Boosting Color Saliency in Image Feature Detection

J. van de Weijer[1] , Th. Gevers[2] , A.D. Bagdanov[3]

index terms: I.4.8.a Color, I.4.6.a Edge and feature detection, I.4.10.d Statistical

Abstract

The aim of salient feature detection is to find distinctive local events in images. Salient features are generally determined from the local differential structure of images. They focus on the shape-saliency of the local neighborhood. The majority of these detectors is luminance based which has the disadvantage that the distinctiveness of the local color information is completely ignored in determining salient image features. To fully exploit the possibilities of salient point detection in color images, color distinctiveness should be taken into account in addition to shape distinctiveness. In this paper, color distinctiveness is explicitly incorporated into the design of saliency detection. The algorithm, called color saliency boosting, is based on an analysis of the statistics of color image derivatives. Color saliency boosting is designed as a generic method easily adaptable to existing feature detectors. Results show that substantial improvements in information content are acquired by targeting color salient features.

I. INTRODUCTION

Indexing objects and object categories as an ordered collection of salient image points has been successfully applied to image matching, content-based retrieval, learning and recognition [1], [2], [3], [4], [5], [6]. Salient points are local features in the image which exhibit geometrical structure, such as T-junctions, corners, and symmetry points. The aim of salient point detection is to represent objects more concisely and being robust to varying viewing conditions, such as changes due to camera zooming, object rotation, and illumination changes.

Although the majority of image data is in color format nowadays, most salient point detectors are luminance based. They typically focus on shape saliency rather than color saliency [7], [8].

1. Lear Group, GRAVIR-INRIA, 655 avenue de l'Europe, 38330 Montbonnot, France,. E-mail: van-de-weijer@inrialpes.fr

2. ISIS group, University of Amsterdam, Kruislaan 403, 1098 SJ Amsterdam, The Netherlands. E-mail: gevers@science.uva.nl

3. Dipartimento di Sistemi e Informatica, Università degli Studi di Firenze, Italy. E-mail: bagdanov@gmail.com

For example, they focus on corner points without distinguishing (low-salient) black-and-white corners from (high-salient) red-green corners. Only recently color information has been incorporated in the detection phase. Montesinos et al. [9] propose an extension of the luminance Harris corner detector to color [10]. Heidemann [11] incorporates color into the generalized symmetry transform proposed by Reisfeld et al. [12]. Both methods achieve a performance gain for near isoluminant events. However, since the luminance axis remains the major axes of variation in the RGB-cube, results do not differ greatly from luminance based feature detection. Itti et al. [13] use color contrast as a clue for saliency. Their method is based on a zero-order signal (normalized red, green, blue, yellow), and is not easily extendable to differential-based features.

For the evaluation of salient point detectors, Schmid et al. [14] propose two criteria: 1. *repeatability*, salient point detection should be stable under varying viewing conditions. 2. *distinctiveness*, salient points should focus on events with a low probability of occurrence. Most salient point detectors are designed according to these criteria. They focus on two dimensional structures, such as corners, which are stable and distinctive at the same time. Although color is considered to play an important role in attributing image saliency [15], the explicit incorporation of color distinctiveness into the design of salient points detectors has been ignored.

Therefore, in this paper, we aim to incorporate color distinctiveness into salient point detection. The extension should be general and hence be easy to incorporate in existing salient point detectors. For a color image, with values $\mathbf{f} = (R, G, B)^T$, salient points are the maxima of the saliency map, which compares the derivative vectors in a neighborhood fixed by scale σ,

$$s = H^{\sigma}\left(\mathbf{f}_x, \mathbf{f}_y\right) \qquad (1)$$

where H is the saliency function and the subscript indicates differentiation with respect to the parameter. This type of saliency maps include [7], [10], [11], [16], [17]. The impact of a derivative vector on the outcome of the local saliency depends on its vector norm, $\|\mathbf{f}_x\|$. Hence, vectors with equal norm have an equal impact on the local saliency. Rather than deriving saliency from the vector norm, the novelty of this paper is to adapt the saliency function in order that vectors with equal color distinctiveness have equal impact on the saliency function.

II. COLOR DISTINCTIVENESS

The efficiency of salient point detection depends on the distinctiveness of the extracted salient points. At the salient points' positions, local neighborhoods are extracted and described by

local image descriptors. The distinctiveness of the descriptor defines the conciseness of the representation and the discriminative power of the salient points. The distinctiveness of interest points is measured by its information content [14].

For luminance-based descriptors, the information content is measured by looking at the distinctiveness of the differential structure described by the local 2-jet [18] at the detected points [4]. Montesinos et al. [9] argue that, due to the extra information available in color images, the color 1-jet is sufficient for the local structure description. The color 1-jet descriptor is given by

$$\mathbf{v} = \left(\begin{array}{ccccccccc} R & G & B & R_x & G_x & B_x & R_y & G_y & B_y \end{array} \right)^T. \tag{2}$$

From information theory, it is known that the information content of an event is dependent on its frequency or probability

$$I(\mathbf{v}) = -\log\left(p\left(\mathbf{v}\right)\right), \tag{3}$$

where $p(\mathbf{v})$ is the probability of the descriptor $\mathbf{v}$. i.e. events which occur rarely are more informative. The information content of the descriptor, given by Eq. 2, is approximated by assuming independent probabilities of the zeroth order signal and the first order derivatives

$$p(\mathbf{v}) = p(\mathbf{f}) p(\mathbf{f}_x) p(\mathbf{f}_y). \tag{4}$$

Hence, the information content of the salient point detector, defined by Eq. 1, will increase if the probability of the derivatives, $p(\mathbf{f}_x)$, is small.

By adapting the saliency map to focuss on rare color derivatives, the color distinctiveness of the detector is improved. Traditionally, for saliency maps based on Eq. 1, derivatives with an equal vector norm $\|\mathbf{f}_x\|$ have equal influence on the saliency map. We now adapt this by requiring vectors with equal information content to have equal influence on the saliency map. Hence, the aim is to find a transformation $g : \Re^3 \to \Re^3$ for which holds that

$$p(\mathbf{f}_x) = p\left(\mathbf{f}'_x\right) \leftrightarrow \|g(\mathbf{f}_x)\| = \left\|g\left(\mathbf{f}'_x\right)\right\|. \tag{5}$$

The transformation, attained by function g, is called *color saliency boosting*. Once a function g has been found, the color boosted saliency can be computed by

$$s = H^\sigma\left(g(\mathbf{f}_x), g(\mathbf{f}_y)\right). \tag{6}$$

The traditional saliency map, which derives saliency from the gradient strength of the derivatives, is after color boosting based the information content of these derivatives. Gradient strength has been replaced by information content, thereby aiming for higher saliency.

coordinate transformation	transformed derivative	decorrelated variation
$\begin{pmatrix} \theta \\ \varphi \\ r \end{pmatrix} = \begin{pmatrix} \arctan(\frac{G}{R}) \\ \arcsin\left(\frac{\sqrt{R^2+G^2}}{\sqrt{R^2+G^2+B^2}}\right) \\ r = \sqrt{R^2+G^2+B^2} \end{pmatrix}$	$S\left(\mathbf{f}_x\right) = \mathbf{f}_x^s = \begin{pmatrix} r\sin\varphi\,\theta_x \\ r\varphi_x \\ r_x \end{pmatrix}$	shadow-shading
$\begin{pmatrix} o1 \\ o2 \\ o3 \end{pmatrix} = \begin{pmatrix} \frac{R-G}{\sqrt{2}} \\ \frac{R+G-2B}{\sqrt{6}} \\ \frac{R+G+B}{\sqrt{3}} \end{pmatrix}$	$O\left(\mathbf{f}_x\right) = \mathbf{f}_x^o = \begin{pmatrix} o1_x \\ o2_x \\ o3_x \end{pmatrix}$	specular
$\begin{pmatrix} h \\ s \\ i \end{pmatrix} = \begin{pmatrix} \arctan\left(\frac{o1}{o2}\right) \\ \sqrt{o1^2+o2^2} \\ o3 \end{pmatrix}$	$H\left(\mathbf{f}_x\right) = \mathbf{f}_x^h = \begin{pmatrix} s\,h_x \\ s_x \\ i_x \end{pmatrix}$	shadow-shading-specular

TABLE I

THE COLOR COORDINATE TRANSFORMATIONS, THEIR COLOR DERIVATIVES AND THE PHYSICAL EVENT RELATED TO THE

TRANSFORMATION.

III. STATISTICS OF COLOR IMAGES

As discussed in Section II, the information content of a feature descriptor depends on the probability of the derivatives. In this section, we investigate the statistics of color derivatives to find a mathematical description of surfaces of equal probability, so called isosalient surfaces. A description of these surfaces leads to the solution of Eq. 5.

The channels of $\mathbf{f}_x$, $\{R_x, G_x, B_x\}$, are correlated due to the physics of the world. Photometric events in the real-world, such as shading, shadows, and specularities influence RGB values in a well defined manner [19]. Before investigating the statistics of color derivatives, the derivatives need to be transformed to a color space which is uncorrelated with respect to these photometric events. For this purpose, we apply the color transformation as proposed in [20]. An overview is given in Table I. These coordinate transformations contain axes which are photometric variant with respect to a physical cause (see column three of Table I), and photometric invariant axes which are invariant with respect to this cause. For more information on the derivation and the assumptions from which these color spaces are derived we refer to [19], [20].

The statistics of color images are shown for the Corel database, which consists of 40,000 images of 256x384 pixels (for a more extensive elaboration on the Corel set see e.g. [21]). In Fig. 1 the distributions of the first order derivatives, $\mathbf{f}_x$, are given for the various color coordinate
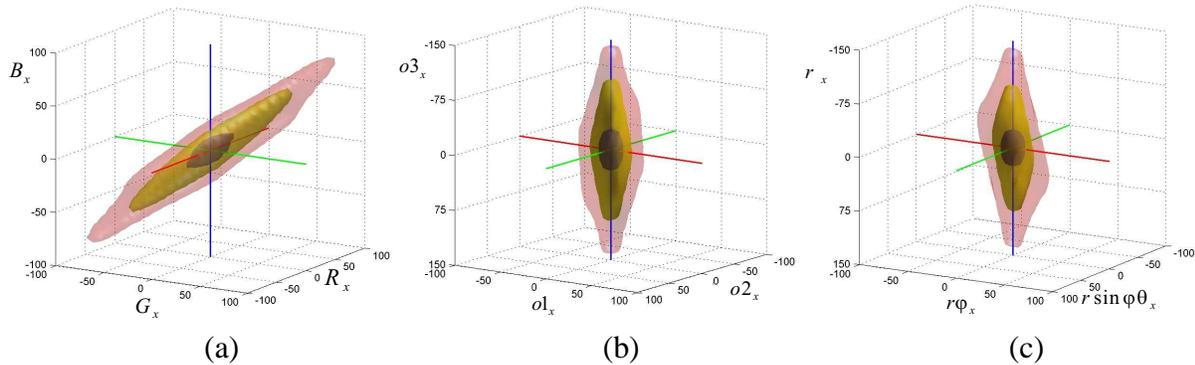
Fig. 1. The histograms of the distribution of the transformed derivatives of the Corel image database in respectively the (a) RGB coordinates, (b) the opponent coordinates and (c) the spherical coordinates. The three planes correspond with the isosalient surfaces which contain (from dark to light) respectively 90%, 99%, 99.9%t of the total number of pixels.

systems. The isosalient surfaces form simple structures similar to ellipsoids. For all three color spaces, the third coordinate coincides with the axis of maximum variation (i.e. the intensity). For the opponent and the spherical coordinate system, the distribution on the plane spanned by the first and second coordinate form an ellipse of which the axes do not align with the coordinates. To accomplish a correct alignment between our coordinate axes and the axes of the ellipsoid, we rotate, with rotation matrix R, the color coordinate system to coincide with the axes of the ellipsoid:

$$
\begin{aligned}
\left(r \sin \tilde{\varphi} \; \tilde{\theta}_x, r\tilde{\varphi}_x\right)^T &= R^\phi \left(r \sin \varphi \theta_x, r\varphi_x\right)^T, \\
\left(\tilde{o}1_x, \tilde{o}2_x\right)^T &= R^\phi \left(o1_x, o2_x\right)^T.
\end{aligned}
\tag{7}
$$

The tilde is used to indicate that all axes are aligned with the axes of the ellipsoid. Consequently, the aligned transformations are given by $\tilde{S}\left(\mathbf{f}_x\right) = \mathbf{f}_x^{\tilde{s}}$ and $\tilde{O}\left(\mathbf{f}_x\right) = \mathbf{f}_x^{\tilde{o}}$.

After the alignment of the axes, isosalient surfaces of the derivative histograms can be approximated by ellipsoids

$$
\left(\alpha h_x^1\right)^2 + \left(\beta h_x^2\right)^2 + \left(\gamma h_x^3\right)^2 = R^2
\tag{8}
$$

where $\mathbf{h}_x = h\left(\mathbf{f}_x\right) = (h_x^1, h_x^2, h_x^3)^T$ and h is one of the color transformations $\tilde{S}$, $\tilde{O}$, or H.

IV. Boosting Color Saliency

In this section, the goal is to incorporate color distinctiveness into salient point detection. Or mathematically, to find the transformation for which vectors with equal information content have

equal impact on the saliency function. In the previous section, it was shown that derivatives of equal saliency form ellipsoids. Since Eq. 8 is equal to

$$\left(\alpha h_x^1\right)^2 + \left(\beta h_x^2\right)^2 + \left(\gamma h_x^3\right)^2 = \|\mathbf{\Lambda} h\left(\mathbf{f}_x\right)\|^2, \tag{9}$$

the following holds

$$p\left(\mathbf{f}_x\right) = p\left(\mathbf{f}_x'\right) \leftrightarrow \|\mathbf{\Lambda} h\left(\mathbf{f}_x\right)\| = \left\|\mathbf{\Lambda}^T h\left(\mathbf{f}_x'\right)\right\|, \tag{10}$$

where $\mathbf{\Lambda}$ is a 3x3 diagonal matrix with $\mathbf{\Lambda}_{11} = \alpha$, $\mathbf{\Lambda}_{22} = \beta$, and $\mathbf{\Lambda}_{33} = \gamma$. $\mathbf{\Lambda}$ is restricted to $\Lambda_{11}^2 + \Lambda_{22}^2 + \Lambda_{33}^2 = 1$. The desired color saliency boosting function (see Eq. 5) is obtained by

$$g\left(\mathbf{f}_x\right) = \mathbf{\Lambda} h\left(\mathbf{f}_x\right), \tag{11}$$

where h is one of the color transformations $\tilde{S}$, $\tilde{O}$, or H. By a rotation of the color axes followed by a rescaling of the axis, the oriented isosalient ellipsoids are transformed into spheres, and thus vectors of equal saliency are transformed into vectors of equal length.

A. Influence of Color Saliency Boosting on Repeatability

The two criteria for salient point detection are distinctiveness and repeatability. The color boosting algorithm is designed to focus on color distinctiveness, while adopting the geometrical characteristics of the operator to which it is applied. In this section, we examine the influence of color boosting on the repeatability. We identify two phenomena which influence the repeatability of $g\left(\mathbf{f}_x\right)$. Firstly, by boosting the color saliency, an anisotropic transformation is carried out, which will negatively reduce the signal-to-noise ratio. Secondly, by boosting the photometric invariant directions (more than the photometric variant directions), the robustness is improved with respect to scene accidental changes.

Loss of signal-to-noise ratio: for isotropic uncorrelated noise, ε, the measured derivative $\hat{\mathbf{f}}_x$ can be written as $\hat{\mathbf{f}}_x = \mathbf{f}_x + \varepsilon$ and after color saliency boosting by

$$g\left(\hat{\mathbf{f}}_x\right) = g\left(\mathbf{f}_x\right) + \mathbf{\Lambda}\varepsilon. \tag{12}$$

Note that isotropic noise remains unchanged under the orthonormal curvilinear transformations. Assume the worst case in which $\mathbf{f}_x$ only has signal along the photometric variant axis. In this case, the noise can be written as

$$\frac{\|g\left(\mathbf{f}_x\right)\|}{\|\mathbf{\Lambda}\varepsilon\|} \approx \frac{\Lambda_{33}\|\mathbf{f}_x\|}{\Lambda_{11}\|\varepsilon\|}. \tag{13}$$

	$\mathbf{f}_x$	$\|\mathbf{f}_x\|_1$	$\mathbf{f}_x^{\tilde{s}}$	$\tilde{\mathbf{S}}_x^c$	$\mathbf{f}_x^{\tilde{o}}$	$\tilde{\mathbf{O}}_x^c$	$\mathbf{f}_x^h$	$\mathbf{H}_x^c$
λ_1	0.577	1	0.851	0.856	0.850	0.851	0.858	1
λ_2	0.577	-	0.515	0.518	0.524	0.525	0.509	0
λ_3	0.577	-	0.099	0	0.065	0	0.066	0

TABLE II

THE DIAGONAL ENTRIES OF Λ FOR THE COREL DATA SET COMPUTED FOR GAUSSIAN DERIVATIVES WITH $\sigma = 1$.

Hence, the signal-to-noise ratio reduces by $\frac{\Lambda_{11}}{\Lambda_{33}}$. The loss of signal will negatively influence repeatability to geometrical and photometrical changes.

Gain in photometric robustness: by boosting color saliency the influence of the photometric variant direction diminishes while the influence of the invariant directions increases. As a consequence the repeatability under photometric changes, such as changing illumination and viewpoint, increases.

Depending on the task at hand, color distinctiveness may be less desired than signal-to-noise. For this purpose the α parameter is proposed, which allows for choosing between best signal-to-noise characteristics, $\alpha = 0$, and best information content, $\alpha = 1$:

$$g^\alpha \left(\mathbf{f}_x \right) = \alpha \mathbf{\Lambda} h \left(\mathbf{f}_x \right) + \left(1 - \alpha \right) h \left(\mathbf{f}_x \right). \tag{14}$$

V. EXPERIMENTS AND ILLUSTRATIONS

Color saliency boosting is tested on information content and repeatability. The salient points based on color saliency boosting are compared to luminance $\|\mathbf{f}_x\|_1$, RGB gradient, $\mathbf{f}_x$, and the quasi-invariant-based salient point detectors. The quasi-invariants are derived from the same color transformation as given in Table I by only using the invariant coordinates of the transformation: the shadow-shading quasi-invariant $\tilde{\mathbf{S}}_x^c = (r \sin \varphi \, \theta_x, r \varphi_x, 0)$, the specular quasi-invariant $\tilde{\mathbf{O}}_x^c = (o1_x, o2_x, 0)$, and the shadow-shading-specular quasi-invariant $\mathbf{H}_x^c = (s \, h_x, s_x, 0)$. An extensive analysis of the quasi-invariants can be found in [17], [20]. Finally, the generality of the approach is illustrated by applying color boosting to several existing feature detectors.

A. Initialization

Experiments are performed on a subset of 1000 randomly chosen images from the Corel data set. Before color saliency boosting can be applied, the Λ-parameters (Eq.9) have to be initialized

<center>(a)   (b)   (c)   (d)</center>

Fig. 2. (a) Example Soil data set and (c) frame from table-tennis sequence. (b) and (d) results of Harris detector (red dots) and the Harris detector with color boosting (yellow dots). The red dots mainly coincide with black and white events, while the yellow dots are focussed on colorful points.

by fitting ellipses to the histogram of the data set. The axes of the ellipsoid are derived by fitting the isosaliency surface which contains 99 percent of the pixels of the histogram of the Corel data set. Changing this parameter to 99.9 or 99.99 percent changes matrix Λ only slightly. The results for the various transformations are summarized in Table II. The relation between the axes in the various color spaces clearly confirms the dominance of the luminance axis in the RGB-cube, since Λ_{33}, the multiplication-factor of the luminance axis, is much smaller than the color-axes multiplication factors, Λ_{11} and Λ_{22}.

To give an idea on how the Λ-parameters vary when changing the data set, we have estimated the Λ parameters for two other data sets, the Soil data set [22] which is an uncompressed set of object images and a table-tennis sequence, given in Fig. 2a,c. For the Soil data set and the opponent color model, the Λ-parameters are $\Lambda_{11} = 0.542$, $\Lambda_{22} = 0.780$, and $\Lambda_{33} = 0.313$. Since this set consists of colorful objects the luminance axis is less suppressed than for the Corel set. For the tennis sequence the difference with the Corel dataset is smaller, $\Lambda_{11} = 0.588$, $\Lambda_{22} = 0.799$, and $\Lambda_{33} = 0.124$. A change in Λ-parameters can have various causes such as the quality of the camera, the applied compression and the different color content of the image data.

To test the influence of compression on the shape of the ellipsoids, we have repeated the ellipse fitting procedure for the same Corel images but after JPEG compression with a quality of 30%. For the opponent color model, we obtained: $\Lambda_{11} = 0.822$, $\Lambda_{22} = 0.567$, and $\Lambda_{33} = 0.062$, which only slightly differ from the parameters found in Table II. Hence, JPEG compression has been found to have little influence on the shape of the fitting procedure of the ellipses.

We have chosen the color Harris point detector [9], [10] to test color boosting in following

experiments. It is computed with

$$H^{\sigma}\left(\mathbf{f}_x, \mathbf{f}_y\right) = \overline{\mathbf{f}_x \cdot \mathbf{f}_x \ \mathbf{f}_y \cdot \mathbf{f}_y} - \overline{\mathbf{f}_x \cdot \mathbf{f}_y}^2 - k\left(\overline{\mathbf{f}_x \cdot \mathbf{f}_x} + \overline{\mathbf{f}_y \cdot \mathbf{f}_y}\right)^2 \qquad (15)$$

by substituting $\mathbf{f}_x$ and $\mathbf{f}_y$ by $g\left(\mathbf{f}_x\right)$ and $g\left(\mathbf{f}_y\right)$ and with $k = 0.04$. The bar $\bar{\ }$ indicates convolution with a Gaussian filter and the dot indicates the inner product. We applied Gaussian derivatives of $\sigma = 1$ and Gaussian smoothing with $\sigma = 3$.

B. Color Distinctiveness

Here, the extend in which color boosting improves the color distinctiveness of the Harris detector is examined. In [14], the Harris detector has been shown to outperform other detectors both on 'shape' distinctiveness and repeatability. The color distinctiveness of salient point detectors is described by the information content of the descriptors extracted at the locations of the salient points. From the combination of Eq. 3 and Eq. 4, it follows that the total information is computed by summing up the information of the zeroth and first order part, $I\left(\mathbf{v}\right) = I\left(\mathbf{f}\right) + I\left(\mathbf{f}_x\right) + I\left(\mathbf{f}_y\right)$. The information content of the parts is computed from the normalized histograms by

$$I\left(\mathbf{f}\right) = -\sum_i p_i \log\left(p_i\right) \qquad (16)$$

where p_i are the probabilities of the bins of the histogram of $\mathbf{f}$.

The results for 20 and 100 salient points per image are shown in Table III. Next to the absolute information content, we have also computed the relative information gain with respect to the information content of the color gradient based Harris detector. For this purpose, the information content of a single image is defined as

$$I = -\sum_{j=1}^{n} \log\left(p\left(v_j\right)\right), \qquad (17)$$

where $j = 1, 2, ...n$ and n is the number of salient points in the image. Here $p\left(v_j\right)$ is computed from the global histograms, which allows comparison of the results per image. The information content change is considered substantially for a 5 percent increase or decrease.

The highest information content is obtained with $\mathbf{f}_x^{\tilde{o}}$, which is the color saliency boosted version of the opponent derivatives. The boosting results in an increase of 7% to 13% of the information content compared to the color gradient based detector. On the images of the Corel set this resulted in a substantial increase for 22% to 63% of the images. The advantage of color

	standard descriptor						normalized descriptor					
	20 points			100 points			20 points			100 points		
method	inf.	incr(%)	decr(%)	inf.	incr.	decr.	inf.	incr.	decr.	inf.	incr.	decr.
$\mathbf{f}_x$	20.4	-	-	20.0	-	-	13.2	-	-	13.9	-	-
$\|\mathbf{f}_x\|_1$	19.9	0	1.4	19.8	0	0.8	13.0	0	2.7	13.8	0	1.0
$\tilde{\mathbf{S}}_x^c$	22.2	45.5	10.1	20.4	9.1	17.7	17.9	92.9	0.9	16.2	69.8	2.8
$\mathbf{f}_x^{\tilde{s}}$	22.3	49.4	.6	20.8	13.1	1.3	16.9	86.9	0.6	15.5	57.6	.7
$\tilde{\mathbf{O}}_x^c$	22.6	51.4	12.9	20.5	12.0	34.2	18.9	92.5	1.3	16.5	64.6	10.8
$\mathbf{f}_x^{\tilde{o}}$	23.2	62.6	0.0	21.4	21.5	0.9	18.4	88.2	0.3	16.4	65.0	1.7
$\mathbf{H}_x^c$	21.0	21.7	43.4	19.0	1.8	77.4	17.3	77.1	10.9	14.8	31.7	37.9
$\mathbf{f}_x^h$	23.0	57.2	0.3	21.3	16.7	1.1	18.3	87.4	0.5	16.2	62.3	2.2
rand.	14.4	0	99.8	14.4	0	100	10.1	2.7	89.1	10.2	.6	96.7

TABLE III

THE INFORMATION CONTENT OF SALIENT POINT DETECTORS. MEASURED IN 1. INFORMATION CONTENT AND 2. THE PERCENTAGE OF IMAGES FOR WHICH A SUBSTANTIAL DECREASE (-5%) OR INCREASE ($+5\%$) OF THE INFORMATION CONTENT OCCURS. THE EXPERIMENT IS PERFORMED WITH BOTH 20 OR 100 SALIENT POINTS PER IMAGE. THE EXPERIMENT IS REPEATED WITH A NORMALIZED DESCRIPTOR WHICH IS INVARIANT FOR LUMINANCE CHANGES.

boosting diminishes when increasing the number of salient points per image. This is caused by the limited number of color clues in many of the images, which is especially visible for the results of the photometric quasi-invariants, $\tilde{\mathbf{S}}_x^c$, $\tilde{\mathbf{O}}_x^c$, or $\mathbf{H}_x^c$. Note that these detectors discard all intensity information, which in the case of 100 salient points per image results in many images with a substantial decrease in information content. Finally, it is noteworthy to observe how small the difference is between luminance and RGB-based Harris detection. Since the intensity direction also dominates the RGB derivatives, using RGB-gradient instead of luminance-based Harris detection only results in a substantial increase in information content in 1% of the images.

It is often desirable for the descriptor to be invariant for scene incidental events like shading and shadows. In these cases the information content of the normalized descriptor, which is invariant to luminance changes, better reflects the information content of the salient point detector

$$v = \left(\frac{R}{\|\mathbf{f}\|}, \frac{G}{\|\mathbf{f}\|}, \frac{B}{\|\mathbf{f}\|}, \frac{R_x}{\|\mathbf{f}_x\|}, \frac{G_x}{\|\mathbf{f}_x\|}, \frac{B_x}{\|\mathbf{f}_x\|}, \frac{R_y}{\|\mathbf{f}_y\|}, \frac{G_y}{\|\mathbf{f}_y\|}, \frac{B_y}{\|\mathbf{f}_y\|} \right). \tag{18}$$

The results of the normalized descriptor are given in the right half of Table III. The increase in information content of the quasi-invariants and the color boosted detectors stands out even
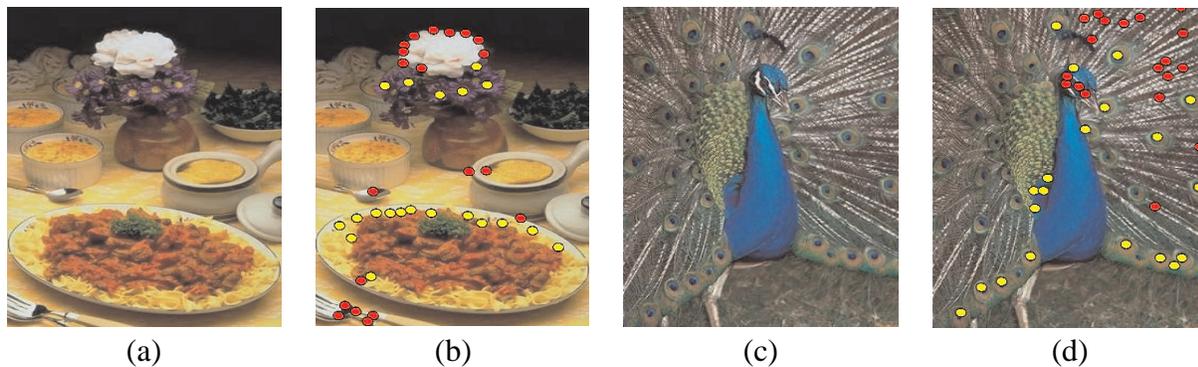
Fig. 3. (a) and (c) Corel images. (b) and (d) results of Harris detector (red dots) and the Harris detector with color boosting (yellow dots). The red dots mainly coincide with black and white events, while the yellow dots are focussed on colorful points.

more, with substantial gains in information content up to 90%. Here the quasi-invariants based detectors outperform the other detectors.
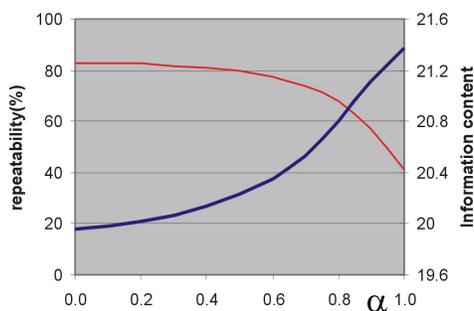
In Fig. 3 results of the RGB-gradient based and color boosted Harris detector are depicted. From a color information point of view, the performance of the RGB-gradient based method is poor. Most of the salient points have a black and white local neighborhood. The salient points after color boosting, focus on more distinctive points. Similar results are depicted in Fig. 2b,d, where the results are shown computed with the Λ-parameters belonging to their data sets.

C. Repeatability: signal-to-noise

Repeatability measures the stability with respect to varying viewing conditions. As indicated in section IV-A, color saliency boosting reduces the signal-to-noise ratio. Repeatability with respect to geometrical changes, scaling, and affine transformations are considered an inherent

method	20 points	100 points
$\mathbf{f}_x$	88	84
$|\mathbf{f}_x|_1$	88	83
$\tilde{\mathbf{S}}_x^c$	53	42
$\mathbf{f}_x^{\tilde{s}}$	62	54
$\tilde{\mathbf{O}}_x^c$	46	34
$\mathbf{f}_x^{\tilde{o}}$	51	41
$\mathbf{H}_x^c$	35	25
$\mathbf{f}_x^h$	52	42

(a)

(b)

Fig. 4. (a) The percentage of Harris points which remain detected after adding Gaussian uncorrelated noise. (b) The information content (blue line) and the repeatability (red line) as a function of the amount of color saliency boosting.

Fig. 5. (a), (b) Two frames from two sequences with changing illumination conditions. (c) Repeatability as a function of the amount of color saliency boosting. Dotted line for the sequence (a) and the continuous line for sequence (b).

property of the detector and will not be considered here. The loss of repeatability caused by color saliency boosting is examined by adding uniform, uncorrelated Gaussian noise of $\sigma = 10$. This yields a good indication of loss in signal-to-noise, which in its turn will influence the results of repeatability under other variations, such as zooming, illumination changes, and geometrical changes. Repeatability is measured by comparing the Harris points detected in the noisy image to the points in the noise-free images. The results in Fig. 4a correspond to the expectation made by Eq. 13. The larger the difference between Λ_{11} and Λ_{33}, the poorer the repeatability. In Fig. 4b, the information content and repeatability as a function of the amount of color boosting, determined by the α-parameter, is given for the opponent color space (see Eq.14). The results show that information content increases at the cost of stability.

D. Repeatability: photometric variation

Photometric robustness increases with color boosting, as discussed in Section IV-A. In Fig. 5 the dependance of repeatability is tested on two sequences with changing illumination conditions [23]. The experiment was performed by applying boosting to the spherical color space, $\mathbf{f}_x^{\tilde{s}}$, since changes due to shadow-shading will be along the photometric variant direction of the spherical system. For these experiments two intertwining phenomena can be observed: the improved photometric invariance and the deterioration of signal-to-noise ratio with increasing α. For the nuts-sequence, with very prominent shadows and shading, the photometric invariance is dominant, while for the fruit-basket the gained photometric invariance only improves performance slightly for medium α values. For total color saliency boosting, $\alpha = 1$ the loss of repeatability,
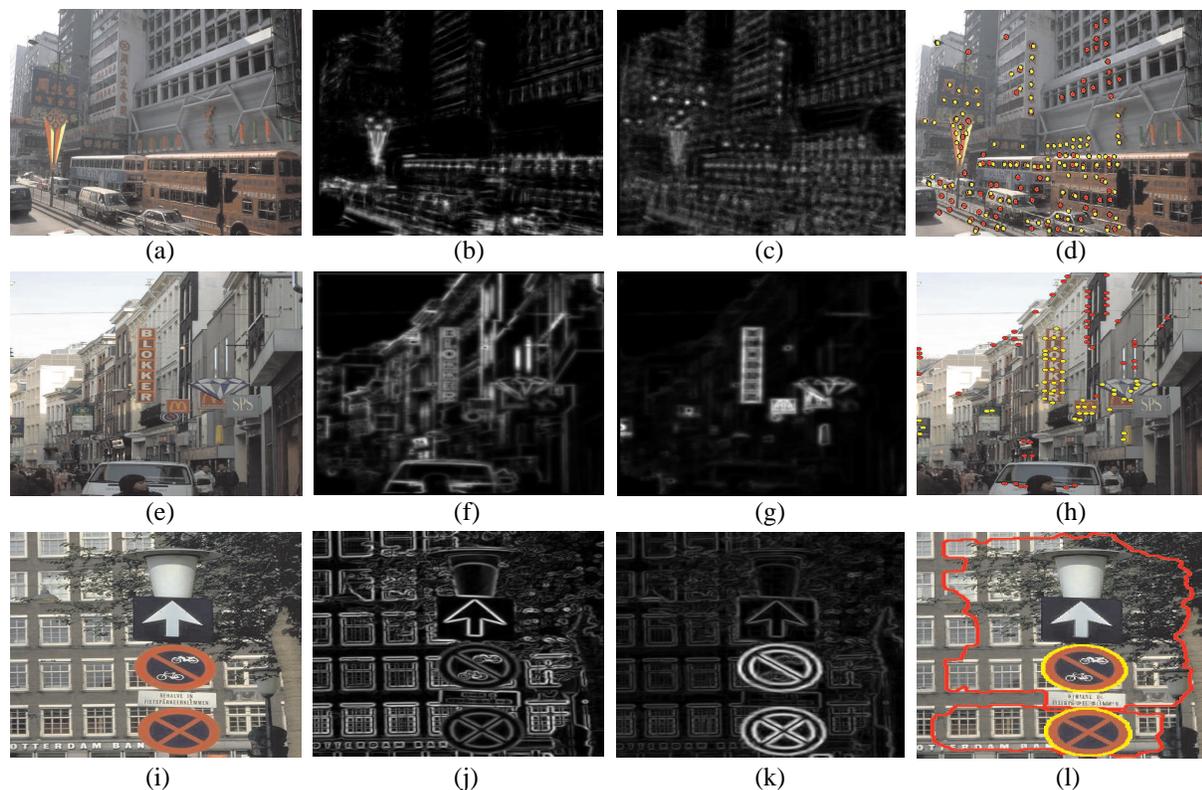
Fig. 6.   Horizontally, respectively, the input image, RGB-gradient based saliency map, the color boosted saliency map and the results with red dots (lines) for the gradient-based method and yellow dots (lines) for the salient points after color saliency boosting. Row one (a,b,c,d): results after [11], row two (e,f,g,h): results after [16], and row 3 (i,j,k,l): results after [24]

due to loss of signal-to-noise, is substantial.

E. Generality: Illustrations

Color saliency boosting can be applied on functions which can be written as a function of the local derivatives. Here we apply it to three different feature detectors. First, the focus point detector which was originally proposed by Reisfeld et al. [12] and recently extended to color by Heidemann [11]. The detector focuses on the center of locally symmetric structures. On the first row of Fig. 6, the result of the focus point detector are shown. Fig. 6b shows the saliency map as proposed in [11]. In Fig. 6c the saliency map after saliency boosting is depicted. Although focus point detection is already an extension from luminance to color, black-and-white transition still dominate the result. Only after boosting the color saliency, the less interesting black-and-white structures in the image are ignored and most of the red Chinese signs are found, see Fig. 6d. Similar difference in performance is obtained by applying color boosting to the star detector

proposed by Bigün [16]. This detector focuses on corner and junction like structures. The RGB gradient based method (Fig. 6f) focuses mainly on black-and-white events while the more salient signboards (Fig. 6g) are found only after color saliency boosting.

As a final illustration, we illustrate that color saliency boosting can be applied to gradient based methods. In the third row of Fig. 6, color boosting is applied to a gradient based segmentation algorithm proposed by Jermyn and Ishikawa [24]. The algorithm finds globally optimal regions and boundaries. In Fig. 6b and c respectively the RGB gradient and the color boosted gradient are depicted. While the RGB-gradient based segmentation is distracted by the many black-and-white events in the background, the color boosted segmentation finds the salient traffic signs.

VI. Conclusions

In this paper, color distinctiveness is explicitly integrated in the design of salient point detectors. The method, called color saliency boosting, can be incorporated into existing detectors which are mostly focused on shape distinctiveness. Saliency boosting is based on the analysis of the statistics of color image derivatives. Isosalient derivatives form ellipsoids in the color derivative distributions., which is exploited to adapt derivatives in such a way that equal saliency implies equal impact on the saliency map. Experiments show that color saliency boosting substantially increases the information content of the detected points. A substantial information content increase is obtained on up to $20 - 60\%$ of the Corel images.

VII. Acknowledgements

References

[1] R. Fergus, P. Perona, and A. Zisserman, "Object class recognition by unsupervised scale-invariant learning," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, vol. 2, June 2003, pp. 264–271.

[2] D. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[3] K. Mikolajczyk and C. Schmid, "Scale and affine invariant interest point detectors," *Int. J. Computer Vision*, vol. 60, no. 1, pp. 62–86, 2004.

[4] C. Schmid and R. Mohr, "Local grayvalue invariants for image retrieval," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 19, no. 5, pp. 530–534, May 1997.

[5] N. Sebe, Q. Tian, E. Loupias, M. Lew, and T. Huang, "Evaluation of salient point techniques," *Image and Vision Computing*, vol. 21, no. 13-14, pp. 1087–1095, 2003.

[6] L. van Gool, T. Tuytelaars, and A. Turina, "Local features for image retrieval," in *State-of-the-Art in Content-Based Image and Video Retrieval*. Kluwer Academic Publishers, 2001, pp. 21–41.

[7] M. Lee and G. Medioni, "Grouping into regions, curves, and junctions," *Computer Vision Image Understanding*, vol. 76, no. 1, pp. 54–69, 1999.

[8] L. Williams and K. Thornber, "A comparision of measures for detecting natural shapes in cluttered backgrounds," *International Journal of Computer Vision*, vol. 34, no. 2/3, pp. 81–96, 1999.

[9] P. Montesinos, V. Gouet, and R. Deriche, "Differential Invariants for Color Images," in *Proceedings of 14th International Conference on Pattern Recognition*, Brisbane, Australia, 1998, pp. 838–840.

[10] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proc. 4th Alvey Vision Conf.*, vol. 15, 1988, pp. 147–151.

[11] G. Heidemann, "Focus-of-attention from local color symmetries," *IEEE Trans. on pattern analysis and machine intelligence*, vol. 26, no. 7, pp. 817–847, 2004.

[12] D. Reisfeld, H. Wolfson, and Y. Yeshurun, "Context free attentional operators: the generalized symmetry," *International Journal of Computer Vision*, vol. 14, pp. 119–130, 1995.

[13] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1254–1259, Nov 1998.

[14] C. Schmid, R. Mohr, and C. Bauckhage, "Evaluation of interest point detectors," *International Journal of Computer Vision*, vol. 37, no. 2, pp. 151–172, 2000.

[15] L. Itti, C. Koch, and E. Niebur, "Computation modeling of visual attention," *Nature Reviews Neuroscience*, vol. 2, no. 11, pp. 194–203, March 2001.

[16] J. Bigun, "Pattern recognition in images by symmetry and coordinate transformations," *Computer Vision and Image Understanding*, vol. 68, no. 3, pp. 290–307, 1997.

[17] J. van de Weijer, T. Gevers, and A. Smeulders, "Robust photometric invariant features from the color tensor," *IEEE Trans. Image Processing*, accepted 2004.

[18] J. J. Koenderink and A. J. van Doom, "Representation of local geometry in the visual system," *Biol. Cybern.*, vol. 55, no. 6, pp. 367–375, 1987.

[19] S. Shafer, "Using color to seperate reflection components," *COLOR research and application*, vol. 10, no. 4, pp. 210–218, Winter 1985.

[20] J. van de Weijer, T. Gevers, and J. Geusebroek, "Edge and corner detection by photometric quasi-invariants," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 27, no. 4, pp. 625–630, 2005.

[21] O. Chapelle, P. Haffner, and V. Vapnik, "Support vector machines for histogram-based image classification," *IEEE Trans. on Neural Networks*, vol. 10, no. 5, pp. 1055–1064, 1999.

[22] D. Koubaroulis, J. Matas, and J. Kittler, "Evaluating colour-based object recognition algorithms using the soil-47 database," in *Asian Conference on Computer Vision*, 2002.

[23] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," in *International Conference on Computer Vision and Pattern Recognition*, Madison, USA, 2003.

[24] I. Jermyn and H. Ishikawa, "Globally optimal regions and boundaries as minimum ratio weight cycles," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23, no. 10, pp. 1075–1088, Oct 2001.