Selection and Fusion of Color Models for Image Feature Detection

Harro Stokman and Theo Gevers, *Member*, *IEEE*

Abstract—The choice of a color model is of great importance for many computer vision algorithms (e.g., feature detection, object recognition, and tracking) as the chosen color model induces the equivalence classes to the actual algorithms. As there are many color models available, the inherent difficulty is how to automatically select a single color model or, alternatively, a weighted subset of color models producing the best result for a particular task. The subsequent hurdle is how to obtain a proper fusion scheme for the algorithms so that the results are combined in an optimal setting. To achieve proper color model selection and fusion of feature detection algorithms, in this paper, we propose a method that exploits nonperfect correlation between color models or feature detection algorithms derived from the principles of diversification. As a consequence, a proper balance is obtained between repeatability and distinctiveness. The result is a weighting scheme which yields maximal feature discrimination. The method is verified experimentally for three different image feature detectors. The experimental results show that the fusion method provides feature detection results having a higher discriminative power than the standard weighting scheme. Further, it is experimentally shown that the color model selection scheme provides a proper balance between color invariance (repeatability) and discriminative power (distinctiveness).

Index Terms—Color, learning, feature detection, scene analysis.

✦

1 INTRODUCTION

THE choice of a color space is important for many computer vision algorithms (e.g., feature detection, image classification, object recognition, and visual tracking). It induces the equivalence classes to the actual algorithms. For example, a standard segmentation algorithm will divide a single yellow object, which is partially shaded due to its orientation with respect to a light source, into two separate parts when the algorithm is applied on the RGB sensor color model. In contrast, when the same algorithm is applied on the rgb color model, which is insensitive to shading variations, the yellow object will be detected as an entire object [1].

However, no color space can be considered as universal because color can be interpreted and modeled in different ways. With the large variety of available color spaces (e.g., $RGB, CMY, YIQ, YUV, XYZ, I_1I_2I_3, rg$, CIE $L^*a^*b^*$, $L^*u^*v^*$, HSV, etc.), the inevitable question that arises is how to select the color model that produces the best result for a particular computer vision task. It is possible that several color spaces are equally good candidates (for instance, the color spaces CIE $L^*a^*b^*$ or $L^*u^*v^*$ are both perceptually uniform) or that different color channels have similar properties (for instance, the V and G channels both encode the intensity information for green colors). In these cases, the subsequent question is how to obtain a proper weighting scheme to combine color

spaces or color channels. To this end, a weighting scheme for color models is proposed by Angulo and Serra [2], where a color space representation is derived of type hue-luminance-saturation in which the saturation plays an important role to merge the chromatic and achromatic information during the segmentation process. Macaire et al. [3] provide a hybrid color space to yield high discrimination between the pixel classes in the context of soccer image analysis. In this paper, we focus on selection and fusion of color models for image feature detection.

Similar questions arise in the context of selecting and weighting different feature detection methods, in particular in the case of color edge detection. Here, the default method to combine edges is to use equal weights for the different color feature detectors. In this way, to achieve color edge detection, intensity-based edge detection techniques are extended by taking the sum or Euclidean distance from the individual gradient maps [4], [5]. This default approach is used by other feature detectors. For example, Cardei and Funt [6] combine the output of different color constancy algorithms by simply averaging, based on optimizing the least mean square error, and by using a neural network. By combining the results of several existing algorithms, better illumination estimates are obtained for color constancy.

In this paper, the goal is to select and fuse color models and feature detection algorithms. Feature selection has been used in computer vision where one or more visual features are chosen from a given initial set of candidates. Based on the notion of class separability, various methods have been proposed to select the feature subset that optimizes the discriminatory power regarding the ground truth consisting of positive and negative data labels [7], [8] [9], [10], [11]. For example, model selection by SVM [7] is used to select the most discriminating appearance features. For image retrieval, a method is provided by Tieu and Viola [10] based on AdaBoost to select a small number of features from a large feature set. In the context of face

---

- H. Stokman is with the Intelligent System Laboratory Amsterdam, Faculty of Science, University of Amsterdam, Kruislaan 403, 1098 SJ, Amsterdam, The Netherlands. E-mail: stokman@science.uva.nl.
- T. Gevers is with the Intelligent System Laboratory Amsterdam, Faculty of Science, University of Amsterdam, Kruislaan 403, 1098 SJ, Amsterdam, The Netherlands, and ICREA, Computer Vision Center (UAB), 08193 Bellaterra, Barcelona, Spain.
  E-mail: gevers@science.uva.nl, gevers@cvc.uab.es.

recognition, PCA is applied to reduce the dimensionality [8]. Most of the classifiers are dichotomic (two-class) learning from positive and negative data labels. In contrast to these two-class methods, we focus on the one-class problem (only positive labels) as it is not evident how to choose proper negative data labels in the context of image feature detection. Contradictory chosen negative data labels result in an invalidated impact on the classification accuracy [12], [13].

Therefore, in this paper, to achieve proper color model selection and fusion of feature detection algorithms, we introduce the weighting scheme proposed by Markowitz [14] using only positive training samples. The method is generic and exploits nonperfect correlation between color models or feature detection algorithms derived from the principles of diversification. The result is a weighting scheme which yields maximal feature discrimination.

Further, two criteria are used to assess the performance of the color feature detectors: 1) (*repeatability*) they should be *invariant* (stable) under varying viewing conditions, such as illumination, shading, highlights, and 2) (*distinctiveness*) they should have high *discriminative power*. It has been shown that there exists a trade-off between color invariant models and their discriminative power [15]. For example, color constant models have been proposed [1] which are invariant to all possible light sources assuming a diagonal model for illumination changes. However, such a strong assumption will significantly reduce the discriminative power. For a particular computer vision task that assumes only a few different light sources, color models should be selected which are invariant (only) to these few light sources resulting in an augmentation of the discriminative power of the algorithm. Therefore, the aim is to automatically weight color models to arrive at a proper balance between color invariance (repeatability) and discriminative power (distinctiveness).

The paper is organized as follows: In Section 2, we interpret the method of Markowitz in the context of image processing. In Section 3, we apply the methodology to image feature detection. In Section 4, the method is evaluated empirically. Conclusions are given in Section 5.

2   FUSION OF OBSERVATIONS

First, in Section 2.1, the problem is formulated. Then, in Section 2.2, the weights are derived by exploiting nonperfect correlation between the observation methods. The notion of optimality is given in Section 2.3. Finally, an example is given in Section 2.4.

2.1   Problem Statement

In this paper, we pose the following abstract problem: "Given sets of observations of the same quantity which are expressed in the same unit but which are obtained by N different methods, how are these methods to be combined in order to obtain the most accurate measurement of the process?" Here, we assume no knowledge about the particular observation methods except that the distribution of the observations itself is unimodal. As such, we assume that the observations may be stated as:

$$u = E(u) \pm \sigma_u, \tag{1}$$

where $E(u)$ is the central value for the observation of random variable u (e.g., the average value) and σ_u is a measure of dispersion of u (e.g., the standard deviation).

To solve the problem, we propose combining the observations obtained with the N different methods by the following standard weighting scheme:

$$E(u) = \sum_{i}^{N} w_i E(u_i), \tag{2}$$

where $E(u_i)$ is the estimate of a particular method i as expressed by (1). Taking a linear fusion scheme given by (2) is a choice motivated by its simplicity. Next, the particular problem we address is: "How to derive proper weighting values w_i?" A method to solve this will be discussed in the next section.

2.2   Derivation of Weights by the Exploitation of Nonperfect Correlations

The variance of several combined observation methods follows from (2) by

$$\sigma_u^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j \text{cov}(u_i, u_j), \tag{3}$$

or, equivalently,

$$\sigma_u^2 = \sum_{i=1}^{N} w_i^2 \sigma_{u_i}^2 + \sum_{i=1}^{N} \sum_{j=1, i \neq j}^{N} w_i w_j \text{cov}(u_i, u_j), \tag{4}$$

where w_i denotes the weight assigned to observation method i, u_i corresponds to the average output of observation method i, σ_{u_i} denotes the standard deviation of quantity u_i of observation method i, and, finally, $\text{cov}(u_i, u_j)$ corresponds to the covariance between observation methods i and j.

This latter expression for the standard deviation reveals the usefulness of diversification over various observation methods to reduce the variance due to the correlation that may exist between them. To achieve optimal diversification, we apply the mathematical selection model of Markowitz [14]. The model involves minimizing the standard deviation for a given expected estimate of quantity u or maximizing the expected estimate of a given standard deviation σ_u. To this end, the selection model is given by

$$\text{minimize } \sigma_u \tag{5}$$

under the condition stated by (2). Furthermore, the following conditions are imposed:

$$\sum_{i=1}^{N} w_i = 1, \tag{6}$$

$$-1 \le w_i \le 1, i = 1, \cdots, N. \tag{7}$$

The goal is to minimize the variance imposed by (5) while achieving the minimally acceptable estimate u. Constraint (6) ensures that the available observation methods are fully allocated. Constraint (7) limits the search space when determining values for w_i. Note that it may well be possible that, for observations of the same process, some of the measured quantities obtain negative values, whereas other quantities obtain positive values. By allowing the weights to

take on negative values instead of restricting the weights to take zero values, the selection model can still exploit the negative values to maximize the expected estimate.

The objective function (5) is quadratic with linear constraints. Therefore, the selection problem can be solved by linear programming [16]. When σ_u is varied parametrically, solutions for (5) result in mean-standard deviation pairs of points which represent different weightings of observation methods. The space of attainable mean-standard deviation pairs will be limited and, thus, be bounded by a frontier.

Enlarging the universe of observation methods from which the set of weights is selected will result in a frontier with larger expected values. This is because new observation methods can always be included at a level of zero weighting. As a final step, a single set of weights needs to be selected from the infinitely large universe of weightings making up this frontier. Such selection will, in general, be application dependent.

2.3 The Notion of Optimality

The notion of an *optimal* set of weights can be defined as follows:

1. For any level of standard deviation, consider all the sets of weights which have that standard deviation. From among them all, select the set which has the highest expected output.
2. For any expected output, consider all the sets of weights which have that expected output. From among them all, select the one which has the lowest standard deviation.

Each definition produces a collection of optimal sets of weights. Definition (1) produces an optimal set of weights for each possible level of standard deviation. Definition (2) produces an optimal set of weights for each expected output. Actually, the two definitions are equivalent. The collections of optimal sets of weights obtained using one definition is exactly the same set which is obtained from the other. The collections of optimal sets of weights generate a curved line in the mean-standard deviation plane. This curve will be referred to as the optimal frontier.

2.4 Example

For clarity, an illustration is given where we derive an explicit formula that computes a set of weights on the optimal frontier. The example assumes, for simplicity, that only two hypothetical observation methods D and E exist. As such, the objective is to find the weights w_D and w_E. We solve the problem for a point on the optimal frontier that is of particular interest in many practical applications: The solution yields the maximal ratio between the expected combined output $E(u)$ and the expected standard deviation σ_u. As such, the objective is to maximize an objective function S_p for weight allocation p. The constraint that the weights sum up to 1.0 has to be satisfied, that is, $w_D + w_E = 1$. Therefore, the mathematical problem to be solved is formally written as:

$$\text{maximize } S_p, \tag{8}$$

where

$$S_p = \frac{E(u_p)}{\sigma_p}, \tag{9}$$

subject to constraint (6). This is a standard problem in optimization. In the case of two observation methods, the solution for the weights of the optimal weight allocation is obtained by

$$E(u) = w_D E(u_D) + w_E E(u_E), \tag{10}$$

which follows from (2). Furthermore,

$$\sigma_p^2 = w_D^2 \sigma_D^2 + w_E^2 \sigma_E^2 + 2 w_D w_E \sigma_D \sigma_E \text{cov}(u_D, u_E), \tag{11}$$

which follows from (3). Next, substitute (10) for $E(u_p)$ and (11) for σ_p. Further, substitute $1 - w_D$ for w_E. Then, differentiate the resulting expression for S_p with respect to w_D, set the derivative equal to zero, and, solving w_D, we obtain:

$$w_D = \frac{E(u_D)\sigma_E^2 - E(u_E)\text{cov}(u_D, u_E)}{E(u_D)\sigma_E^2 + E(u_E)\sigma_D^2 - [E(u_D) + E(u_E)]\text{cov}(u_D, u_E)}, \tag{12}$$

$$w_E = 1 - w_D, \tag{13}$$

where the ratio between $E(u)$ and σ_u exceeds the ratio of any other set of weights.

3 APPLICATION TO IMAGE FEATURE DETECTION

In this section, the weighting scheme is applied to fuse color models and image feature detectors. To this end, an algorithm is proposed in Section 3.1 to fuse different color models. In Section 3.2, the weighting scheme is provided to fuse different feature detection methods. A discussion is given in Section 3.3.

3.1 Selection and Fusion of Color Models

To apply the results obtained in the previous section for the fusion of color models, we interpret the variables of (4) as follows: w_i denotes the weight assigned to a particular color plane i, u_i is a central value for the observation of random variable u (which we interpret here as the average pixel output for a target color in color plane i), σ_{u_i} denotes the standard deviation of quantity u of color plane i, and, finally, $\text{cov}(u_i, u_j)$ corresponds to the covariance between color planes i and j.

For training, the following steps are performed:

- For all transformed color planes in a region of interest in a training image, estimate the expected mean $E(u)$, the standard deviation σ_u, and the covariances between all the color planes.
- Use the obtained results to select the optimal weightings of the color planes by the method proposed in Section 2.2.

For testing, perform the following method:

- Apply the weights obtained in the training phase to the standard weighting scheme of (2), thereby combining linearly the color planes to a gray-value image. Let the obtained intensity values in the gray-value image be denoted by $t(x, y)$, where (x, y) are the image coordinates.
- The local zeroth and first order statistics are computed by selecting a window of $M \times M = N$ pixels, which is moved over $t(x, y)$. First, the weighted mean is

computed at (x, y) by a convolution with a Gaussian kernel G^{σ_F} with smoothing parameter σ_F as follows:

$$G(x, y) = t(x, y) \otimes G^{\sigma_F}(x, y). \qquad (14)$$

The mean value is subtracted from each of the pixel values in the window and the result is squared:

$$r(x, y) = \sum_{a=-M/2}^{M/2} \sum_{b=-M/2}^{M/2} (t(x + a, y + b) - g(x, y))^2. \qquad (15)$$

We compute the variance by

$$h(x, y) = r(x, y) \otimes G^{\sigma_F}(x, y). \qquad (16)$$

Finally, the standard deviation v at image coordinate (x, y) is obtained by:

$$v(x, y) = \sqrt{\frac{h(x, y)}{N - 1}}. \qquad (17)$$

- The signal-to-noise image, SNR, is obtained by dividing, at each pixel, the local mean value $g(x, y)$ by the local standard deviation $v(x, y)$:

$$SNR(x, y) = g(x, y)/v(x, y). \qquad (18)$$

  For training colors, the value of SNR will be high.
- Compute the χ-squared error between the signal-to-noise ratio $E(u)/\sigma_u$ based on the training data and the locally measured signal-to-noise SNR value at every pixel (x, y) by

$$\chi(x, y) = \left[SNR(x, y) - \frac{E(u)}{\sigma_u} \right]^2. \qquad (19)$$

- Threshold the χ-squared image and assign pixel values close to zero to the target color, i.e., for training colors, the χ-squared error will attain small values compared to other (arbitrary) colors:

$$C(x, y) = \begin{cases} 1 & \text{if } \chi(x, y) < T \\ 0 & \text{otherwise.} \end{cases} \qquad (20)$$

  The appropriate value of T can be obtained, for example, by the isodata method [17].

In this way, the described algorithm classifies pixels using the training information.

3.2 Selection and Fusion of Feature Detection Algorithms

The application of the results obtained in Section 2 to fuse feature detection algorithms is now straightforward. To this end, we interpret the variables of (4) as follows: w_i denotes the weight assigned to detection algorithm i, u_i denotes the average detection error (in pixels) of algorithm i, σ_u denotes the standard deviation of quantity u of algorithm i, and, finally, $\mathrm{cov}(u_i, u_j)$ corresponds to the covariance between the results obtained with algorithms i and j.

3.3 Discussion

Using the term *optimality* in the context of image processing might be misleading. In this paper, we follow the approach in which different color planes are processed separately before combining individual results. By making the decision to process the color planes separately, we can only claim that our results are optimal within this context. Stated this way, it may very well be that a (true) optimal result could be obtained by processing the color planes simultaneously.

Still, an attractive feature of the fusing algorithm, given by (2), is that it allows for the combination of features coming from very different domains. In this paper, we use the fusing algorithm to combine different image processing algorithms which detect the same feature (e.g., edges or corners). Alternatively, we use the algorithm to combine different color transformation algorithms which transform standard RGB color images into, for instance, the normalized red and green rg planes (invariant to shading and illumination intensity [1]) or the opponent color models RG and YB planes (invariant to highlights, assuming a white light source [18]). These applications are discussed in detail in the next section.

In conclusion, the output of a weighted set of observation methods is the weighted average of its component expected outputs, whereas its standard deviation is less than the weighted average of the component standard deviations. As a consequence, the weighted sets of observation methods of less than perfectly correlated observation methods always offer better variance-output ratio's than the individual component observation methods on their own. In this way, an optimal balance is obtained by the proposed method between repeatability and distinctiveness. The results are verified experimentally in the next section.

4 EXPERIMENTS

In this section, the proposed fusion method is tested on different data sets. First, in Section 4.1, the accuracy of the method to select color models is examined for feature detection. To this end, in Section 4.1.1, color models are automatically selected for image segmentation yielding the highest discrimination power between foreground-background objects. Then, in the context of color edge detection, the performance of the method is tested to provide the highest edge value and minimal variance in Section 4.1.2. Finally, in Section 4.2, the performance of the proposed method is tested to fuse various image feature detection algorithms.

4.1 Selection and Fusion of Color (Invariant) Models

Experiments are conducted on a series of images taken from real-world scenes. For all experiments described in this section, the training RGB images are transformed into the following color channels: First, the normalized red and green channels are, respectively, defined as $r = 255R/(R + G + B)$ and $g = 255G/(R + G + B)$. Then, the opponent color channels, red-green RG and yellow-blue YB, are obtained by $RG = R - G$ and $YB = (2B - R + G)/4$. The CIE $L^*a^*b^*$ color channels are defined by

$$L^* = 116 \left(\frac{Y}{Y_n} \right)^{1/3} - 16,$$

$$a^* = 500 \left[\left(\frac{X}{X_n} \right)^{1/3} - \left(\frac{Y}{Y_n} \right)^{1/3} \right],$$
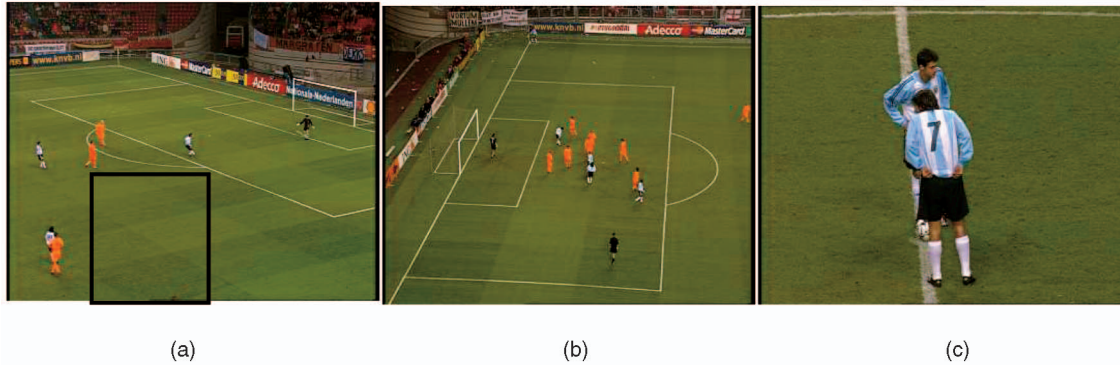
Fig. 1. (a) Training image. The marked section denotes the region of grass used to train the color model. (b) Test image. (c) Test image.

and

$$b^* = 200 \left[\left(\frac{Y}{Y_n} \right)^{1/3} - \left(\frac{Z}{Z_n} \right)^{1/3} \right],$$

where $X_n Y_n Z_n$ is the reference white point. Finally, the saturation S and value V channels are obtained as $V = \max(R, G, B)$ and $S = V - \min(R, G, B)$. The hue color channel is omitted as it is expressed in degrees, which complicates straightforward computation of its mean and variance. All 12 color models are used in the experiments.

These models are selected as they are commonly encountered in color image processing. Further, these color models contain both variant and invariant properties with regard to the imaging conditions. From [1], it is derived that RGB, CIE L^*, and SV are all sensitive to shadows, shading, illumination, and highlights. Further, rg and CIE a^*b^* are invariant to shadows, shading, and illumination intensity. The opponent RG and YB color models are invariant to highlights, assuming a white light source. As these color channels provide both color invariance, i.e., repeatability (rg, CIE a^*b^*, RG, and YB), and variance, i.e., distinctiveness (RGB, CIE L^* and S, and V), it allows us to test whether the proposed method will yield an optimal balance between repeatability and discriminative power by choosing the proper weights for the color channels.

4.1.1 Color Region Detection: Zeroth Order Statistics

The first experiment is conducted on images taken from a football match. The goal is to detect the grass. The task is complicated due varying imaging conditions and the texture of the grass, which cause nonuniform pixel values. The training and test images are shown in Fig. 1. The size of the images is 288×360 pixels.

From the training sample, the correlation coefficients are computed for the various color channels. The coefficients are shown in Table 1. Inspection of the table shows that various color channels show strong negative correlations such as, e.g., the red channel and the normalized colors. The optimal frontier for the grass is illustrated in Fig. 2. From the frontier, the weighting is selected giving the best signal-to-noise ratio, obtained by solving (8). Table 2 shows the weights assigned to the various color channels. The expected intensity value obtained by the weighting method is 91.0 ± 0.3. Here, the standard deviation has a very small value. Next, a small moving window of size $N = 9$ is used. The smoothing value for the Gaussian filter is set to $\sigma_F = 1.5$. The obtained signal-to-noise and chi-squared images are shown in Figs. 3 and 4.

The experiment is repeated for images depicting faces, as shown in Fig. 6. The goal is to detect skin color. The weights assigned to the different color models are specified in Table 2. The optimal frontier is plotted in Fig. 5. The expected intensity value obtained through the weighting scheme is 107.5 ± 1.0. Again, the standard deviation is very small. In Fig. 6a, is the image from which the training data is taken, denoted by the rectangular box. Fig. 6b shows the test image. The images are taken under different imaging conditions. Fig. 6c shows the signal-to-noise image where high values denote large confidence in skin presence.

A third experiment is conducted on a series of images taken from object number 25 of the Amsterdam Library of Images [19]. The image shows a ball with red, green, yellow, and blue colors against a black background. Images are taken under various viewpoints and illuminations. The training and test images are shown in Fig. 7. Three training patches are obtained containing various kinds of blue colors: normal, very dark, and a patch containing highlights. The optimal frontier is shown in Fig. 9. The figure shows that the blue sensor shows very high variance. As a result, the blue color is assigned a weight close to zero by the proposed weighting method, as is shown in Table 2. The expected intensity, after applying (2), for the blue patch is 116.1 ± 0.4. The obtained signal-to-noise image is shown in Fig. 8a. Clearly, the proposed method augments the discriminative power for the blue color.

4.1.2 Color Edge Detection: First Order Statistics

In this experiment, the goal is to detect color transitions of interest. The proposed weighting scheme is used to select the weights such that the highest possible gradient value is selected, while, at the same time, the smallest variance of gradient values is obtained. To achieve color edge detection, training samples are selected manually under varying imaging conditions.

To obtain the training edge set, a single color channel is selected and threshold values are determined manually to obtain the color transition of interest. The values of the edge pixels of the 12 transformed color channels are used to train the optimal weights as described in Section 2. The resulting optimal frontiers in the gradient-standard deviation space is shown as a curve in Fig. 10 for red-cyan and yellow-purple transitions. The precise weights for the color models are shown in Table 3. In Fig. 11, the training image and the test image are shown for the yellow-green color transition. Clearly, training on yellow-green transitions yields the color transition in the test image.

TABLE 1
Correlation Coefficients for Grass

	R	G	B	r	g	RG	YB	L^*	a^*	b^*	S	V
R	1.00											
G	0.96	1.00										
B	0.96	0.95	1.00									
r	-0.83	-0.92	-0.93	1.00								
g	-0.94	-0.86	-0.95	0.80	1.00							
RG	0.31	0.05	0.20	0.14	-0.45	1.00						
YB	0.47	0.45	0.67	-0.66	-0.65	0.13	1.00					
L^*	0.97	0.99	0.96	-0.90	-0.88	0.11	0.45	1.00				
a^*	0.21	-0.04	0.14	0.19	-0.39	0.95	0.19	0.02	1.00			
b^*	-0.31	-0.25	-0.51	0.45	0.51	-0.23	-0.90	-0.27	-0.32	1.00		
S	-0.95	-0.93	-0.99	0.92	0.97	-0.23	-0.70	-0.94	-0.18	0.53	1.00	
V	0.96	1.00	0.95	-0.91	-0.86	0.06	0.45	0.99	-0.04	-0.26	-0.93	1.00

4.2 Selection and Fusion of Image Feature Detection Algorithms

In this section, the goal is to investigate whether robust edge detection is obtained through diversification over different edge detectors. Therefore, edge detection is performed on 1,001 different colors from the PANTONE[1] color system. The 1,001 PANTONE colors are recorded by an RGB-camera (Sony DXC-930P), under a 5,200K daylight simulator (Little Light, Grigull, Jungingen, Germany). Color edges are formed by digitally combining two images into one single image as described by [15]. As such, one yellow image is aligned to the other images, yielding 1,000 different color edges. In this way, the exact location of the color transition is known. As such, statistics for the detection error can easily be obtained. An example image is shown in Fig. 12.

The color information of the red and green channel is discarded in order to illustrate that the proposed method also works on gray-value images. The blue color channel is selected instead of the red or green channel as it is the most noisy channel, which may disturb the detection of the color transition. For each image (i.e., color transition), a single horizontal line scan is taken from the blue color (Fig. 13). The goal is to detect the location of the transition. Images are recorded such that the exact location of the transition is known, yielding the ground-truth.

Three standard detectors are used to compute the color transition. These detectors are:

Thresholding: Thresholds the signal by a chosen value. If the signal value exceeds the threshold value, then the output is set to one, otherwise to zero. Output is the mean grid position.

Edge detection: The signal is processed by a differential filter. The output of the operation is the input to the aforementioned threshold operation.

Erosion and dilation operation (morphology): A mask of a chosen length is used to perform an erosion and dilation operation, after which the difference is taken. The output of the operation is the input to the threshold operation.
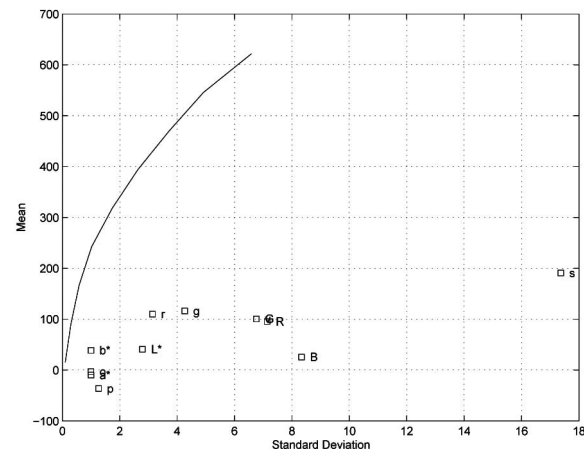

Fig. 2. Mean—standard deviation space for grass. The curved line is the optimal frontier.

TABLE 2
Weights (Expressed in Percentages) Assigned
to Color Models for Detection of Various Targets

Target	R	G	B	r	g	RG	YB	L^*	a^*	b^*	S	V
Grass	-17	-12	21	68	39	0	-6	1	-20	33	-18	12
Skin	92	28	-25	54	73	27	12	16	29	-100	-5	-100
Blue patch	15	7	-15	52	16	-2	3	0	9	-12	26	2

(a)   (b)
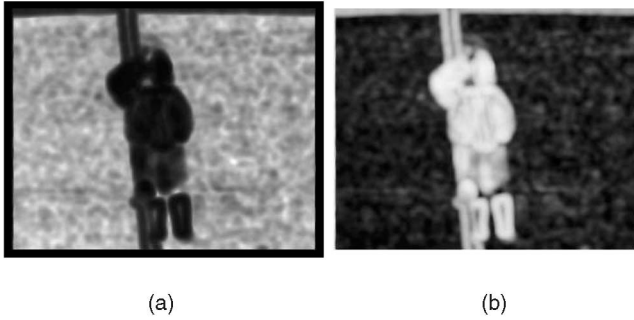
Fig. 3. (a) Output signal-to-noise image and (b) chi-squared image for the first test image.

The filter size of the edge detection algorithm and the size of the structuring element of the morphology operators are comparable. As a final postprocessing step, the error e between the output of each detector and the ground truth is computed as

$$e = x_m - x_i, \qquad (21)$$

with x_m denoting the x position where method m detected the intensity transition and x_i denotes the ideal, golden standard position of the transition.

4.2.1 Determining the Optimal Weighting

To estimate the performance of the edge detectors, a training set consisting of 500 images is used. The data set is processed by the three detectors after which the mean and standard deviation of the error for each detector is computed. The results are listed in Table 4. The table shows that the edge detection method has a very high mean precision, although the standard deviation is quite high. The thresholding and morphology methods both have a systematic error as the locations of the detected color transitions are, in general, biased to the left. Considering the
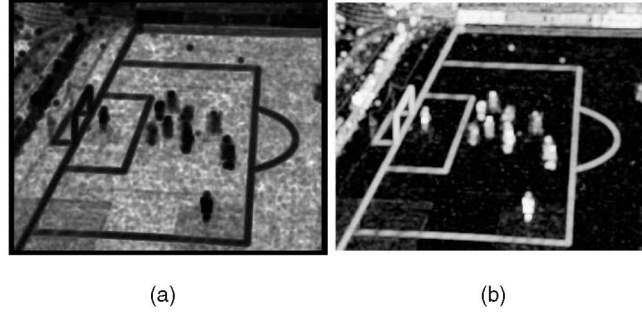

(a)   (b)

Fig. 4. (a) Output signal-to-noise image and (b) chi-squared image for the second test image.
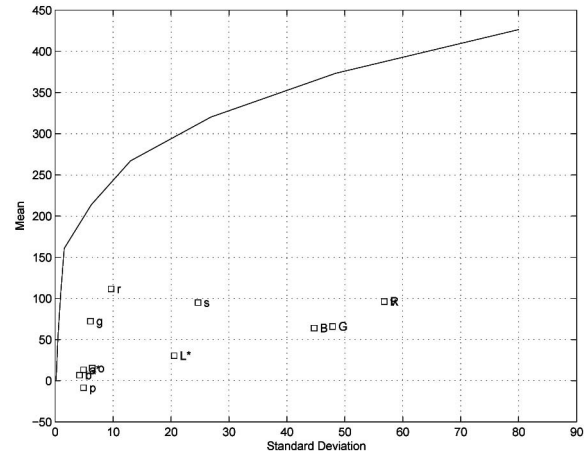

Fig. 5. Mean—standard deviation space for skin. The curved line is the optimal frontier.

mean error, the default way of combining the three detectors would be to equally weight each method.

However, according to the proposed fusion model, the variance of the detectors is also important. Therefore, the autocorrelations are shown in Table 5. From the table, it can be derived that the three detection methods have a relatively low correlation. The correlation between the threshholding detector and the morphology-based method is the lowest. This correlation is important as it indicates that a proper diversification over these edge detectors will yield a significant decrease in the variance. As the expected errors and covariances are known in advance from the training set, the optimal frontier is computed for 10 optimal sets of detectors. The results are shown in Fig. 14. The weightings of the specific sets are given in Table 6. Further, the expected error and the standard deviations are given in Table 7. From the table, it can be derived that a global minimal error

(a)   (b)   (c)

Fig. 6. (a) Training image depicting the region from which training data is obtained. (b) The image used for testing. (c) Signal-to-noise image.
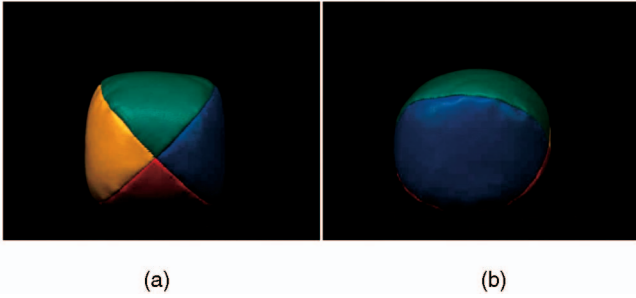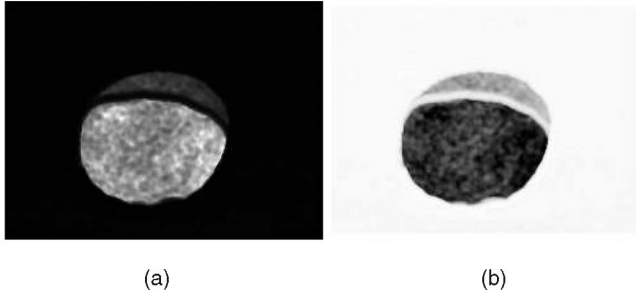
Fig. 7. (a) Training image. (b) Test image.

Fig. 8. (a) Signal-to-noise image. (b) Chi-squared image.

weighting of detectors is obtained by assigning 100 percent of the weight to the edge and thresholding methods which
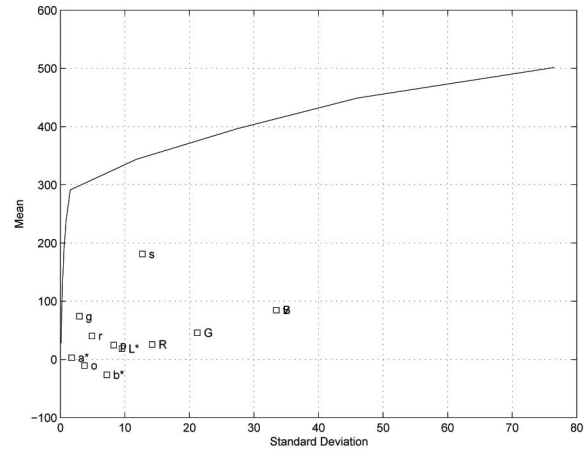

Fig. 9. Optimal frontier for the color blue. The blue color model yields very high variance value and is therefore not a good candidate for detecting the color blue.

corresponds to a assigning w_{edge} and $w_{threshold}$ the value one. Simultaneously, the morphology method is assigned the weight $w_{morphology} = -1$, which is denoted as -100 percent.

4.2.2 Verification of the Weighting Scheme

A new set of five-hundred images is used to verify the obtained weighting factors. The performance of the
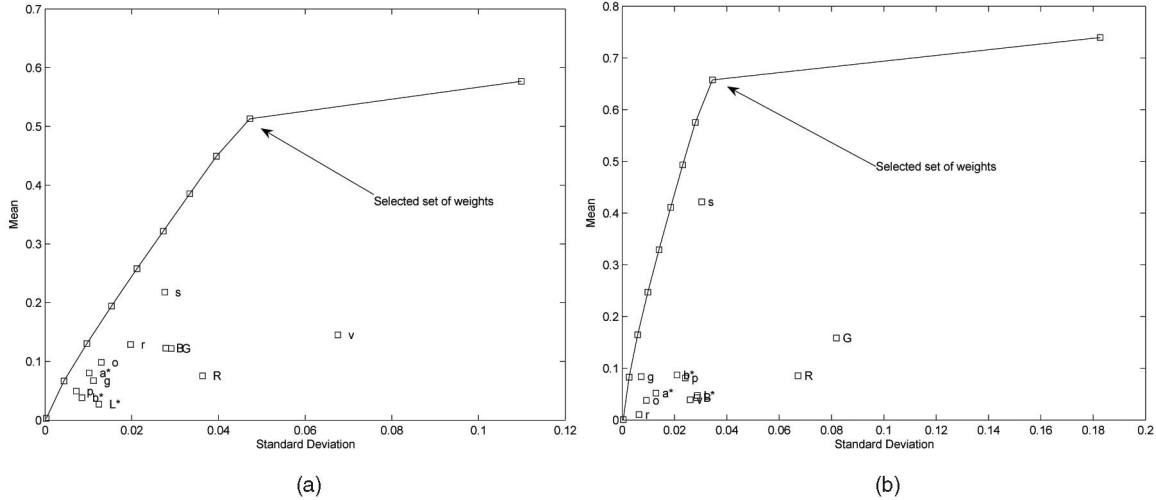

Fig. 10. Mean—standard deviation space for (a) red-cyan color transitions and (b) yellow-purple transitions.

TABLE 3
Weights Expressed as Percentages Assigned by the Proposed Weighting Scheme to Detect Color Transitions

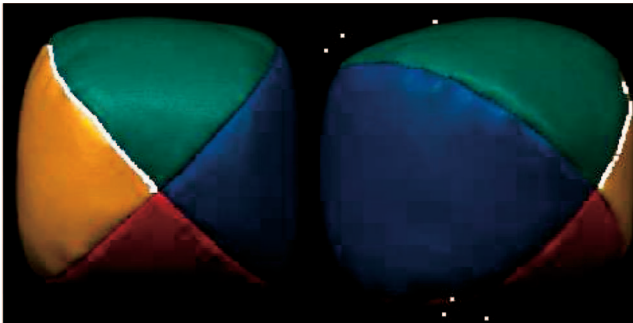| Transition	R	G	B	r	g	RG	YB	L^*	a^*	b^*	S	V
Yellow-Purple	-65	100	63	-100	100	-100	100	-100	-100	100	100	2
Yellow-Blue	100	14	31	100	-45	-100	-100	-100	-100	100	100	100
Red-Cyan	88	100	100	100	-100	95	-100	-100	-100	-100	100	17
Green-Yellow	100	100	-51	72	-100	100	100	-16	-100	-100	-100	95

Fig. 11. Training edges between green and yellow colors are shown on the left. Verification of selected edges is shown on the right. For display purposes, the actual edges are enlarged by a dilation operation.
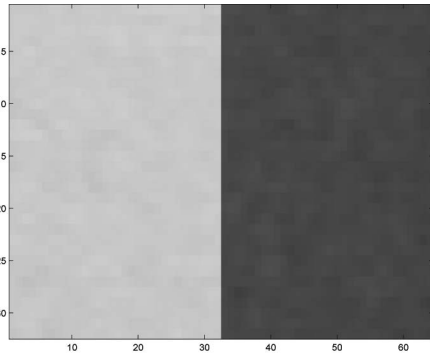

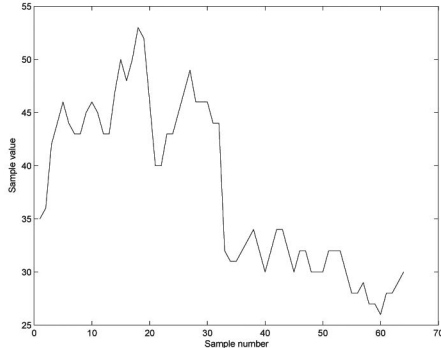Fig. 12. Example image consisting of two homogeneous colors.

Fig. 13. Example of the blue color model for the image shown in Fig. 12.

TABLE 4
Error for Various Detectors for the Training Set

Detector	Error (pixels)
Edge	0.03 ± 1.29
Threshold	-0.23 ± 0.65
Morphology	-0.27 ± 1.22

The edge detection method yields the smallest error.

individual detectors for this new set is shown in Table 8. The results are comparable with the results obtained for the training set as shown in Table 4.

We now determine how the (optimized) weighting sets of detectors perform out-of-sample. The results are computed

TABLE 5
Autocorrelation Values of the
Three Detectors for the Training Set

	Edge	Morphology	Thresholding
Edge	1		
Morphology	0.33	1	
Thresholding	0.28	0.13	1

The detectors are uncorrelated, which means, in theory, that combining of the detectors may significantly reduce the overall variance.
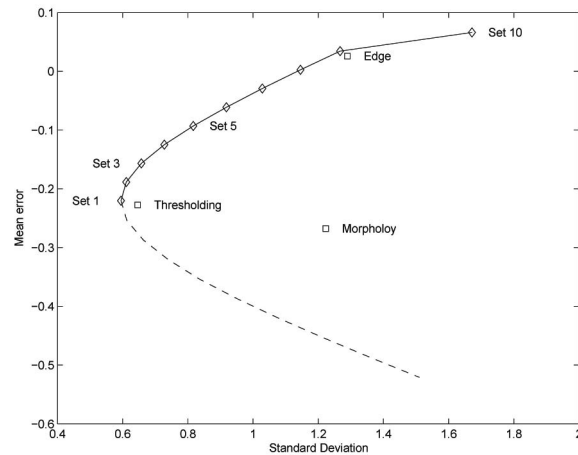

Fig. 14. Error-standard deviation plane. The optimal frontier is obtained by combining three edge detection algorithms.

for the default weighting of algorithms and for the minimal variance detectors. The results are shown in Table 8. The mean error for the minimal variance detector set is 8 percent smaller than for the default weighting scheme. As such, the experiment shows that the fusion model can be successfully applied to decrease the standard deviation of detection algorithms while obtaining a minimal error.

4.3 Discussion of the Results

The proposed method has been tested on a wide variety of applications.

First, in Section 4.1.1, the proposed selection method is evaluated for zeroth order statistics. The correlation coefficients between the different color channels reveal that taking combinations of color channels performs better than taking only a single color space. This is of interest as commonly a single color space is chosen for digital image processing, e.g., the CIE $L^*a^*b^*$ or the HSV space. The obtained weights yield the best ratio between the mean intensity of the training color and the variance thereof. The ratio is exploited through computing signal-to-noise images. The experiments clearly show that the appearance of the training color is boosted significantly.

Second, in Section 4.1.2, the proposed method is evaluated for first-order statistics. The curve of the optimal frontier shows that, for edge detection, the ratio between the gradient strength and the variance thereof is significantly smaller than

TABLE 6
Weights Assigned to the Thresholding, Edge,
and Morphology Detectors, Expressed in Percentages,
for Various Points along the Optimal Frontier

Set	w_{edge} (%)	$w_{morphology}$ (%)	$w_{thresholding}$ (%)
1	5	17	78
2	17	11	72
3	29	5	66
4	40	-0	60
5	52	-6	54
6	64	-11	47
7	76	-17	41
8	87	-22	35
9	99	-28	29
10	100	-100	100

Set 1 yields the smallest variance in the detection, set 8 yields the smallest average error.

TABLE 7
Expected Variance and Errors for Optimal Frontier Sets
of Detectors, as Shown in Fig. 14

Set	Expected Error (pixels)
1	-0.22 ± 0.60
2	-0.19 ± 0.61
3	-0.16 ± 0.66
4	-0.12 ± 0.73
5	-0.09 ± 0.82
6	-0.06 ± 0.91
7	-0.03 ± 1.03
8	0.00 ± 1.15
9	0.03 ± 1.27
10	0.07 ± 1.67

for the case of zeroth order statistics or for the case of combining different feature detectors. As such, the minimal variance set of weights was not selected, but the more extreme weights, yielding very high gradient values. Therefore, Table 3 shows that many color channels are assigned weights close to 100 percent or -100 percent, whereas the weights in Tables 6 and 2 are more evenly assigned.

Third, in Section 4.2, the method is evaluated for different feature detectors. Due to the low correlation between the detectors, combining the detectors yields, in theory, a better performance than using only a single detector. This theoretical result was successfully confirmed in practice. Even more, the proposed diversification method outperformed the default weighting method in practice.

TABLE 8
The First Three Rows Depict the Mean and Standard Deviation of the Error for Various Detectors for the Test Set

Detection Method	Error (pixels)	Range of 3σ error (pixels)
Morphology (no diversification)	-0.27 ± 1.07	$-3.5 \cdots 2.9$
Edge (no diversification)	$+0.30 \pm 0.92$	$-2.5 \cdots 3.1$
Threshold (no diversification)	-0.22 ± 0.64	$-2.1 \cdots 1.7$
Equal weights assigned to the three detectors	-0.06 ± 0.61	$-1.9 \cdots 1.8$
Weights assigned using the proposed method	-0.11 ± 0.53	$-1.6 \cdots 1.4$

The results for the morphology-based detection method, edge detector, and threshold-based method are quite different than those obtained for the training set, as shown in Table 4. The last two rows specify the error for the default and minimal variance edge detection sets in the test set. The experiment shows empirically that diversification over different feature detection methods gives much better results than using only a single detector. Moreover, the proposed weighting method outperforms the default method as 99 percent of all transitions are detected within a range of 3.65 pixels for the default method and 3.21 pixels for the proposed method. Hence, the performance of the proposed method is 12 percent.

5 CONCLUSION

In this paper, a generic selection model has been proposed to select and weight color (invariant) models for discriminatory and robust image feature detection. Further, the method is used to combine image feature detection algorithms. It has been experimentally verified that the selection model can be applied successfully to construct the optimal weighting of color features. Further, the results are successfully applied to edge detection.

The extensive experiments conducted on a wide variety of applications show that the proposed method is widely applicable, yielding an optimal balance between repeatability and distinctiveness. The generic method can be used to achieve both color model weighting as well as the fusion of image feature detectors.

ACKNOWLEDGMENTS

REFERENCES

[1] T. Gevers and A.W.M. Smeulders, "Color Based Object Recognition," *Pattern Recognition,* vol. 32, pp. 453-464, Mar. 1999.

[2] J. Angulo and J. Serra, "Color Segmentation by Ordered Mergings," *Proc. IEEE Int'l Conf. Image Processing,* vol. 2, pp. 125-128, 2003.

[3] L. Macaire, N. Vandenbroucke, and J.G. Postaire, "Color Image Segmentation by Pixel Classification in an Adapted Hybrid Color Space: Application to Soccer Image Analysis," *Computer Vision and Image Understanding,* vol. 90, no. 2, pp. 190-216, 2003.

[4] S. Di Zenzo, "A Note on the Gradient of a Multi-Image," *Computer Vision, Graphics, and Image Processing,* vol. 33, pp. 116-125, 1986.

[5] G. Sapiro and D.L. Ringach, "Anisotropic Diffusion of Multi-valued Images with Applications to Color Filtering," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 18, no. 11, pp. 1582-1586, Nov. 1996.

[6] V. Cardei and B. Funt, "Committee-Based Color Constancy," *Proc. IS&T/SID Seventh Color Imaging Conf.: Color Science, Systems, and Applications,* pp. 311-313, 1999.

[7] N. Cristianini and J. Shawe-Taylor, *Support Vector Machines and Other Kernel-Based Learning Methods.* Cambridge Univ. Press, 2000.

[8] A. Pentland, B. Moghaddam, and T. Starner, "View-Based and Modular Eigenspaces for Face Recognition," *Proc. Conf. Computer Vision and Pattern Recognition,* pp. 84-91, 1994.

[9] S. Stearns, "On Selecting Features for Pattern Recognition," *Proc. Int'l Conf. Pattern Recognition,* pp. 71-75, 1976.

[10] K. Tieu and P. Viola, "Boosting Image Retrieval," *Proc. Conf. Computer Vision and Pattern Recognition,* pp. 228-235, 2000.

[11] A. Whitney, "A Direct Method of Nonparametric Measurement Selection," *IEEE Trans. Computers,* vol. 20, pp. 1100-1103, 1971.

[12] M. Moya and D. Hush, "Network Constraints and Multi-Objective Optimizatiom for One-Class Classification," *Neural Networks,* vol. 9, no. 3, pp. 463-474, 1996.

[13] D.M.J. Tax and R.P.W. Duin, "Uniform Object Generation for Optimizing One-Class Classifiers," *J. Machine Learning Research,* vol. 2, no. 2, pp. 155-173, 2002.

[14] H. Markowitz, "Portfolio Selection," *J. Finance,* vol. 7, 1952.

[15] J.M. Geusebroek, R. van den Boomgaard, and A.W.M. Smeulders, "Color Invariance," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 23, no. 12, pp. 1338-1350, Dec. 2001.

[16] P. Wolfe, "The Simplex Method for Quadratic Programming," *Econometrica,* vol. 27, no. 3, 1959.

[17] T. Ridler and S. Calvard, "Picture Thresholding Using an Iterative Selection Method," *IEEE Trans. Systems, Man, and Cybernetics,* vol. 8, pp. 630-632, 1978.

[18] T. Gevers and H. Stokman, "Robust Histogram Construction from Color Invariants for Object Recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 26, no. 1, pp. 113-118, Jan. 2004.

[19] J.M. Geusebroek, G.J. Burghouts, and A.W.M. Smeulders, "The Amsterdam Library of Object Images," *Int'l J. Computer Vision,* vol. 61, no. 1, pp. 103-112, 2005.

Harro Stokman received the MS degree in 1996 and the PhD degree in 2000, both in computer science from the University of Amsterdam. In the past years, he has built up an extensive track record for industrial knowledge transfer projects. Currently, he holds two part-time positions, one as project leader for the University of Amsterdam and one as principal consultant for a consultancy firm. His scientific interests include video processing and analysis, computer vision, statistical pattern recognition, and statistical learning.

Theo Gevers is an associate professor of computer science at the University of Amsterdam, The Netherlands, and an ICREA research professor at the Computer Vision Center (UAB), Barcelona, Spain. At the University of Amsterdam, he is a teaching director of the MSc in Artificial Intelligence. His main research interests are in the fundamentals of content-based image retrieval, color image processing, and computer vision specifically in the theoretical foundation of geometric and photometric invariants. He is cochair of the Internet Imaging Conference (SPIE 2005, 2006), coorganizer of the First International Workshop on Image Databases and Multi Media Search (1996), the International Conference on Visual Information Systems (1999, 2005), and the Conference on Multimedia and Expo (ICME, 2005). He is the guest editor of the special issue on content-based image retrieval for the *International Journal of Computer Vision* and the special issue on color for image indexing and retrieval for the journal of *Computer Vision and Image Understanding.* He has published more than 100 papers on color image processing, image retrieval, and computer vision. Further, he is a lecturer for postdoctoral courses given at various major conferences (CVPR, ICPR, SPIE, CGIV). He is a member of the IEEE.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/publications/dlib.