

Robust Color Contour Object Detection Invariant to Shadows

Jorge Scandaliaris¹, Michael Villamizar¹, Juan Andrade-Cetto¹, and Alberto Sanfeliu^{1,2}

¹ Institut de Robòtica i Informàtica Industrial (UPC-CSIC)

² Dept. System Engineering and Automation, Universitat Politècnica de Catalunya (UPC)
Barcelona, Spain

Abstract. In this work a new robust color and contour based object detection method in images with varying shadows is presented. The method relies on a physics-based contour detector that emphasizes material changes and a contour-based boosted classifier. The method has been tested in a sequence of outdoor color images presenting varying shadows using two classifiers, one that learnt contour object features from a simple gradient detector, and another that learnt from the photometric invariant contour detector. It is shown that the detection performance of the classifier trained with the photometric invariant detector is significantly higher than that of the classifier trained with gradient detector.

Key words: color invariance, shadow removal, object detection, boosting

1 Introduction

The motivation of this work is to reduce the effect of shadows when detecting objects in a sequence of outdoor images. We show that the photometric invariant used in this paper is more sensitive to the contours of objects that are not shadows while neglecting shadow contours. Our experiments show a comparison between the proposed method and one using image intensity gradient information only.

There have been other approaches for removing shadows from images. Nadimi *et al.* [1] use a multistage approach based on physical models to detect moving shadows in video. Input video frames are passed through a moving object detection stage and then through a series of classifiers which distinguish object pixels from shadow pixels. They show extensive experimental results demonstrating the usefulness of their approach. Salvador *et al.* [2] exploit spectral and geometrical properties of shadows to segment cast shadows from still and moving images. They make initial hypotheses assuming that cast shadows darken the surfaces where they are cast. They further validate the initial hypotheses using complex hypotheses based on color invariance and geometrical properties, to end with an integration stage that confirms or rejects the hypotheses made.

In this work we also focus on a physical model and strive to remove the effects of shadows, but unlike the aforementioned methods we are concerned with producing a contour image invariant to shadows. This is because we base our boosting algorithm for object detection on contour information and having them invariant to shadows greatly improves the robustness of the detection process.

Our objective is to detect objects in image sequences where there are changes in illumination due to the presence of varying shadows. We use a gradient-like image to perform object detection based on contours in a way different to classical methods. Instead of calculating the gradient modulus from the color images, we detect contours that correspond to material changes using a modification to the approach proposed by Geyer *et al.* [3] based on a combination of photometric invariant contours and an automatic local noise-adaptive thresholding.

Boosting algorithms are very well known methods for fast object detection which are based on building robust classifiers from simple (weak) features [4, 5]. We follow the framework addressed in [6], but based on contours instead of intensity images. The use of contour images allows the use of inner and outer object contours to perform robust detection without the drawback of background. Contour features are encoded by Haar operators so that they can be computed in constant time using the intensity integral image. However not all local contours are taken into account for modelling the object as a constellation of Haar operators, since this will require for a large number of weak classifiers. Therefore a learning boosting phase is used in order to select the most discriminant operators and then to linearly combine them for establishing a robust classifier.

To validate our method we have used a sequence of outdoor color images presenting varying shadows to perform object detection. Two boosting classifiers were used, one using simple intensity-based gradient images and the other using the contour images obtained with the proposed method, and their results were compared. We have also compared the effects of shadows in the appearance of spurious contours for intensity-based gradient images and photometric invariant contour images.

2 Robust Physics-Based Contour Detection

2.1 Basic Definitions

Images are the result of complex physical interactions between the light incident over the scene, the surfaces of the objects and the device that acquires the images. Several models of these processes have been developed during the years. One that is commonly used in computer vision applications is the dichromatic reflection model [7]. This model has two terms corresponding to two reflection processes. The light reflected from a surface is a combination of the light reflected at the interface, and the light which enters the substrate and is subsequently reflected back as the result of scattering. It is common to refer to these two reflection components as the interface reflection and the body reflection. The model can be further simplified if the illumination source is assumed to be white or spectrally smooth and the interface reflectance is assumed to be neutral, i.e. the Fresnel reflectance does not depend on wavelength. Under these assumptions the reflection model, expressed in term of the sensor responses, is given by

$$V_k = G_b(\mathbf{n}, \mathbf{s})E \int_{\lambda} B(\lambda)F_k(\lambda) d\lambda + G_i(\mathbf{n}, \mathbf{s}, \mathbf{v})ESF \quad (1)$$

where V_k is the k th sensor response, G_b and G_i are geometric terms denoting the geometric dependencies of the body and surface reflection component, that is sur-

face normal, \mathbf{n} , illumination direction, \mathbf{s} , and viewing direction, \mathbf{v} . $B(\lambda)$ is the surface albedo, E denotes the illumination source, and S denotes the Fresnel reflectance, both assumed independent of λ . $F_k(\lambda)$ denotes the k th sensor spectral sensitivity and $F_k = \int_{\lambda} F_k(\lambda) d\lambda$.

2.2 Color Models

Three color models are used because of their different and complementary properties regarding their response against parameters of the reflection model: RGB , $c1c2c3$ [8] and $o1o2$. In the RGB color model $\{R, G, B\}$ values correspond directly with V_k in (1). The $c1c2c3$ color model is defined by

$$c1(R, G, B) = \arctan(R/\max(G, B)) \quad (2)$$

$$c2(R, G, B) = \arctan(G/\max(R, B)) \quad (3)$$

$$c3(R, G, B) = \arctan(B/\max(R, G)) \quad (4)$$

and the $o1o2$ color model is defined by

$$o1(R, G, B) = (R - G)/2 \quad (5)$$

$$o2(R, G, B) = (R + G)/4 - B/2 \quad (6)$$

It follows from (1) that the RGB color model is sensitive to all parameters of the dichromatic reflection model. Gevers *et al.* [3, 8] showed that under the assumptions included in (1) the $c1c2c3$ color model depends only on the sensor spectral sensitivities and the surface albedo or material for dull objects, being independent of shadows and geometry (E and G_b in the model). $c1c2c3$ still vary in the presence of highlights. They also showed that the $o1o2$ color model is invariant to highlights for shiny objects under the same assumptions. $o1o2$ is still dependent on geometry (G_b). These results are summarized in Table 1.

Table 1. Color model sensitivity to parameters of the image formation process. + denotes sensitivity and - invariance of the color model to a particular parameter.

	shadow geometry material highlights			
RGB	+	+	+	+
$c1c2c3$	-	-	+	+
$o1o2$	+	+	+	-

2.3 Contour Detection

To compute contours we start by calculating the x and y derivatives for each channel of the three aforementioned color models using Gaussian derivatives. Then the color

gradient magnitude for each color model is computed using the Euclidean metric over the various channel derivatives:

$$\nabla C = \sqrt{\sum_{i=1}^N \left[\left(\frac{\partial c_i}{\partial x} \right)^2 + \left(\frac{\partial c_i}{\partial y} \right)^2 \right]} \quad (7)$$

with C representing each color model, N being their dimensionality, and c_i the particular color channels.

The presence of noise in the images can lead to the appearance of maxima in the gradient modulus that are not related to any parameter of the image formation process (1). If it is assumed that the sensor noise is normally distributed, and that we know the parameters for each particular sensor, then using (14) and (15) (see appendix) the uncertainties associated to the $c1c2c3$ and $o1o2$ color models, as well as the different gradient moduli, can be propagated from the a priori known sensor uncertainties. Once we have the associated uncertainty of a measure, we can use it to eliminate noise, as it is shown in Sect. 2.4.

2.4 Contour Invariance

Once the gradient modulus of each color model is available, it is necessary to combine them to obtain the invariance against the undesired parameters on the image formation process.

Gevers *et al.* classified the edges into shadow-geometry, material and highlights [3]. To achieve this they first calculate the gradient magnitude of the RGB , $c1c2$ and $o1o2$ color models. Then they propagate the RGB uncertainties through the color models up to the gradient magnitudes, and local thresholding was used to binarize the gradient magnitudes to obtain C^b . The assumption that the noise is normally distributed implies that 99% of the values fall within a 3σ margin. If the value of the gradient moduli ∇C is greater than $3\sigma_{\nabla C}$ at a particular (x, y) location then the probability of that contour being due to noise is only of 1%:

$$\nabla C^b(x, y) = \begin{cases} 1 & \text{if } \nabla C(x, y) > 3\sigma_{\nabla C(x, y)} \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

with C^b representing each color model used, that is RGB , $c1c2$ and $o1o2$. Finally a rule-base classifier based on the sensitivity of each color model to the dichromatic reflection model parameters, see Table 1, was used to label the resultant image as follows:

```

if  $\nabla C_{RGB}^b \neq 0$  and  $\nabla C_{c1c2}^b = 0$  then
    classify as shadow or geometry edge
else if  $\nabla C_{c1c2}^b \neq 0$  and  $\nabla C_{o1o2}^b = 0$  then
    classify as highlight edge
else
    classify edge as material edge
end if

```

We have taken a different approach to obtain the invariant contour image. Rather than classifying edges according to its physical nature we only pursue to detect object contours that are due to material changes, which are the contours relevant for the task we want to solve. Besides, we have realized that a binarized contour image discards information that might be valuable for the detection stage. Then, the contour image that we calculate is a gray-level image, where the image intensity gives a measure similar to a signal to noise ratio.

As in [3], we calculate the gradient magnitudes of the three color spaces defined in Sect. 2.2 and propagate the *RGB* uncertainties using (15) to obtain the uncertainties associated with the gradient magnitude of each color space, $\sigma_{\nabla C}$. Instead of performing the local thresholding defined in (8) at this stage, we define a function M in the following way:

$$M = \nabla RGB \cdot \nabla c1c2c3 \cdot \nabla o1o2 \quad (9)$$

M will have a maximum value when the gradient moduli of all color models have simultaneously a maximum, and will have low values when the gradient modulus of any of the color models is low. By looking at Table 1 it is evident that the response of M emphasizes material changes in the image, while minimizing those due to shadow-geometry and highlights.

Then, the uncertainty in the function M is also computed using (15) to yield

$$\sigma_M \leq \left(\frac{\partial M}{\partial(\nabla RGB)} \sigma_{\nabla RGB} + \frac{\partial M}{\partial(\nabla c1c2c3)} \sigma_{\nabla c1c2c3} + \frac{\partial M}{\partial(\nabla o1o2)} \sigma_{\nabla o1o2} \right) \quad (10)$$

with $\sigma_{\nabla C}$ being calculated from(15) for each color model

$$\sigma_{\nabla C} \leq \frac{\sum_i \left[\left| \frac{\partial c_i}{\partial x} \right| \cdot \sigma_{\frac{\partial c_i}{\partial x}} + \left| \frac{\partial c_i}{\partial y} \right| \cdot \sigma_{\frac{\partial c_i}{\partial y}} \right]}{\sqrt{\sum_i \left[\left(\frac{\partial c_i}{\partial x} \right)^2 + \left(\frac{\partial c_i}{\partial y} \right)^2 \right]}} \quad (11)$$

with C representing each color model and c_i the particular color channels. The uncertainties $\sigma_{\frac{\partial c_i}{\partial x}}$ and $\sigma_{\frac{\partial c_i}{\partial y}}$ are calculated taking into account that the derivatives are approximated by filtering with a mask, gaussian derivatives in this case. Using (15), it results that the uncertainties can be computed by filtering the uncertainty planes with the absolute value of the mask used for the derivatives. The uncertainty planes are nothing more than the propagation of the *RGB* uncertainties to the other color models using (14).

The assumption that the noise is normally distributed used in Sect. 2.4 is also used here to obtain a local noise-adaptive threshold for removing noisy measurements from M .

$$M' = \begin{cases} M & M > 3\sigma_M \\ 0 & \text{otherwise} \end{cases} \quad (12)$$

The final result is a gray-level contour image that emphasizes the contribution of material changes and at the same time reduces that of shadow-geometry and highlights on the input images. Note that while a value of zero in M' means there is a probability

of 1% of *being wrong*, higher values reduce that probability further. Thus intensity in M' is directly correlated with the probability of a given (x, y) location being a material change.

3 Experiments

In order to evaluate the robustness of the proposed method we have made experiments of detecting an object in a sequence of images, where there are changes in the illumination of the objects due to varying shadows. For the experiments we have used two boosting classifiers that learn contour object features. One boosting classifier uses contour features coming from a simple gradient detector. The second boosting classifier uses our method, that is the photometric invariant contour detector.¹

Using both methods, the learning boosting step selected 100 weak classifiers for each method from a learning set of 50 object images and 200 background images. The background images were extracted from patches of outdoor and indoor images using a randomized process. The learning set of images included objects with small variations of position and scale which make this classifier robust against small object transformations.

The two boosting classifiers were tested over a sequence of 934 images where one static object under varying shadows appears. Some frames are shown in Figure 1, where we can also appreciate some detection results. The input color images are on the left column, the results of the classifier based on the simple gradient detector are on the middle column, and the results of the proposed method are on the right column. Each green square in the images represents one object detection. We can see how the classical method is perturbed by the varying shadows, being unable to detect the object under these illumination variations. On the other hand, the classifier based on photometric invariant contours achieves a correct detection thanks to the shadow-free contours, being the contour object features reliable over the sequence. The Figure 3 shows the ROC (*Receiver Operating Characteristic*) curves for both methods. The proposed method overcomes the classical one, achieving 100% detection, meanwhile the simple gradient method performs 79% without false positives.

In Figure 4, there are some test frames when the threshold in boosting classifier is reduced from $\beta = 0.75$ (Figure 1) to $\beta = 0.65$. We can notice that the object is detected in all cases though with many false positives for the method based on simple gradients, unlike the proposed method that continues detecting correctly, demonstrating its robustness and reliability with a relaxed classification threshold.

We have evaluated the influence of shadows both in the proposed photometric invariant contour images and the simple gradient images. The motivation of the experiment was to test the relative increase in contour pixels caused by shadows. To this purpose, a subset of the outdoor image sequence consisting of 87 images from the total of 934 was randomly selected. This subset was enlarged by 3 images selected with the constraint of being shadow free. Because this constraint could not be met over the entire image, the images were cropped around the carton recycle bin used as target object in the detection experiments.

¹ Information about the boosting classifier can be found in [6].

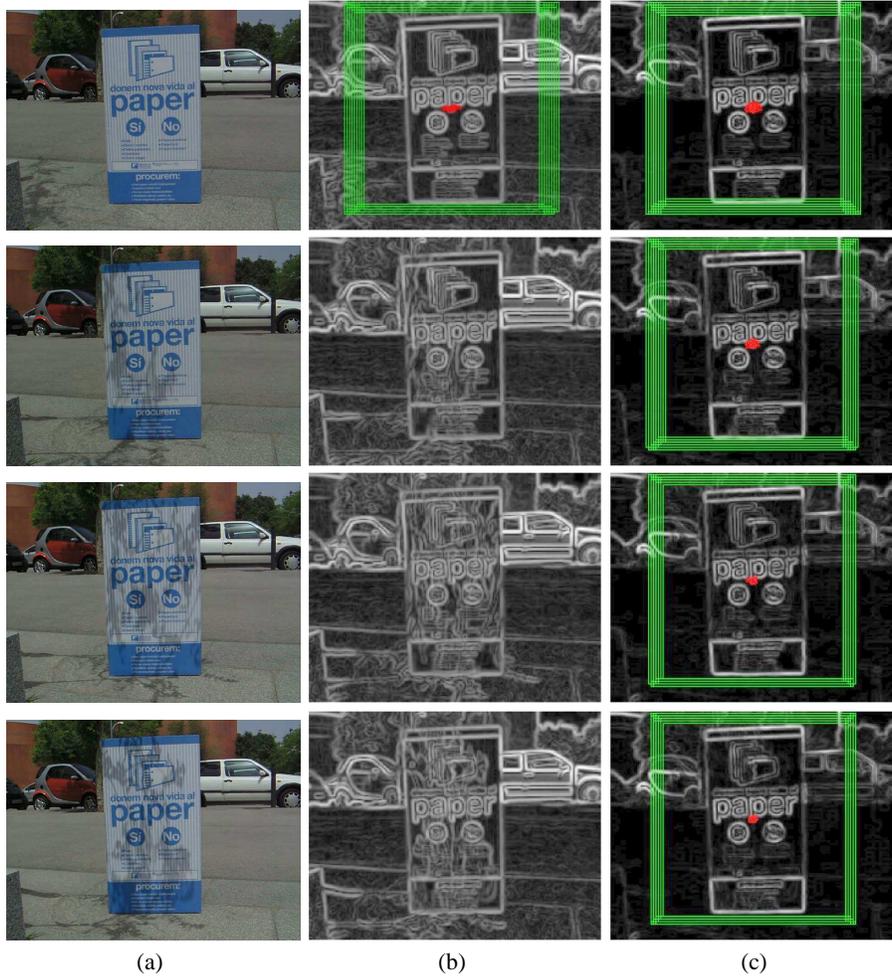


Fig. 1. Test 1. Frames when the threshold β in the boosting classifier is set to 0.75. (a) Input images, (b) gradient based detections and (c) invariant based detections.

The images were processed and both the gradient and the photometric invariant contour images were obtained. One shadow free contour image was visually inspected for each method. The inspection determined a threshold for each method that produced a binarized image with a similar aspect between them and good correlation between object features and contours. This shadow-free image was used as the reference image, and then the following metric was used for all images

$$E_{bd} = \text{count}(\text{abs}(\Delta I) > k) / E_r \quad (13)$$

with $\Delta I = I_i - I_r$ the difference between a given image and the reference, k the aforementioned thresholds, one for each method and $E_r = \text{count}(I_r > k)$ the number of contours in the reference image. E_{bd} is the ratio of the number of incorrectly detected contours to the number of contours in the reference image. The results are shown in Fig. 2. As can be seen, the ratio is small and stable for the proposed method, with a maximum number of misdetections of around 3% of the number of contours in the reference image. For the gradient contour image the ratio is unstable and the number of misdetections ranges from 35 to 65 percent of the number of contours in the reference image. This simple metric clearly shows the impact of the proposed method in the presence of shadows.

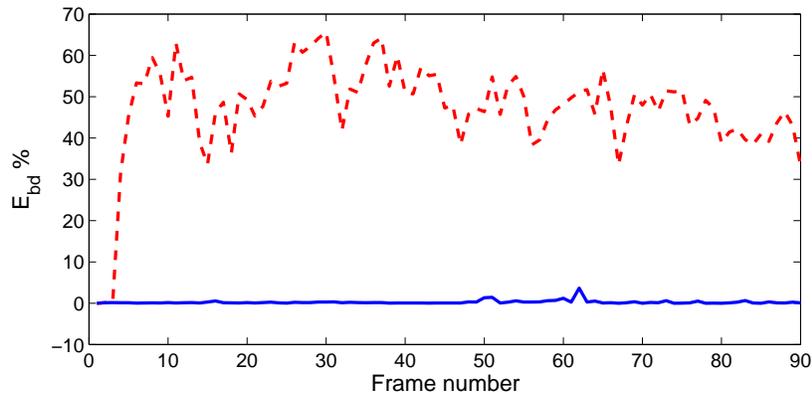


Fig. 2. Percent ratio of the number of misdetections to the number of contours in the reference image for the photometric invariant contour image (—), and for the simple gradient contour image (- -).

4 Conclusions

The proposed method has demonstrated to perform robust object detection in outdoor images under varying shadows and illumination changes, overcoming the classical method relying on a simple gradient detector. This latter method fails due to varying

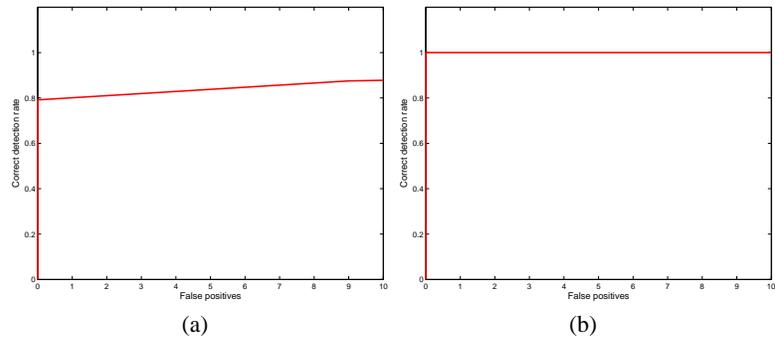


Fig. 3. ROC curves. (a) Simple gradient and (b) photometric invariant contours.

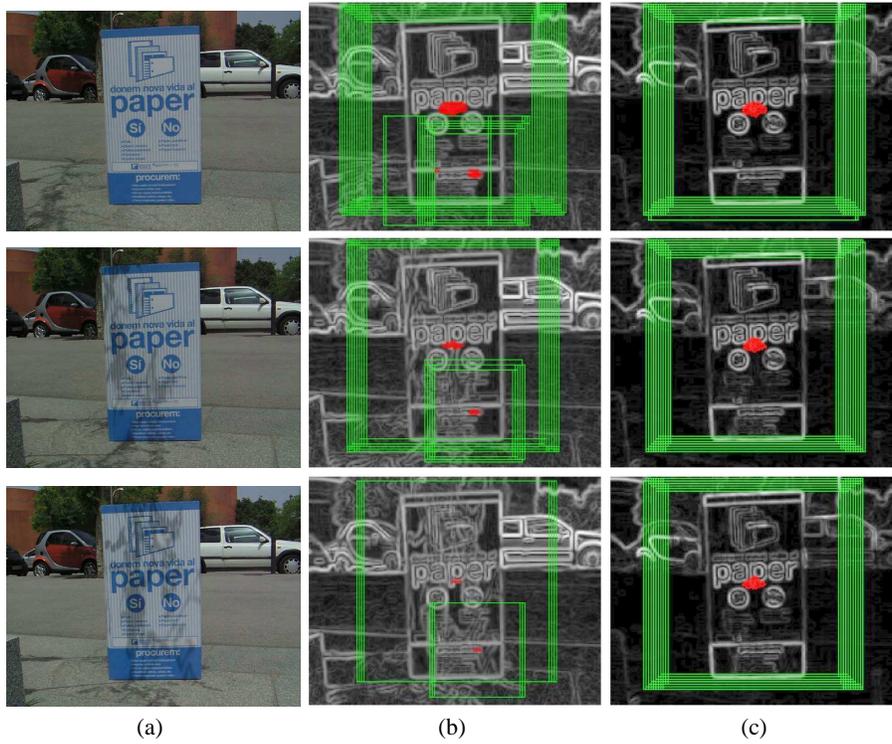


Fig. 4. Test 2. Frames when the threshold β in the boosting classifier is reduced to 0.65. (a) Input images, (b) gradient based detections and (c) invariant based detections.

shadows producing *new* object contour features that eventually mislead the classifier. On the other hand, the proposed method based on photometric invariant gives a contour image without shadow effects. This facilitates the identification task, as the classifier focuses on actual object features. The experiments showed the usefulness of the shadow invariance of the method in a sequence of outdoor images for object detection where the illumination conditions were not controlled.

Acknowledgements This work is supported in part by the Spanish Ministry of Education and Science under project DPI 2004-5414 and the EU URUS project FP6-IST-045062 to JS, MV, JAC and AS; by the Generalitat of Catalonia's Department of Education and Universities and the European Social Fund to JS; and by the Technical University of Catalonia to MV. JAC is a Ramón y Cajal Postdoctoral Fellow.

References

1. Nadimi, S., Bhanu, B.: Physical models for moving shadow and object detection in video. *IEEE Trans. Pattern Anal. Machine Intell.* **26**(8) (Aug. 2004) 1079–1087
2. Salvador, E., Cavallaro, A., Ebrahimi, T.: Cast shadow segmentation using invariant color features. *Comput. Vis. Image Und.* **95**(2) (Aug. 2004) 238–259
3. Gevers, T., Stokman, H.: Classifying color edges in video into shadow-geometry, highlight, or material transitions. *IEEE Trans. Multimedia* **5**(2) (Jun. 2003) 237–243
4. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: *Proc. 15th IEEE Conf. Comput. Vision Pattern Recog., Kauai* (Dec. 2001) 511–518
5. J. Shotton, A. Blake, R.C.: Multiclass object recognition with sparse, localized features. In: *Proc. IEEE Int. Conf. Comput. Vision, Beijing* (Oct. 2005) 503 – 510
6. Villamizar, M., Sanfeliu, A., Andrade-Cetto, J.: Computation of rotation local invariant features using the integral image for real time object detection. In: *Proc. 18th IAPR Int. Conf. Pattern Recog. Volume 4., Hong Kong, IEEE Comp. Soc.* (Aug. 2006) 81–85
7. Shafer, S.A.: Using color to separate reflection components. *Color Res. Appl.* **10**(4) (1985) 210–218
8. Gevers, T., Smeulders, W.M.: Color based object recognition. *Pattern Recogn.* **32** (1999) 453–464
9. Taylor, J.R.: *An introduction to error analysis.* 2nd edn. University Science Books, Sausalito, CA (1997)

Appendix: Error Propagation

Suppose that x, \dots, z are measured values with uncertainties $\sigma_x, \dots, \sigma_z$ and the measured values are used to compute the function $q(x, \dots, z)$. If the uncertainties in x, \dots, z are independent and random, then the uncertainty in q is [9]

$$\sigma_q = \sqrt{\left(\frac{\partial q}{\partial x} \sigma_x\right)^2 + \dots + \left(\frac{\partial q}{\partial z} \sigma_z\right)^2} \quad (14)$$

In any case the uncertainty is never larger than the ordinary sum

$$\sigma_q \leq \left|\frac{\partial q}{\partial x}\right| \sigma_x + \dots + \left|\frac{\partial q}{\partial z}\right| \sigma_z \quad (15)$$