does not depend on the pattern's outline. An evaluation of the recognition performance obtained with these invariants is also presented in [22] and shows that rather good results can be obtained.

A limitation of these approaches is that one may have to let grow the order of the moments beyond the point where they remain stable, in order to create a sufficient number of moment invariants. These problems are remedied by introducing powers of the intensities in the individual color bands and combinations thereof in the expressions for the moments. This solution was introduced by Mindru et al. [21], where invariants were built as rational expressions of generalized color moments rather than the traditional moments. Also the work reported here is based on generalized color moments.

3.1. Generalized color moments

A color pattern can be represented as a vector-valued function \( I \) defined on a region \( \Omega \) in the (image) plane and assigning to each image point \((x, y) \in \Omega\) the 3-vector \( I(x, y) = (R(x, y), G(x, y), B(x, y)) \) containing the RGB-values of the corresponding pixel. The generalized color moment \( M_{pq}^{abc} \) is defined by

\[
M_{pq}^{abc} = \int \int \alpha^p \beta^q [R(x, y)]^a [G(x, y)]^b [B(x, y)]^c \, dx \, dy.
\]

(5)

\( M_{pq}^{abc} \) is said to be a (generalized color) moment of order \( p + q \) and degree \( a + b + c \). Observe that generalized color moments \( M_{pq}^{000} \) of degree 0 are in fact the \((p, q)\)-shape moments of the image region \( \Omega \); and, that the generalized color moments of degree 1, viz. \( M_{pq}^{001}, M_{pq}^{010}, M_{pq}^{100} \), are just the \((p, q)\)-intensity moments of respectively the \(R\)-, \(G\)- and \(B\)-color band. On the other hand, the generalized color moments \( M_{pq}^{abc} \) of order 0 are the non-central \((a, b, c)\)-moments of the (multivariate) color distribution of the RGB-values of the pattern. Hence, these generalized color moments generalize shape moments of planar shapes, intensity moments of greylevel images, and non-central moments of the color distribution in the image. A large number of generalized color moments can be generated with only small values for the order and the degree. This is key to the extraction of robust moment invariants. In our work, only generalized color moments up to the first order and the second degree are considered, thus the resulting invariants are functions of the generalized color moments \( M_{000}^{abc}, M_{001}^{abc}, M_{010}^{abc}, M_{100}^{abc} \) with \((a, b, c) \in \{(0, 0, 0), (1, 0, 0), (0, 1, 0), (0, 0, 1), (2, 0, 0), (0, 2, 0), (0, 0, 2), (1, 1, 0), (1, 0, 1), (0, 1, 1)\} \).

3.2. Geometric and photometric effects on moments

A first remark concerning the effect of the transformations on the set of moments is concerned with projective transformations. Due to the fact that a (finite) set of moments cannot be closed under the action of the projective group (the presence of a moment of order \( p + q \) forces the measurement set to also contain a moment of order \( p + q + 1 \) if it is to be closed under projective transformations), projective invariant
moment invariants do not exist [39]. As a consequence, if one has to deal with perspective deformations (and not just affine), these deformations have to be eliminated first through shape normalization.

Another remark is that the actions of the affine geometric and the photometric changes commute for these moments. As a consequence, the overall group of geometric–photometric transformations is a direct product of the affine group and the photometric group (Eqs. (2)–(4)). Thus, invariants exist if the number of moments surpasses the sum of the orbit dimensions of both actions taken separately and they are found as common expressions in the sets of affine and photometric moment invariants separately.

Another consequence of this remark is that, since the actions commute, one might first normalize against one type of transformation and then against the other. Alternatively, one may normalize against one and switch to the use of invariants for the other. The photometric offset can, e.g., be eliminated through the use of intensity minus average intensity and the photometric scale parameters can be eliminated by normalizing the resulting intensity's variance (as done by Reiss [27]). After these normalizations one has to deal with geometric deformations exclusively. In [22] a compromise was made by normalizing against photometric offset alone and using invariance under affine geometric transformations and photometric scaling. When the geometric transformations are dealt with through normalization, one has to deal with photometric changes exclusively.

3.3. Classification of the generalized color moment invariants

The moment invariants are obtained by Lie group methods (for details on Lie methods in computer vision we refer to [24,39]). Following the Lie group approach, invariants are found as solutions of systems of partial differential equations.

Our goal is to build generalized color moment invariants, i.e., rational expressions of the generalized color moments (5), that do not change under the selected geometric and photometric transformations. Moreover, we prefer to only use those moments that are of a simple enough structure to be robust under noise. That means that high orders and high degrees should be avoided. We recall that we only consider moments up to the first order and the second degree, in the 3 bands.

The invariants can be classified according to three parameters: the order, the degree and the number of color bands of the moments involved. To allow maximal flexibility for the user in the choice of the color bands we consider moments involving one, two or three color bands. At the same time, the set of moments is gradually built including first the lowest-order moments that deliver invariants, and then increasing the order (up to the second order) to expand the set of invariants. Of course, the separation of the color bands is only possible when considering photometric transformations of Type D (Eq. (2)) or SO (Eq. (3)), since the AFF transformations (Eq. (4)) imply that a certain color band depends on all 3 color bands, thus they cannot be separated.

As a result of this systematic procedure, a classification of the basis invariants involving 1, 2 or 3 band moments is generated. A basis of invariants for a particular
c 2-bands invariants of order 10

Table 1

<table>
<thead>
<tr>
<th>Type 1 (GPD)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_{02} = \frac{M_{00}^4 M_{02}^2}{M_{00}^4 M_{02}^2}$</td>
<td>$D_{02} = \frac{M_{00}^{10} M_{00}^6}{M_{00}^{10} M_{00}^6}$</td>
</tr>
</tbody>
</table>
| $S_{12} = M_{10}^2 M_{02}^2 M_{00}^1 + M_{10}^2 M_{02}^2 M_{00}^1$ | $
- M_{10}^2 M_{02}^2 M_{00}^1 + M_{10}^2 M_{02}^2 M_{00}^1 - M_{10}^2 M_{02}^2 M_{00}^1$ |
| $D_{11} = M_{01}^{10} M_{10}^6 M_{00}^1 + M_{01}^{10} M_{10}^6 M_{00}^1 + M_{10}^{10} M_{01}^6 M_{00}^1$ | $D_{11} = M_{01}^{10} M_{10}^6 M_{00}^1 + M_{01}^{10} M_{10}^6 M_{00}^1 + M_{10}^{10} M_{01}^6 M_{00}^1$ |
| $D_{12} = M_{10}^{10} M_{01}^6 M_{00}^1 + M_{10}^{10} M_{01}^6 M_{00}^1 + M_{10}^{10} M_{01}^6 M_{00}^1$ | $D_{12} = M_{01}^{10} M_{10}^6 M_{00}^1 + M_{01}^{10} M_{10}^6 M_{00}^1 + M_{01}^{10} M_{10}^6 M_{00}^1$ |
| $D_{13} = M_{10}^{10} M_{01}^6 M_{00}^1 + M_{10}^{10} M_{01}^6 M_{00}^1 + M_{10}^{10} M_{01}^6 M_{00}^1$ | $D_{13} = M_{01}^{10} M_{10}^6 M_{00}^1 + M_{01}^{10} M_{10}^6 M_{00}^1 + M_{01}^{10} M_{10}^6 M_{00}^1$ |
| $D_{14} = M_{10}^{10} M_{01}^6 M_{00}^1 + M_{10}^{10} M_{01}^6 M_{00}^1 + M_{10}^{10} M_{01}^6 M_{00}^1$ | $D_{14} = M_{01}^{10} M_{10}^6 M_{00}^1 + M_{01}^{10} M_{10}^6 M_{00}^1 + M_{01}^{10} M_{10}^6 M_{00}^1$ |
| $D_{15} = M_{10}^{10} M_{01}^6 M_{00}^1 + M_{10}^{10} M_{01}^6 M_{00}^1 + M_{10}^{10} M_{01}^6 M_{00}^1$ | $D_{15} = M_{01}^{10} M_{10}^6 M_{00}^1 + M_{01}^{10} M_{10}^6 M_{00}^1 + M_{01}^{10} M_{10}^6 M_{00}^1$ |
| $D_{16} = M_{10}^{10} M_{01}^6 M_{00}^1 + M_{10}^{10} M_{01}^6 M_{00}^1 + M_{10}^{10} M_{01}^6 M_{00}^1$ | $D_{16} = M_{01}^{10} M_{10}^6 M_{00}^1 + M_{01}^{10} M_{10}^6 M_{00}^1 + M_{01}^{10} M_{10}^6 M_{00}^1$ |
| $D_{17} = M_{10}^{10} M_{01}^6 M_{00}^1 + M_{10}^{10} M_{01}^6 M_{00}^1 + M_{10}^{10} M_{01}^6 M_{00}^1$ | $D_{17} = M_{01}^{10} M_{10}^6 M_{00}^1 + M_{01}^{10} M_{10}^6 M_{00}^1 + M_{01}^{10} M_{10}^6 M_{00}^1$ |
| $D_{18} = M_{10}^{10} M_{01}^6 M_{00}^1 + M_{10}^{10} M_{01}^6 M_{00}^1 + M_{10}^{10} M_{01}^6 M_{00}^1$ | $D_{18} = M_{01}^{10} M_{10}^6 M_{00}^1 + M_{01}^{10} M_{10}^6 M_{00}^1 + M_{01}^{10} M_{10}^6 M_{00}^1$ |
| $D_{19} = M_{10}^{10} M_{01}^6 M_{00}^1 + M_{10}^{10} M_{01}^6 M_{00}^1 + M_{10}^{10} M_{01}^6 M_{00}^1$ | $D_{19} = M_{01}^{10} M_{10}^6 M_{00}^1 + M_{01}^{10} M_{10}^6 M_{00}^1 + M_{01}^{10} M_{10}^6 M_{00}^1$ |
| $D_{20} = M_{10}^{10} M_{01}^6 M_{00}^1 + M_{10}^{10} M_{01}^6 M_{00}^1 + M_{10}^{10} M_{01}^6 M_{00}^1$ | $D_{20} = M_{01}^{10} M_{10}^6 M_{00}^1 + M_{01}^{10} M_{10}^6 M_{00}^1 + M_{01}^{10} M_{10}^6 M_{00}^1$ |

$M_{ij}$ stands for either $M_{pq}^{00}$, $M_{pq}^{01}$ or $M_{pq}^{02}$, depending on which color band is used; $M_{ij}$ stands for either $M_{ij}^{00}$, $M_{ij}^{01}$ or $M_{ij}^{02}$, depending on which 2 of the 3 color bands are used.
contains only independent invariants. There are 24 invariants in the collection obtained by applying the 1- and 2-band GPD invariants to all the combinations of the 3 color bands, and there are only 21 independent invariants in the basis. The 21 independent invariants in the basis are obtained by removing the following three invariants: $D_{12}^{(RB)}$, $D_{12}^{(RG)}$, and $D_{12}^{(GB)}$.

### 3.3.2. GPSO invariants

For affine geometric deformations and photometric transformations of Type SO, all Geometric/Photometric invariants (GPSO Type) involving generalized color moments up to the first order and second degree are functions of the invariants defined in Table 2.

There are 18 basis invariants involving generalized color moments in all 3 color bands. Again, every invariant involving all 3 color bands is a function of invariants $M_{pq}^c$ stands for either $M_{pq}^{R0}$, $M_{pq}^{G0}$ or $M_{pq}^{B0}$, depending on which color band is used; $M_{pq}^d$ stands for either $M_{pq}^{R0}$, $M_{pq}^{G0}$ or $M_{pq}^{B0}$, depending on which 2 of the 3 color bands are used.

<table>
<thead>
<tr>
<th>Type 2 (GPSO)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_{12} = \frac{(M_{00}^{R0}M_{00}^{G0} - M_{00}^{R0}M_{00}^{B0})^2 - M_{00}^{R0}M_{00}^{G0}M_{00}^{B0} + M_{00}^{R0}M_{00}^{B0}M_{00}^{G0} + M_{00}^{R0}M_{00}^{B0}M_{00}^{G0})}{(M_{00}^{G0})^2[M_{00}^{R0}M_{00}^{G0} - (M_{00}^{B0})^2]^{3/2}}$</td>
</tr>
<tr>
<td>$D_{02} = \frac{[(M_{00}^{G0})^2 - M_{00}^{G0}M_{10}^{G0}]^2}{[M_{00}^{G0}M_{00}^{R0} - (M_{00}^{B0})^2]^2}$</td>
</tr>
<tr>
<td>$D_{12} = \frac{[M_{00}^{G0}M_{00}^{B0}M_{10}^{G0} - M_{00}^{R0}M_{00}^{B0}M_{00}^{G0} + M_{10}^{R0}M_{10}^{B0}M_{00}^{G0} + 2M_{10}^{R0}M_{10}^{B0}M_{00}^{G0} - 2M_{10}^{R0}M_{10}^{B0}M_{00}^{G0})^2}{(M_{00}^{G0})^2[M_{00}^{R0}M_{00}^{G0} - (M_{00}^{B0})^2]^{3/2}}$</td>
</tr>
<tr>
<td>$D_{12}^2 = \frac{[M_{00}^{R0}M_{00}^{G0}M_{00}^{B0} - M_{00}^{R0}M_{00}^{G0}M_{00}^{B0} + M_{00}^{R0}M_{00}^{G0}M_{00}^{B0} + 2M_{00}^{R0}M_{00}^{G0}M_{00}^{B0} - 2M_{00}^{R0}M_{00}^{G0}M_{00}^{B0})^2}{(M_{00}^{G0})^2[M_{00}^{R0}M_{00}^{G0} - (M_{00}^{B0})^2]^{3/2}}$</td>
</tr>
<tr>
<td>$D_{12}^3 = \frac{[M_{00}^{R0}M_{00}^{G0}M_{00}^{B0} - M_{00}^{R0}M_{00}^{G0}M_{00}^{B0} + M_{00}^{R0}M_{00}^{G0}M_{00}^{B0} + 2M_{00}^{R0}M_{00}^{G0}M_{00}^{B0} - 2M_{00}^{R0}M_{00}^{G0}M_{00}^{B0})^2}{(M_{00}^{G0})^2[M_{00}^{R0}M_{00}^{G0} - (M_{00}^{B0})^2]^{3/2}}$</td>
</tr>
<tr>
<td>$D_{12}^4 = \frac{[M_{00}^{R0}M_{00}^{G0}M_{00}^{B0} - M_{00}^{R0}M_{00}^{G0}M_{00}^{B0} + M_{00}^{R0}M_{00}^{G0}M_{00}^{B0} + 2M_{00}^{R0}M_{00}^{G0}M_{00}^{B0} - 2M_{00}^{R0}M_{00}^{G0}M_{00}^{B0})^2}{(M_{00}^{G0})^2[M_{00}^{R0}M_{00}^{G0} - (M_{00}^{B0})^2]^{3/2}}$</td>
</tr>
</tbody>
</table>
which involve only 2 of the 3 bands. The basis of (independent) invariants contains all the invariants $S_{pq}^{(K)}$ and $D_{pq}^{(KL)}$ defined in Table 2, evaluated in the color bands $K$ and $L$, without the following three invariants: $D_{12}^{2(RG)}$, $D_{12}^{2(GB)}$ and $D_{12}^{3(RB)}$.

3.3.3. PSO invariants

The photometric invariants PSO are meant for cases when no geometric deformations are present (i.e., they are either absent or canceled by normalization) and the photometric transformations are of Type SO. All photometric invariants involving generalized color moments up to the first order and second degree are functions of the invariants defined in Table 3.

There are 24 basis invariants involving generalized color moments in all 3 color bands. These 24 invariants are the invariants defined in Table 3 applied to all combinations of 1 and 2 color bands, and the following three moment invariants: $M_{00}^0$ and $M_{pq}^m$, with $pq \in \{01, 10\}$. Again, the basis only consists of 2-band moment invariants.

3.3.4. PSO stabilized invariants (PSO*)

When examining the PSO invariant functions we notice that, except for the trivial 0th degree invariants (i.e., $M_{pq}^0$, $pq \in \{00, 01, 10\}$), they all represent rational functions of moment based expressions. Their denominators represent determinants of matrices whose elements are moments. If these determinants get very small, the moment invariants get (numerically) unstable. Such instabilities were observed in experiments only in the case of PSO invariants. The denominators in the case of GPD, GPSO, and PAFF invariants do not cause instabilities.

The numerical instability of PSO invariants can be circumvented by computing a new set of invariants as functions of the basis invariants, such that the unstable denominators get eliminated. Such correction is only needed for the PSO invariants.

Table 3

Invariants of Type PSO involving 1 or 2 color bands; $S_{cd}$ stands for 1-band invariants, and $D_{cd}$ for 2-bands invariants of order $c$, and degree $d$, respectively

<table>
<thead>
<tr>
<th>Type 3 (PSO)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_{11} = \frac{M_{00}^0 M_{10}^0 - M_{00}^0 M_{10}^0}{M_{00}^0 M_{10}^0 - M_{00}^0 M_{10}^0}$</td>
</tr>
<tr>
<td>$S_{12} = M_{pq}^0 M_{pq}^0 - (M_{pq}^0)^2$, $pq \in {01, 10}$</td>
</tr>
<tr>
<td>$S_{22} = \frac{M_{pq}^0 M_{pq}^0}{M_{00}^0 M_{10}^0 - M_{00}^0 M_{10}^0}$</td>
</tr>
<tr>
<td>$D_{11} = \frac{M_{00}^0 M_{10}^0 - M_{00}^0 M_{10}^0}{(M_{00}^0 M_{10}^0 - M_{00}^0 M_{10}^0)}$</td>
</tr>
<tr>
<td>$D_{12} = \frac{M_{pq}^0 M_{pq}^0}{M_{pq}^0 M_{pq}^0 - M_{pq}^0 M_{pq}^0}$, $pq \in {01, 10}$</td>
</tr>
</tbody>
</table>

$M_{pq}^0$ stands for either $M_{pq}^{00}$, $M_{pq}^{01}$, or $M_{pq}^{10}$, depending on which color band is used; $M_{pq}^0$ stands for either $M_{pq}^{00}$, $M_{pq}^{01}$, or $M_{pq}^{10}$, depending on which 2 of the 3 color bands are used.
Examining the PSO basis invariants with common denominators it becomes clear that several combination strategies are possible. One strategy that maintains the maximum number of resulting independent invariants has the following structure:

\[
SO_1 = \frac{(S_{12}^2)^2}{(S_{12}^1[pq = 10])}, \quad SO_2 = \frac{S_{11}D_{12}^1[pq = 10]}{D_{12}^1},
\]

\[
SO_3 = \frac{S_{11}D_{12}^1}{D_{12}^1[pq = 01]},
\]

and the resulting invariants are the following, which we will hereafter refer to as P3O*:

\[
SO_1 = \frac{(M_{10}^0 M_{00}^0 - M_{00}^1 M_{10}^0)^2}{(M_{10}^0 M_{00}^0 - M_{10}^1 M_{00}^0)(M_{01}^0 M_{10}^0 - M_{01}^1 M_{10}^0)},
\]

\[
SO_2 = \frac{M_{10}^0 M_{10}^1 - M_{10}^0 M_{10}^0}{M_{01}^0 M_{10}^0 - M_{01}^0 M_{00}^0}, \quad SO_3 = \frac{M_{00}^0 M_{00}^1 - M_{00}^0 M_{00}^0}{M_{01}^0 M_{01}^0 - M_{01}^0 M_{01}^0};
\]

3.3.5. PAFF invariants

The photometric invariants PAFF are meant for cases when no geometric deformations are present (i.e., they are either absent or canceled by normalization) and the photometric transformations are of Type AFF. Due to the complexity of the computations only 11 out of the existing 14 photometric invariants PAFF with moments up to the first order and second degree were retrieved and they are given in Table 4. Of the 11 PAFF invariants, 2 are the invariants \(T_{12}^1\), \(T_{21}^1\), and the remaining 9 are of type \(T_{12}^1(i, j)\) with \(pq \in \{00, 01, 10\}\) and \(ij \in \{11, 12, 22\}\). All invariants are functions of the 3 color bands, since color band separation is not possible.

The next section investigates the practical use of the moment invariants in terms of discriminant power and robustness to noise. The combinations of geometric and

| Invariants of Type PAFF involving all 3 color bands; \(T_{ij}\) stands for 3-bands invariants of order \(c\), and degree \(d\), respectively; \(pq \in \{00, 01, 10\}\) and \(ij \in \{11, 12, 22\}\) |
|---|---|---|---|---|
| Type 4 (PAFF) |

\[
k_j^1 = M_{10}^0 M_{00}^0 - M_{10}^0 M_{00}^0, \quad k_j^2 = M_{10}^0 M_{00}^0 - M_{10}^0 M_{00}^0, \quad \ldots
\]

\[
J_{pq} = \begin{pmatrix} d_{11} & d_{12} & d_{13} \\ d_{21} & d_{22} & d_{23} \\ d_{31} & d_{32} & d_{33} \end{pmatrix} \begin{pmatrix} \frac{M_{pq}^{00} M_{pq}^{00} - (M_{pq}^{00})^2}{M_{pq}^{00} M_{pq}^{00} - M_{pq}^{00} M_{pq}^{00}} \\ \frac{M_{pq}^{00} M_{pq}^{00} - (M_{pq}^{00})^2}{M_{pq}^{00} M_{pq}^{00} - M_{pq}^{00} M_{pq}^{00}} \end{pmatrix},
\]

\[
U_j = \begin{pmatrix} k_j^1 & d_{12} & d_{13} \\ k_j^2 & d_{22} & d_{23} \\ k_j^3 & d_{32} & d_{33} \end{pmatrix} U_j = \begin{pmatrix} d_{11} & k_j^1 & k_j^2 \\ d_{21} & k_j^2 & k_j^3 \\ d_{31} & k_j^3 & k_j^1 \end{pmatrix},
\]

\[
T_{12}^1 = \frac{|J_{12}|}{|J_{00}|}, \quad T_{21}^1 = \frac{|J_{01}|}{|J_{00}|}, \quad T_{12}^1(i, j, pq) = \frac{k_j^1 |U_j|^1 + k_j^2 |U_j|^2 + k_j^3 |U_j|^3}{|J_{pq}|}.
\]
photometric models are compared in terms of the discriminant power of their invariants and the classification performance under different experimental settings, since it is this overall performance that is of real interest.

4. Performance evaluation

The recognition performance is estimated using classifiers based on feature vectors consisting of moment invariants. Each Type of moment invariants form a separate feature vector. The experiments involve the following steps:
(1) Extract the data (region of interest) from all images, with manual delineation of planar parts, if necessary.
(2) Extract the moment invariants (Type GPD, Type GPSO, stabilized Type PSO, and Type PAFF), after a normalization has been applied, if the type of invariants require so.
(3) Statistical analysis of the overall sample population, and extraction of the five main canonical variables following a MANOVA [16] (i.e., five linear combinations of the moment invariants, separately for each Type of moment invariants. Canonical variables are linear combinations of the moment invariants within the feature vector that maximize the separation between classes).
(4) Recognition following a leave-one-out strategy: each time a sample is singled out and all the others are used as training set. The sample is then assigned to a class based on the classification scheme. This process is repeated for all samples in the data set. The following classification schemes are used: quadratic discriminant functions (QDF) using the first five canonical variables, and $k$ nearest neighbors ($k$NN), with $k = 1$ or $k = 3$, based on the Mahalanobis distance, using the entire feature vector.

4.1. Synthetic transformations

In a first experiment, the patterns are transformed synthetically. The tests aim at assessing the constancy, robustness, and discriminant power of the invariants under supervised model conditions. In particular, the tests are aimed at evaluating the robustness of the invariants under deviations from the ideal model conditions. To this end, geometric and photometric transformations were applied to a set of 30 real color images, which are shown in Fig. 1.

The following types of synthetic transformations were applied to the patterns:
1. 24 photometric transformations of type SO
2. 24 photometric transformations of type AFF
3. 24 geometric transformations of type affine
4. 24 geometric transformations of type perspective
5. 24 geometric transformations of type affine combined with the 24 photometric transformations of type SO
6. 24 geometric transformations of type perspective combined with the 24 photometric transformations of type SO
7. **24 geometric** transformations of type **perspective** combined with the **24 photometric** transformations of type **AFF**.

Each series of transformations applied to the original patterns generates an image data set $\text{ImSet}_k$, $k = 1, \ldots, 7$.

The diagonal elements of the photometric transformations range between 0.4 and 1.8 and their offsets between $-35$ and $15$, with different values corresponding to the different bands. In the case of the AFF transformations, the off-diagonal elements of the first 12 transformations range between 0.05 and 0.4 and are smaller than the
diagonal elements. The last 12 transformations have off-diagonal elements with larger values, in the range of 0.1 and 0.9. This is in agreement with model fitting experiments on outdoor images [23].

The affine geometric transformations applied to the images are combinations of scaling (resize, 10% less), skewing (10 or 20% along \( y \) axis), and rotations with 10°, 20°, 30° or 45° towards left or right. The perspective transformations are generated by placing the vanishing point on the horizontal or vertical axis, and they are applied to the original pattern or a rotated version of it. The series of transformations (as shown in Figs. 2 and 3) were applied to all 30 original images, resulting in the series of 24 geometrically transformed versions of each of the original patterns.

All four types of moment invariants taken over the 30 images under ideal model conditions yielded 100% recognition performance with both QDF and kNN classification methods, which corroborates the correctness of the invariants and their implementation. It also demonstrates that the moment invariant vectors have discriminant power.

Fig. 2. The series of affine geometric transformations.
The recognition performance obtained with quadratic discriminant functions using the first $c$ canonical variables $QDF(c)$ and with the nearest neighbor method ‘NN’ ($K = 1$) under non-ideal conditions is the following:

- Invariants GPD
  - $ImSet_1$: $NN = 100\%$, $QDF(15) = 100\%$
  - $ImSet_2$: $NN = 98.9\%$, $QDF(15) = 97.9\%$
  - $ImSet_4$: $NN = 100\%$, $QDF(15) = 100\%$
  - $ImSet_5$: $NN = 100\%$, $QDF(15) = 100\%$
  - $ImSet_6$: $NN = 99.9\%$, $QDF(15) = 98.9\%$
  - $ImSet_7$: $NN = 97.2\%$, $QDF(15) = 94.7\%$

- Invariants GPSO
  - $ImSet_2$: $NN = 96.8\%$, $QDF(10) = 96.1\%$
  - $ImSet_4$: $NN = 100\%$, $QDF(10) = 100\%$
  - $ImSet_6$: $NN = 100\%$, $QDF(10) = 100\%$
  - $ImSet_7$: $NN = 83.6\%$, $QDF(10) = 87.8\%$
Invariants PSO
  - ImSet2: $NN = 96.5\%$, $QDF(15) = 83.5\%$

Stabilized invariants PSO*
  - ImSet2: $NN = 100\%$, $QDF(5) = 100\%$

The results given above specify the results under non-ideal conditions for the four types of moment invariants. Again, performance is good, although no longer perfect. This should not come as a surprise, as, e.g., invariance under affine transformations cannot be expected to shield off against the influence of perspective deformations. We notice the improvement in performance and robustness of the stabilized PSO invariants over the PSO basis invariants. The classification results show that although Type SO photometric transformations are more complex than the assumed ideal model for invariants Type GPD, these invariants remain quite stable under such transformations, whereas AFF photometric transformations turn out to be more difficult to withstand. Also GPSO invariants show a decrease in performance under AFF photometric transformations, whereas the stabilized invariants PSO show enhanced stability against them.

4.2. Transformations in real scene images

In order to compare the recognition performance of real scene images for the four Types of moment invariants, experiments were run on two main sorts of images. One kind of data consists of indoor images of scenes under different illuminations and the second database contains digital color images of outdoor advertisement panels.

Since PSO and PAFF moment invariants can only cope with photometric changes, before measuring PSO and PAFF invariants on patterns that suffered geometric deformations, these deformations are dealt with through a normalization step. This is achieved by first selecting the four corners of the frame that contains the region of interest in the image (i.e., the pattern considered for classification). This was easy in the case of our experiments based on the advertisement panels which have a simple rectangular shape. A similar procedure was applied to delineate the planar objects belonging to the database of indoor images. Then the four corners of the frame were brought to four canonical positions (homography matrix based computation) in order to achieve the geometric normalization. Examples of geometrically normalized images are shown in Fig. 4. The normalized shape more or less corresponds to a rectangle with an aspect ratio that one would get in a head-on view.
Geometric normalization is not needed for the GPD and GPSO invariants. But, (only) for the GPD invariants, a photometric normalization against the photometric offset (as described in [22]) needs to be applied before computing these invariants.

4.2.1. Indoor images—artificial light

This section presents the assessment of the recognition performance of the invariant features when applied to the set of indoor images with planar surfaces that is publicly available at http://www.cs.sfu.ca/~colour/data. These data were collected by Lindsay Martin and Kobus Barnard, as part of investigations into computational color constancy algorithms. The images are in TIFF format. Several preprocessing steps were taken to improve the data. First, some fixed pattern noise was removed. Second, images were corrected for a spatially varying chromaticity shift due to the camera optics. Finally, the images were mapped into a more linear space. This included removing the sizable camera black signal. A detailed presentation of the data is available in [2].

The database consists of objects viewed on a black background, under 11 different illuminants with different colors. The set of images with minimal specularities and containing planar surfaces were selected for our tests. That brought us to the collection of the following objects (Fig. 5): books-2, collage, macbeth, munsell1, munsell2, munsell3, munsell4, munsell5, paper1, paper2, sml-mondrian1, sml-mondrian2, book1, book2, book3, book4, cruncheroos, macaroni, rollups, thus 19 in total. When required viewpoint invariance was achieved by a normalization step, as described before.

For these images the photometric changes are reported to be of Type D [2] and the geometric deformations are of type affine (removed by normalization for the
invariants of Type PSO and PAFF). It is then to be expected that all four Types of invariants would perform well, since there are no significant departures from the assumed theoretical models. Indeed, the high recognition rates obtained for all invariants with all classification methods (presented in Table 5), and the maximum recognition rate (above 98%) that can be obtained for each Type prove that the moment invariants are well suited for reliable and efficient recognition using this kind of indoor images.

4.2.2. Outdoor images

In the case of the outdoor advertisement panels, the goal was to recognize to which of the 16 classes (Fig. 6) a panel belongs. For each of the classes between 10 and 18 images were taken, under quite a large variety of viewing conditions (Fig. 7). Samples of the same class also included images of several, physically different panels. We recall that the model that best describes the intensity changes is that of Type AFF [23]. Also, the geometric deformations can often be fairly well approximated by an affine transformation, but certain degrees of perspective deformation are present in the dataset. No geometric normalization has been applied before

<table>
<thead>
<tr>
<th>Table 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recognition rates for indoor images (left) and outdoor images (right)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Indoor images</th>
<th>QDF(5)</th>
<th>NN</th>
<th>KNN</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPD</td>
<td>99.5</td>
<td>100</td>
<td>98.6</td>
</tr>
<tr>
<td>GPSO</td>
<td>93.3</td>
<td>98.6</td>
<td>97.6</td>
</tr>
<tr>
<td>PSO*</td>
<td>98.1</td>
<td>98.6</td>
<td>97.1</td>
</tr>
<tr>
<td>PAFF</td>
<td>96.2</td>
<td>98.6</td>
<td>96.3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outdoor images</th>
<th>QDF(5)</th>
<th>NN</th>
<th>KNN</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPD</td>
<td>57.0</td>
<td>79.9</td>
<td>74</td>
</tr>
<tr>
<td>GPSO</td>
<td>88.6</td>
<td>98.8</td>
<td>96.7</td>
</tr>
<tr>
<td>PSO*</td>
<td>99.2</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>PAFF</td>
<td>93.9</td>
<td>92.2</td>
<td>91.2</td>
</tr>
</tbody>
</table>

Fig. 6. Outdoor images—the 16 different patterns that were used in the classification system.
computing Type GPD and GPSO invariants. This means that all the invariants, except those of Type PAFF, need to cope with a more complex type of changes than they are designed for.

As the results presented in Table 5 show, now a substantially different recognition performance is achieved by the 4 types of invariants. Examining the performance obtained with Type GPD and Type GPSO, we immediately notice a serious decrease in recognition performance as compared to the indoor images. Two main differences between the two datasets cause most of the performance decrease: first, the more complex photometric transformations to be handled in this case, and second, the presence of certain amounts of perspective deformations of the billboard images, whereas these invariants were only designed to be immune against affine deformations. Nevertheless, the average rate of about 90% and the maximum rate of 98.8% obtained with the GPSO invariants suggests that the geometric/photometric type of invariants GPSO can be successfully applied to outdoor images as well, despite the increased complexity of the actual geometric and photometric transformations.

Interestingly, although the Type AFF photometric transformation model offers a better fit to the actual intensity changes than the simpler transformations of Type SO, the stabilized PSO invariants prove to yield slightly better recognition rates than the PAFF invariants. This is due to the higher inherent numerical complexity of the PAFF invariants.

It seems therefore that the invariants based on Type SO photometric transformations are a good choice for both indoor and outdoor images. They provide a good balance between simplicity and performance.

4.3. Non-planar objects

At first, the practical applicability of the moment invariants may seem rather restricted, as only planar patches can be dealt with. Moreover, these patches need to be segmented out of the images first. Here, we show that the invariants can also be used
under more general conditions, e.g., for 3D object recognition tasks. The trick lies in the fact that most 3D objects can locally be approximated by planar patches. The only difficulty consists of delineating those local planar patches in an automatic way and for each image independently. Tuytelaars and Van Gool [35,37] have proposed a method to extract such local surface patches in a way that is independent of the viewpoint and illumination. These ‘affine invariant regions’ change their shape in the image in such a way that they consistently delineate the same physical part of the surface in different views. As these invariant regions carve out small parts of the surface, there is a good chance that these parts are more or less planar, such that they can be described (and hence recognized, i.e., matched between different views) with the moment invariants introduced earlier. First, we’ll shortly describe the method proposed in [35,37] for extracting such affine invariant regions. Next, we demonstrate the strength of this powerful combination of affine invariant regions on the one hand and color moment invariants on the other hand with some wide baseline matching experiments.

Tuytelaars and Van Gool proposed both a geometry-based and an intensity-based method for the extraction of invariant regions. In both cases, they included invariance to affine geometric deformations and photometric transformations of type $SO$. The geometry-based method starts from a Harris corner point $p$ and a pair of edges in its neighbourhood, as shown in Scheme 1. The edges are parametrized using a relative affine invariant parameter. This allows to find for each point $p_1$ on the first edge the corresponding point $p_2$ on the second edge. Together with the corner point, these points define a parallelogram-shaped region for each value of the relative affine invariant parameter. From this, one (or a few) regions are selected, by evaluating a photometric (moment-based) function over the regions and taking the local extrema.

The intensity-based method, on the other hand, starts from local extrema in intensity as anchor points (see Scheme 2). It then evaluates the intensity function along rays emanating from the extremum. On each ray a point is selected, for which a particular function reaches an extremum. Linking these points together yields an affine invariant region, to which an ellipse is fitted, that yields the sought delineating contour. Finally, we double the size of the ellipses found. This leads to more distinctive regions, due to a more diversified texture pattern within the region and hence
facilitates the matching process, at the cost of a higher risk of non-planarity due to the less character of the regions.

Both methods turn out to complement one another very well, in that invariant regions are typically found at different locations in the image. Other methods for extracting affine invariant regions have recently been proposed by Baumberg [3], Matas [19], and Mikolajczyk [20], and could be used for this purpose as well.

Once we have extracted affine invariant regions from an image, we can describe their content using the moment invariants introduced earlier. Based on the resulting invariant feature vectors, it is possible to find correspondences between regions extracted in two very different images of the same object or scene. Moment invariants in combination with invariant neighbourhoods are therefore a powerful tool in various applications. The local character of the descriptor yields robustness to occlusions, changing backgrounds and non-planar objects. The geometric and photometric invariance, on the other hand, allows to deal with large changes in viewpoint and illumination conditions. Several experimental results are reported for a wide range of applications going from fast matching of interest points in the context of wide baseline stereo ([34,37]), over object recognition and content-based image retrieval ([34,36]) to visual servoing of a mobile robot ([34,35]). In these experiments, the matching was based on our Type GPSO invariants calculated over the invariant neighbourhoods. It is beyond the scope of this paper to present in detail all these experiments, so here we will only tackle the use of the local invariant descriptors in the context of the wide baseline correspondence problem.

The primary goal of matching the local invariant descriptors in the context of wide baseline stereo vision is finding corresponding features in different views of the same object or scene. This is typically a much harder problem than the recognition experiments reported earlier, for several reasons. First, the number of 'classes' to be distinguished is undetermined and considerably larger, as several hundreds of invariant regions are usually extracted from a single image. Moreover, several classes may be virtually indistinguishable due to symmetry or repetition—think, e.g., of different windows on the same building. Second, one typically has only a single learning example (the corresponding region in the other image). Third, due to imperfections
in the region extraction as well as non-planarities (i.e., serious deviations from the
local planarity assumption), occlusions, etc., typically only a fraction of the regions
extracted from one view do indeed have a corresponding region extracted from the
other view. Fourth, the information content within the invariant regions is very lim-
ited, due to their local character. Finally, the effect of noise, discretization errors,
etc., is much bigger, again due to the limited size of the regions. Making the regions
larger, e.g., by simply rescaling them, could reduce the effect of the last two prob-
lems, but at the same time increases the risk of partial occlusions or non-planarities,
so a trade-off needs to be made. On the other hand, plenty of regions are extracted
from a single image, and not all of them need to be matched before one can draw a
conclusion. From our experiments, we learnt that the moment invariants do a very
good job in selecting a few good candidate matching regions in a time-efficient way
(even more so if indexing or hashing techniques are applied). Final verification based
on normalized crosscorrelation (after compensating for the affine deformation) as
well as checking some semi-local or global consistency measures can then be used
to remove the remaining false matches if needed.

Fig. 8 shows some examples of regions that have been matched using these tech-
niques. Note how thanks to the local character of the invariant descriptors, the
system can deal with occlusions, changing background and non-planar objects or scenes.

5. Conclusions

The goal of this work was obtaining a set of image features that can be used for viewpoint and illumination invariant recognition of planar color patterns.

The invariant features are functions of generalized color moments, which are a generalization of the traditional moments. These generalized moments combine powers of pixel coordinates and their intensities in the individual color bands, and thus they implicitly characterize the shape and the color distribution of the pattern in a uniform manner. This property represents one of the main advantages of using these invariant functions. Another advantage obtained by generalizing the traditional definition of the moments is that, even for low orders and degrees, a sufficiently large set of such moments can be extracted to build robust invariants.

Achieving viewpoint and illumination invariance means dealing with a combination of geometric and photometric changes of the patterns. In this work, we investigated alternative choices for the geometric and photometric transformations. Four types of combinations of photometric and geometric changes were considered. Two sets of moment invariants deal with a combination of affine geometric and linear photometric transformations, and two sets are photometric invariants to be combined with normalization against geometric deformations of the pattern. For these cases a systematic classification of invariant functions of generalized color moments was given.

Different experimental settings were considered for assessing the applicability of the four types of invariants. First, the invariants were tested in a pattern recognition setting, with the aim of assessing the robustness and discriminative power of the invariants as global descriptors of planar parts in real images. Different classification techniques were employed in order to provide a better view on the classification performance of the invariants. Synthetically transformed patterns and two main sorts of real scene images were used, namely indoor images with internal changes of light and images of outdoor advertisement panels. The synthetic tests were aimed at evaluating the robustness of the invariants under deviations from the ideal model conditions. For the case of indoor images, diagonal photometric transformations had been reported to be a good fit, but it still needed to be proved that illuminants with variable spectral composition are well handled by the invariant features. The high recognition rates obtained for all invariants with all classification methods suggest that the moment invariants are well suited for reliable and efficient recognition of indoor images. The outcome of the recognition tests on outdoor imagery shows that although the Type AFF photometric transformation model is a better fit to the actual intensity changes in outdoor images than the simpler transformation of Type SO, the PSO invariants yield slightly better recognition rates than the PAFF invariants. In conclusion, the invariants based on the SO Type of photometric transformations are a good
choice for use with both indoor and outdoor images, as they provide a good balance between simplicity and performance.

The use of moment invariants need not be limited to planar objects that can easily be delineated. Several types of applications were presented, one of them into more detail, where moment based invariants can be employed as 3D scene or object descriptors. The automatic delineation of the regions of interest is achieved by extracting local affinely invariant regions from images. Moment invariants have been used successfully for the fast matching of such invariant neighbourhoods.

Acknowledgments

The authors gratefully acknowledge support from the GOA project ‘Variability in Human Shape and Speech VHS+’ (Research Council K.U.Leuven) and the IST project VIBES. Florica Mindru would like to thank ERCIM (the European Research Consortium for Informatics and Mathematics) for sponsoring her fellowship. Tinne Tuytelaars is a Postdoctoral Researcher of the Fund for Scientific Research Flanders (Belgium).

References