Author manuscript, published in "International Conference on Image and Video Retrieval (CIVR '06) 4071 (2006) 61--71"
DOI : 10.1007/11788034_7



# Corner detectors for affine invariant salient regions: Is color important?

Nicu Sebe[1], Theo Gevers[1], Joost van de Weijer[2], Sietse Dijkstra[1]

[1]Faculty of Science, University of Amsterdam, The Netherlands
[2]INRIA Rhone-Alpes, France

**Abstract.** Recently, a lot of research has been done on the matching of images and their structures. Although the approaches are very different, most methods use some kind of point selection from which descriptors or a hierarchy are derived. We focus here on the methods that are related to the detection of points and regions that can be detected in an affine invariant way. Most of the previous research concentrated on intensity based methods. However, we show in this work that color information can make a significant contribution to feature detection and matching. Our color based detection algorithms detect the most distinctive features and the experiments suggest that to obtain optimal performance, a tradeoff should be made between invariance and distinctiveness by an appropriate weighting of the intensity and color information.

## 1 Introduction

Corner detection can be traced back to Moravec [2] who measured the average change of intensity by shifting a local window by a small amount in different directions. Harris and Stephens [3] improved the repeatability of Moravec detector under small image variations and near edges. By an analytic expansion of the Moravec detector the local autocorrelation matrix is derived using first order derivatives. The Harris detector, in combination with a rotational invariant descriptor, was also used by Schmid and Mohr [4] when they extended local feature matching to general object recognition.

A low-level approach to corner finding is proposed by Smith and Brady: the SUSAN detector [5]. Their corner detector compares the intensity of a pixel with the intensities of neighboring pixels. If few of the neighboring pixels have approximately the same value, the center pixel is considered a corner point.

Lindeberg [6] proposed an "interesting scale level" detector which is based on determining maxima over scale of a normalized blob measure. The Laplacian-of-Gaussian (LoG) function is used for building the scale space. Mikolajczyk [1] showed that this function is very suitable for automatic scale selection of structures. An efficient algorithm for use in object recognition was proposed by Lowe [7]. This algorithm constructs a scale space pyramid using difference-of-Gaussian (doG) filters. The doG are used to obtain an efficient approximation of the LoG. From the local 3D maxima a robust descriptor is build for matching purposes. The disadvantage of using doG or LoG is that the repeatability is not optimal since they not only respond to blobs, but also to high gradients in one direction. Because of this, the localization of the features may not be very accurate.

An approach that intuitively arises from this observation, is the separation of the feature detector and the scale selection. The original Harris detector [3] shows to be robust to noise and lighting variations, but only to a very limited extend to scale changes [8]. To deal with this Dufournoud et al. [9] proposed the scale adapted Harris operator. Given the scale adapted Harris operator, a scale space can be created. Local 3D maxima in this scale space can be taken as salient points. Mikolajczyk points out that the scale adapted Harris operator rarely attains a maximum over scales [1]. This results in very few points, which are not representative enough for the image. To address this problem, Mikolajczyk [1] proposed the Harris-Laplace detector that merges the scale-adapted Harris corner detector and the Laplacian based scale selection.

All the approaches presented above are intensity based. Since the luminance axis is the major axis of color variation in the RGB color cube, most salient points are found using just intensity. The additional color based salient points might not dramatically increase the number of salient points. The distinctiveness of these color based salient points is however much larger, and therefore color can be of great importance when matching images. Furthermore, color plays an important role in the pre-attentive stage in which features are detected. This means that the saliency value of a point also depends on the color information that is present. Very relevant to our work is the research of van de Weijer et al. [10]. They aim at incorporating color distinctiveness into the design of salient point detectors. In their work, the color derivatives form the basis of a color saliency boosting function since they are used in both the detection of the salient points, and the determination of the information content of the points. Furthermore, the histograms of color image derivatives show distinctive statistical properties which are used in a color saliency boosting function.

Our contribution is twofold. First of all, we are comparing the Harris corner detector used by Mikolajczyk [11] with the SUSAN corner detector [5]. Second we are investigating the use of color in extracting corners. We first used the color extended Harris detector [10] which operates on the same principle as the intensity based Harris detector. The extension to color consists of a transformation of the color model to decorrelate common photometric variations and a saliency boosting function that takes into account the statistics of color image derivatives. Later, we investigate the use of invariant color ratios and we show that by using color information the distinctiveness of the regions is increased, whereas the desirable properties are preserved. The incorporation of color information however increases the detection complexity.

## 2 Corners detectors in the affine invariant framework

In the affine invariant region detection algorithm [1], an initial point with a corresponding detection scale is assumed. Based on the region defined by the initial location and scale, the point is subject to an iterative procedure in which the parameters of the region are adapted until convergence is reached. The affine invariance is obtained by combining a number of existing algorithms. The characteristic scale of a structure is selected using the Laplacian scale selection. The location of a region is determined using the Harris corner detector, and the affine deformation of a structure is obtained by using certain properties of the second moment matrix. Because all parameters (scale, location, and shape) influence each other, they all need to be adjusted in every iteration. If the algorithm converges towards a stable region, the adjustments become smaller. If they become small enough the algorithm halts, and the next initial region is processed. More details on the framework can be found in [11, 1]. Note that we are using this framework as a baseline and we "plug-in" several corner detectors (Harris [3], SUSAN [5], and two color variants of Harris corner detector).

To extend the Harris detector to incorporate color information, the second moment matrix used in the detector should be based on color information. Because of common photometric variations in imaging conditions such as shadows and illumination, we use two invariant color spaces i.e. the opponent color space [10] for the colOppHarris detector and the $m$-color ratio space [12] for the colRatHarris detector. The reason for choosing these color spaces is to investigate whether color invariance plays a role in the repeatability and distinctiveness of the detectors. It has been shown that there exists a trade-off between color invariant models and their discriminative power [12]. While the opponent color space has limited invariance and the intensity information is still present, the color ratio is independent of the illumination, changes in viewpoint, and object geometry [12].

The second moment matrix is computed as follows. The first step is to determine the gradients of $RGB$ by using a convolution with the differentiation kernels of size $\sigma_D$. The gradients are then transformed into the desired color system (i.e. opponent or color ratio system). By the multiplication and summation of the transformed gradients, the components of the second moment matrix are computed. The values are averaged by a Gaussian integration kernel with size $\sigma_I$. Scale normalization is done again using a factor $\sigma_D^2$. This procedure is shown in Eq. 1 where a general notation is used. Color space $C$ is used with its components $[c_1, \ldots, c_n]^T$, where $n$ is the number of color system components and $c_{i,x}$ and $c_{i,y}$ denote the components of the transformed $RGB$ gradients, with $i \in [1, \ldots, n]$, and the subscript $x$ or $y$ indicating the direction of the gradient.

$$\mu(\mathbf{x}) = \sigma_D^2 g_{\sigma_I} \begin{bmatrix} C_x^T(\mathbf{x})C_x(\mathbf{x}) & C_x^T(\mathbf{x})C_y(\mathbf{x}) \\ C_x^T(\mathbf{x})C_y(\mathbf{x}) & C_y^T(\mathbf{x})C_y(\mathbf{x}) \end{bmatrix} \tag{1}$$

If the distribution of the transformed image derivatives is observed for a large set of images, regular structures are formed by points of equal frequency [10]. The planes of these structures are called isosalient surfaces. These surfaces are formed by connecting the points in the histogram that occur the same number of times. Based on the observed statistics a saliency measure can be derived in which vectors with an equal information content have an equal effect on the saliency function. This is called the color saliency boosting function which is based on rotation and normalization [10]. The components of the second moment matrix that incorporate the rotation and normalization, are shown in Eq. 2. The components of the rotated transformed image derivatives are denoted by $\tilde{c}_{i,x}$ and $\tilde{c}_{i,y}$. The normalization of the ellipsoid is done using the diagonal matrix $\Lambda$.

$$\begin{aligned} C_x^T(\mathbf{x})C_x(\mathbf{x}) &= \sum_{i=1}^{n} \Lambda_{ii}^2 \tilde{c}_{i,x}^2(\mathbf{x}, \sigma_D) \\ C_x^T(\mathbf{x})C_y(\mathbf{x}) &= \sum_{i=1}^{n} \Lambda_{ii}^2 \tilde{c}_{i,x}(\mathbf{x}, \sigma_D)\tilde{c}_{i,y}(\mathbf{x}, \sigma_D) \\ C_y^T(\mathbf{x})C_y(\mathbf{x}) &= \sum_{i=1}^{n} \Lambda_{ii}^2 \tilde{c}_{i,y}^2(\mathbf{x}, \sigma_D) \end{aligned} \tag{2}$$

Note that in the case of color ratios, the derivatives are already incorporated in the way the ratios are computed. A brief description is given below.

We focus on the following color ratio [12]:

$$M(c^i_{\boldsymbol{x}_1}, c^i_{\boldsymbol{x}_2}, c^j_{\boldsymbol{x}_1}, c^j_{\boldsymbol{x}_2}) = \frac{c^i_{\boldsymbol{x}_1} c^j_{\boldsymbol{x}_2}}{c^i_{\boldsymbol{x}_2} c^j_{\boldsymbol{x}_1}}, c^i \neq c^j, \tag{3}$$

expressing the color ratio between two neighboring image locations $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, for $c^i, c^j \in C$ giving the measured sensor response obtained by a narrow-band filter with central wavelengths $i$ and $j$.

For a standard $RGB$ color camera, we have:

$$m_1(R_{\boldsymbol{x}_1}, R_{\boldsymbol{x}_2}, G_{\boldsymbol{x}_1}, G_{\boldsymbol{x}_2}) = \frac{R_{\boldsymbol{x}_1} G_{\boldsymbol{x}_2}}{R_{\boldsymbol{x}_2} G_{\boldsymbol{x}_1}} \tag{4}$$

$$m_2(R_{\boldsymbol{x}_1}, R_{\boldsymbol{x}_2}, B_{\boldsymbol{x}_1}, B_{\boldsymbol{x}_2}) = \frac{R_{\boldsymbol{x}_1} B_{\boldsymbol{x}_2}}{R_{\boldsymbol{x}_2} B_{\boldsymbol{x}_1}} \tag{5}$$

$$m_3(G_{\boldsymbol{x}_1}, G_{\boldsymbol{x}_2}, B_{\boldsymbol{x}_1}, B_{\boldsymbol{x}_2}) = \frac{G_{\boldsymbol{x}_1} B_{\boldsymbol{x}_2}}{G_{\boldsymbol{x}_2} B_{\boldsymbol{x}_1}}. \tag{6}$$

Taking the natural logarithm of both sides of Eq. 4 results for $m_1$ (a similar procedure is used for $m_2$ and $m_3$) in:

$$\ln m_1(R_{\boldsymbol{x}_1}, R_{\boldsymbol{x}_2}, G_{\boldsymbol{x}_1}, G_{\boldsymbol{x}_2}) = \ln(\frac{R_{\boldsymbol{x}_1} G_{\boldsymbol{x}_2}}{R_{\boldsymbol{x}_2} G_{\boldsymbol{x}_1}}) = \ln(\frac{R_{\boldsymbol{x}_1}}{G_{\boldsymbol{x}_1}}) - \ln(\frac{R_{\boldsymbol{x}_2}}{G_{\boldsymbol{x}_2}})$$

Hence, the color ratios can be seen as differences at two neighboring locations $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ in the image domain of $\ln(R/G)$:

$$\nabla_{m_1}(\boldsymbol{x}_1, \boldsymbol{x}_2) = (\ln(\frac{R}{G}))_{\boldsymbol{x}_1} - (\ln(\frac{R}{G}))_{\boldsymbol{x}_2}. \tag{7}$$

Differentiation is obtained by computing the difference in a particular direction between neighboring pixels of $\ln R/G$. The resulting derivation is independent of the illumination color, changes in viewpoint, the object geometry, and illumination intensity. To obtain the gradient magnitude, the Canny's edge detector is taken (derivative of the Gaussian with $\sigma = 1.0$).

## 3 Experiments

In this section we compare the different corner detectors according to three criteria: repeatability, information content, and complexity. We are interested in comparing the intensity and color based detectors and in investigating the role color invariance plays in the performance of color based detectors.

### 3.1 Repeatability

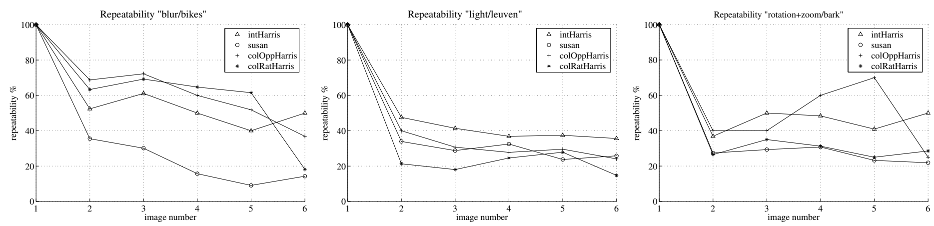
The repeatability is measured by comparing the regions that are detected in an image $I_R$, and in a transformed copy of it, $I_L$. The localizations and shapes of the structures in the images are related by a homography $H$. By comparing the correspondences between the detected regions that cover the same part of the depicted scene, the repeatability rate can be computed as [1]:

$$r = \frac{n_m}{\min(n_R, n_L)} \times 100\%$$

where $n_R$ is the number of regions in the common part of $I_R$, $n_L$ is the number of regions in the common part of $I_L$, and $n_m$ is the number of matches.

In order to determine the robustness of the detectors, the repeatability is measured under common variations in the imaging conditions. For each transformation the detectors are evaluated using a set of images in which in every successive image the transformation is increased a little. The dataset used, is the one used in [1] for determining the repeatability. Test sets are provided for blur, zoom & rotation, viewpoint changes, light changes, and JPEG compression. All images are taken with a digital camera that introduced JPEG compression artifacts.

**Blur** The blur testset consists of two sets of 6 images. In both sets the focus of the camera is gradually varied from sharp in the first image to unsharp in the last image. The successive images are also translated.




(a) (b) (c)

**Fig. 1.** Repeatability for different detectors on the blur set (a), the lighting set (b), and the rotation and scaling set (c)

For most of the images (see Fig. 1(a)), the color-based Harris detectors performed best. The intensity based Harris detector performs about 10% worse for images 2 to 5. The SUSAN based detector performs worse over the whole set of images. This poor performance might be due to the scale of the detectable structures that increases as the images get more blurred. The localization of the SUSAN based detector gets worse as the scale increases. Note that the color Harris detectors have similar repeatability and they only need a fraction of the number of regions that the other detectors need to achieve a similar repeatability.

**Lighting** In this test set the aperture of the camera is varied. This results in a sequence of images that range from light to dark. In theory, this should only affect the intensity component, but since the camera pre-processes the image additional variations might be present. The successive images in this set are also slightly translated. In the test set only the intensity is changed and no other lighting changes like shadow and highlights are present.

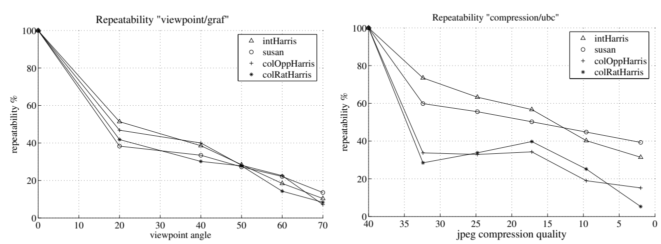
The intHarris detector performs best on this test set (Figure 1(b)). This is probably due to the fact that the Harris corner detector is based on the derivatives of the image instead of on the actual pixel values. The SUSAN detector uses its brightness threshold to determine whether something qualifies as a corner. If the overall image intensity gets lower, the variations in brightness also get lower. As a result the SUSAN detector will pick up less corners. The repeatability of the colOppHarris based detector is similar to that of the SUSAN detector, although it is also based on derivatives. ColRatHarris detector performs the worst probably due to the invariant properties imposed on it. The number of regions needed is the highest for the intHarris detector, whereas the SUSAN and color Harris detectors need a lower number of regions.

**Rotation and scaling** Invariance against rotation and scaling is very important in detecting the same regions in different images of the same scene. Any multi-scale interesting point detector should have good results on this.

The “bark” test set consists of a number of rotated and zoomed images depicting a natural structure. Although corners and edges are present, most of them are found in the texture. Color information is present, be it in a modest way.

The colOppHarris based detector performs best (Figure 1(c)). This might be due to the fact that it only detects 10 regions in the reference image; which might be too few for matching. The intensity based Harris detector also performs well, using more regions. The SUSAN based detector needs the most regions and achieves the lowest repeatability comparable to the one of colRatHarris detector. Note again that by using a more invariant color space (as is the case for colRatHarris detector) we tend to lose in repeatability performance.

**Viewing angle** The “graffiti” test set depicts a planar scenes from different viewpoints and its images contain regions of equal color that have distinctive edges and corners. The images bear similarities to synthetic images as those images in general also have sharp edges and colors with high saturation.




(a) (b)

**Fig. 2.** Repeatability for different detectors on the viewpoint set (a) and the compression set (b)

The repeatability results are shown in Figure 2(a). All detectors perform similar. Overall the repeatability of the SUSAN detector is a few percents lower than those of the other detectors. Again, the number of regions used by the color Harris detectors to achieve this repeatability, is much lower than that of the other detectors.

**JPEG compression** The JPEG compression artefacts introduced are all rectangular regions. These rectangles introduce many additional corners in the image. All salient point detection methods used in the experiments rely on corners. Therefore, these artefacts might have a significant impact on the repeatability. When dealing with color and JPEG compression it is important to know that the lossy compression aims at discarding information the human cannot easily see. The human eye is more sensitive to variations in intensity than to variations in color. This is therefore also used in the JPEG compression scheme.

In the test set the reference image is compressed at a quality of 40%. In the successive images the quality is decreased to 2%. Note that most JPEG compressed images are compressed at a higher quality level; low quality values like these are in practice only used under special circumstances like low bandwidth video.

The intHarris and SUSAN detectors perform similar under compression in this test set, as is shown in Figure 2(b). The intensity based detectors deal significantly better with the artefacts than the color Harris detectors do. The color Harris detectors are clearly confused by the (color) artefacts introduced by the high JPEG compression. This might be due to the fact that the JPEG encoding is more conservative in varying the intensity information than it is with varying the color information of an image.

**Discussion** Overall, from the experiments can be concluded that regions can be detected more reliable under some transformations when using just intensity information. The opponent color model that is used in the colorOppHarris detector decorrelates the intensity and color information. It is possible that by varying the ratio of these two different components, the tradeoff between invariance and distinctiveness can be made. If the weighting of the intensity component is increased, probably more regions are detected. Although the information content of these additional regions might not be very high, they can be detected reliable under varying imaging conditions. On the other extreme, the colRatHarris detector does not use any intensity information and this is reflected in the poor results under most of the transformations. However, this is compensated by a higher distinctiveness of the descriptors as it will be illustrated in the next section.

### 3.2 Information content

The information content of the detected regions is measured using the entropy. The entropy is defined as the average information content of all "messages". The information content of one message $i$ can be computed from the probability of the message $p_i$ according to $I_i = -\log(p_i)$. From the information content of one message, the average information content of all messages can be derived. The entropy of a set of messages is therefore defined as $I = -\sum_i p_i \log(p_i)$.

To estimate the entropy of a number of detected regions, the regions need to be described. In this context, the descriptor of a region acts as the "message" in the entropy estimation. There are numerous ways of describing a region; in this research two common methods are used to describe regions. Both methods are based on convolutions with Gaussian derivatives.

A method to describe a region using derivatives is the "local jet" of order $N$ at point $\mathbf{x}$ [13]. In this research rotational differential invariants up to the second order are used to create the intensity based descriptor $\mathbf{v}_i$ (similar to [8] and [1]):

$$\mathbf{v}_i = \begin{bmatrix} L_x^2 + L_y^2 \\ L_{xx}L_x^2 + 2L_{xy}L_xL_y + L_{yy}L_y^2 \\ L_{xx} + L_{yy} \\ L_{xx}^2 + 2L_{xy}^2 + L_{yy}^2 \end{bmatrix} \tag{8}$$

To determine the entropy of a set of descriptors, the probabilities of the descriptors have to be determined. We implemented the method proposed in [8]. Due to the space limitation we refer the reader to the original work for more details.

The color based descriptor $\mathbf{v}_c$ as used in [10] is given by:

$$\mathbf{v}_c = \left[R, G, B, R_x, G_x, B_x, R_y, G_y, B_y\right]^T \tag{9}$$

This descriptor uses only derivatives up to the first order. Montesinos et al. [14] argue that due to the additional color information the color 1-jet is sufficient for local structure description. Note that this descriptor is not invariant to rotation. To keep the probabilities of the descriptors computable, the probabilities of the $zero^{th}$ order signal and the first order derivatives are assumed independent, as is done in [10]. The probability of descriptor $\mathbf{v}_c$ becomes:

$$p(\mathbf{v}_c) = p((R, G, B)^T) p((R_x, G_x, B_x)^T) p((R_y, G_y, B_y)^T) \tag{10}$$

The information content of such a descriptor can be computed by summing the information content of the three independent components. This is shown in Eq. 11, where $I(L)$, $I(L_x)$, and $I(L_y)$ represent the information content of the $zero^{th}$ and first order color derivatives.

$$I(\mathbf{v}_c) = I(L) + I(L_x) + I(L_y) \tag{11}$$

Since the Harris and SUSAN detector are based on intensity and the other detectors are color based we use two information content measures. The intensity based descriptors are computed as described in [8] (cf. Eq. 8). The color based descriptors are computed according to [10] (cf. Eq. 9).

**Evaluation** A large number of points has to be considered in order to get a statistically significant measure. For this purpose a random selection of 300 images from the Corel dataset was made. The images both depict man made objects as well as images of natural scenes.

After normalization the descriptor space is partitioned in order to determine the probabilities of the descriptors. Because of normalization the same partition size can be used in all dimensions. The size of the partitions is determined by the largest absolute descriptor value of the normalized descriptors. In the experiments, each dimension of the normalized descriptor space is divided into 20 partitions.

For intensity based entropy calculation, the results are shown in Table 1. Note that in this experiments, for each corner detected by one of the methods (intensity or color based detectors), we used the descriptors calculated from derivatives of the intensity function up to the second order (cf. Eq. 8).

| Detector | Entropy | | | | |
|---|---|---|---|---|---|
| | intensity | $L$ | $L_x$ | $L_y$ | Total (color) |
| SUSAN | 3.1146 | 4.9321 | 3.3366 | 2.9646 | 11.2333 |
| Harris | 3.3866 | 4.9799 | 3.2160 | 3.2132 | 11.4091 |
| colOppHarris | 2.5541 | 5.4367 | 4.0040 | 3.9659 | 13.4066 |
| colRatHarris | 2.4505 | 5.4153 | 4.2644 | 4.2865 | 13.9662 |
| Random | 2.3654 | 4.8754 | 2.1534 | 2.2183 | 9.2470 |

**Table 1.** The intensity and color information content for different detectors

Although the color Harris detectors are included here, the intensity based descriptors are too restrictive to draw conclusions since no color information is considered in characterizing the regions. A region that is of equal intensity might be a high salient red-green corner. The color Harris detectors are included here since the occurrence of such corners is quite rare in natural images. Most color corners are also visible in the intensity image, be it with a lower gradient.

As expected, the regions that are detected using the random region generator have the lowest average information content. Also, the intensity based entropies of the regions detected by the color Harris are low. The intensity based descriptor is unable to describe the features that are detected by the color Harris detectors. The entropy of the regions detected by the Harris and SUSAN detectors are the highest.

Although the Harris and SUSAN detector are intensity based we can still use a color descriptor to compare the detected features. These detectors do not only detect pure black/white/gray corners but to a certain extent also color corners. The results for the color based entropy calculation are summarized in Table 1. The columns $L$, $L_x$ and $L_y$ correspond to the components of Eq. 11. The total entropy is computed from this by summing the components.

Again, the regions detected by the random region generator have the lowest entropy. The Harris and SUSAN based detectors perform also approximately the same. The regions detected by the color Harris detectors are by far the most distinctive according to the color based entropy calculation.

**Discussion** The color Harris detectors in combination with the intensity based descriptors are not good choices, as expected. The intensity descriptor is unable to represent the additional color information that is present; an opposite effect can be seen in the results of the color based entropy calculation.

It is clear that a good balance between repeatability and distinctiveness has to be made. By increasing the repeatability, the distinctiveness of the regions decreases. To make the regions more distinctive, color information can be used to describe the regions. When introducing the color information in the descriptor, the detector becomes less invariant to changes in illumination. By using a color model in the detector that is invariant to common changes in illumination, a tradeoff between invariance and distinctiveness can be made. This is exactly what the colRatHarris detector does. The experiments clearly show the advantages of this approach. If color information is used to describe the regions, the color Harris regions are significantly more distinctive than the regions detected by the intensity based detectors.

### 3.3 Complexity

The complexity of the complete system depends on two parameters: color or intensity framework; and Harris or SUSAN corner detector. The complexity of the color based framework is due to the additional color channels, about 3 times larger than that of the intensity based framework.

Computing the Harris cornerness measure is equally expensive in the color or intensity based framework. The SUSAN corner detector is only used in the intensity based framework. In the intensity based framework, the SUSAN corner detector operates the fastest. The Harris corner detector needs to perform more and larger convolutions to determine the cornerness measure of one pixel. If recursive filters are used to perform the convolutions, the size of the kernel does not matter anymore. In this case, the difference in speed between the detectors within the framework becomes very small.

The greater part of the total running time is spent in the framework. Also, the SUSAN and Harris corner detectors (intensity) perform similar in terms of speed. For these two reasons the choice of the corner detector should be based on performance in terms of repeatability and entropy.

The choice between using color or intensity information depends on more criteria. The complexity is increased by a factor of 3 compared to the intensity based framework. At this cost the distinctiveness of the regions is increased whereas the repeatability is decreased. Possibly, the matching complexity should also be considered, since this involves the number of regions needed for matching. The experiments have shown that when using the color information, less regions are needed to obtain a similar repeatability.

## 4 Conclusion

Based on our extensive experiments a number of conclusions can be drawn. The invariance to common image transformations is in general similar for the intensity and color Harris detectors. The intensity based detectors have the lowest computational cost. The color based detection algorithms detect the most distinctive features. Furthermore, the experiments suggest that to obtain optimal performance, a tradeoff can be made between invariance and distinctiveness by an appropriate weighting of the intensity and color information. To conclude, color information can make a significant contribution to (affine invariant) feature detection and matching.

## References

1. Mikolajczyk, K., Schmid, C.: Scale & affine invariant interest point detectors. IJCV **60** (2004) 63–86
2. Moravec, H.: Visual mapping by a robot rover. In: Int. Joint Conf. on Artif. Intell. (1979) 598–600
3. Harris, C., Stephens, M.: A combined corner and edge detector. In: 4th Alvey Vision Conf. (1988) 147–151
4. Schmid, C., Mohr, R.: Local grayvalue invariants for image retrieval. PAMI **19** (1997) 530–535
5. Smith, S., Brady, J.: SUSAN - a new approach to low level image processing. IJCV **23** (1997) 45–78
6. Lindeberg, T.: Feature detection with automatic scale selection. IJCV **30** (1998) 79–116
7. Lowe, D.: Distinctive image features from scale-invariant keypoints. IJCV **60** (2004) 91–110
8. Schmid, C., Mohr, R., Bauckhage, C.: Evaluation of interest point detectors. IJCV **37** (2000) 151–172
9. Dufournaud, Y., Schmid, C., Horaud, R.: Matching images with different resolutions. In: CVPR. (2000) 612–618
10. van de Weijer, J., Gevers, T., Geusebroek, J.M.: Edge and corner detectors by photometric quasi-invariants. PAMI **27** (2005) 625–630
11. Mikolajczyk, K., Schmid, C.: An affine invariant interest point detector. In: ECCV. Volume 1. (2002) 128–142
12. Gevers, T., Smeulders, A.: Color based object recognition. Patt. Recogn. **32** (1999) 453–464
13. Koenderink, J., van Doorn, A.: Representation of local geometry in the visual system. Biol. Cybern. **55** (1987) 367–375
14. Montesinos, P., Gouet, V., Deriche, R.: Differential invariants for color images. In: ICPR. (1998) 838–840