Extracting Scale and Illuminant Invariant Regions through Color

Ranjith Unnikrishnan Martial Hebert

Robotics Institute, Carnegie Mellon University Pittsburgh, Pennsylvania {ranjith, hebert}@cs.cmu.edu

Abstract

Despite the fact that color is a powerful cue in object recognition, the extraction of scale-invariant interest regions from color images frequently begins with a conversion of the image to grayscale. The isolation of interest points is then completely determined by luminance, and the use of color is deferred to the stage of descriptor formation. This seemingly innocuous conversion to grayscale is known to suppress saliency and can lead to representative regions being undetected by procedures based only on luminance. Furthermore, grayscaled images of the same scene under even slightly different illuminants can appear sufficiently different as to affect the repeatability of detections across images. We propose a method that combines information from the color channels to drive the detection of scale-invariant keypoints. By factoring out the local effect of the illuminant using an expressive linear model, we demonstrate robustness to a change in the illuminant without having to estimate its properties from the image. Results are shown on challenging images from two commonly used color constancy datasets.

1 Introduction

The representation of objects and scenes by sparse local patches has proven to be well suited to the practical challenges of occlusion, clutter and variation in viewing conditions. Approaches employing this representation [14] typically start with the detection of a set of distinguished points from the image. These "interest" points are often chosen as the extrema of functions of the input specifically constructed to have some desirable properties, e.g. high entropy, stability and invariance to geometric transformations. The detection step is followed by the computation of a descriptor at each interest point, possibly after the patch surrounding the point at its appropriate scale has been suitably normalized [13].

When the input is a color image, a common preprocessing step is the conversion of the image to grayscale. This conversion is convenient for several reasons. Firstly, by moving from our original vector-valued image to a scalar-valued image we make the available theory [11] and its proven guarantees of scale-invariant behavior directly applicable to our data. Secondly, it is well recognized that while color is a useful cue, its reliable use as a feature is hampered by various practical difficulties. The measured pixel values in a scene are influenced not just by the spectral reflectivity at that point, but also by the material properties, the sensitivities of the camera sensors and the spectral profile of the illuminant. Without additional information, inference of the "true" color at a pixel (defined, say as



Figure 1: Example of the effect of grayscale conversion on a synthetic color image under two perceptually similar illuminants. Note the repeatability of detected scale-invariant LOG features (circled). This figure is best seen in color.

seen with a reference illuminant) is often under-determined and necessitates modeling assumptions on a subset of these unknowns [5]. Reverting to grayscale and postponing the use of color to when computing the descriptor [18] is an attractive alternative.

However the conversion to grayscale has a number of side-effects that are particularly undesirable for reliable interest point detection. It is well recognized in neurobiology and computer graphics that grayscaled versions of color images do not preserve chromatic saliency [8, 12]. Regions that exhibit chromatic variation often lose their distinctive-ness when naively mapped to scalars based on isoluminance. Figure 1 (top row) shows an example of a Laplacian of Gaussian (LOG) based interest point detector applied to a grayscaled image from [8] with added Gaussian noise. It can be seen that the luminance values in the grayscale image do not exhibit the original differences in chrominance between many of the disks and the gray background as in the original. This causes the LOG-based detector to not detect significant stable extrema at several locations.

A related effect occurs due to change in illuminant. Using images of the *cruncheroos* scene under a white Sylvania 75W bulb (*halogen*) and fluorescent Philips Ultralume (*phulm*) tube from the Simon Fraser University (SFU) dataset [5], we can compute the linear transform that best numerically approximates the mapping of colors from *halogen* to *phulm*. These two illuminants are common to indoor scenes and the only perceivable visual difference between the two is in the form of a mild blue tone in the latter. We may then render the image to appear as it would under the *ph-ulm* illuminant in Figure 1 (bottom row). Note the difference between the grayscaled version of this image and that of the original. It can also be seen that the LOG-based detector picks a larger subset of keypoints from the content of the picture.

The above two observations indicate that the use of a grayscale intermediary to guide the selection of keypoints poses two kinds of risks. First, by attenuating differences in chrominance in favor of preserving luminance across the image, it can restrict the representative power of extracted regions. Secondly, by its sensitivity to the global illuminant, it can inconsistently accentuate parts of the scene leading to lowered repeatability of the detected keypoints. This paper addresses the question of how to appropriately combine information from the color channels to detect scale-invariant features in a manner that is robust to change in illuminant.

2 Approach

Our proposed method first constructs a one-parameter family of functions on the color image analogous to the scale-space representation of scalar-valued images [11]. The function at each level is designed to be relative invariant to the transformation group approximating the effect of the illuminant. The extrema of this scale-space representation along with their appropriately sized neighborhoods are chosen as the interest regions.

In Section 2.1 we look at two possible choices of linear transformation groups to approximate the effect of the illuminant. Section 2.2 then constructs normalized differential invariants that are relative invariant to these group and are subsequently used to form a linear scale-space representation of the color image.

2.1 Choice of Illumination Model

We choose to numerically approximate the effect of the illuminant as a linear transformation of the pixel RGB values. Finlayson [4] proposed that under this model, a diagonal transform is sufficient to discount the illuminant if a pre-computed transformation matrix, tuned to a particular illuminant pair, is first applied to the sensor outputs.

The diagonal model is sometimes also defended with Shafer's dichromatic reflectance model [3, 17] under the weaker assumption of sensor responses behaving as delta functions. Consider the "Mondrian" assumption of lambertian surfaces illuminated by a single light source. The reading $c_i(x)$ recorded by sensor *i* at location *x* may be expressed as:

$$c_i(x) = m_b(x) \int b(\lambda, x) e(\lambda) s_i(\lambda) d\lambda$$
(1)

where $m_b(x)$ models the effect of the illuminant geometry on surface reflectance (shadows, incident angle), $b(\lambda, x)$ denotes the surface reflectance, $e(\lambda)$ is the spectral profile of the illuminant and $s_i(\lambda)$ is the spectral sensitivity of the *i*-th sensor.

If the sensors $s_i(\lambda)$ behave as delta functions $\delta(\lambda - \lambda_i)$, the sensor response $c'_i(x)$ to a new illuminant $e'(\lambda)$ can then be trivially related by a scale factor as $c'_i(x) = \alpha c_i(x)$ with $\alpha = e'(\lambda_i)/e(\lambda_i)$, making the diagonal model valid exactly. Unfortunately, most consumer cameras are nowhere near obeying this assumption [1, 19]. Several researchers [6] have subsequently noted that a diagonal matrix model is often insufficient to explain observed data.

We may instead adopt a non-diagonal illumination model for the scene-independent terms $e(\lambda)s_i(\lambda)$ in Eqn. (1) as:

$$e'(\lambda)s_i(\lambda) \approx \sum_j \alpha_{ij} e(\lambda)s_i(\lambda)$$
 (2)

The quality of the diagonal approximation then depends on how well the function $e'(\lambda)s_i(\lambda)$ can be numerically approximated by $\alpha_i e(\lambda)s_i(\lambda)$ for the best possible choice of scaling factor α_i for all channels i = 1, 2, 3.

It is easy to verify that for real sensors and the choice of L_2 cost function, the optimal α_{ij} values reduce to zero for $i \neq j$ only if $s_i(\lambda)s_j(\lambda) = 0$ for all λ . That is, the diagonal

model is L_2 -optimal if the sensor responses are non-overlapping. This is in agreement with the principle behind spectral sharpening [1], and the observation that narrow-band sensors tend to obey the diagonal model better since their profiles are "closer" to delta functions.

To accommodate the common scenario of overlap in spectral response functions, we will consider both the full as well as the diagonal 3×3 matrix models. Hence, the illuminant dependent transformation is modeled as

$$\mathbf{c}'(x) = A\mathbf{c}(x) \tag{3}$$

where $\mathbf{c}'(x)$ and $\mathbf{c}(x)$ are the 3-vector responses at pixel *x* under two illuminants, and $A \in GL(3,\mathbb{R})$, the set of real non-singular 3×3 matrices or $A \in GD(3,\mathbb{R})$, the set of real non-singular diagonal 3×3 matrices.

2.2 Constructing Relative Invariants

Having fixed the transformation groups $GL(3,\mathbb{R})$ and $GD(3,\mathbb{R})$ to model the effect of a change in illuminant on pixel value, we now proceed to construct a 1-parameter scalar function of the input image that is invariant to this transformation as well as exhibits a semi-group property with respect to this parameter. We will then compute a stack of scalar functions of the image corresponding to increasing values of this parameter to form a scale-space representation of the image. The extrema in this 3-D representation will then constitute our desired result of a set of interest points and their scales.

Our procedure to construct a relative differential invariant is similar to that of [15], although that work focused on semi-differential invariants. We will denote the image f(x) by a 3-tuple of functions as $f(x) = \begin{bmatrix} r(x) & g(x) & b(x) \end{bmatrix}^T$ denoting the three color channels.

We will work with the larger group $GL(3,\mathbb{R})$ from which construction of invariants to the subgroup $GD(3,\mathbb{R})$ will be self-evident. We define a relative invariant to the transformation group $GL(3,\mathbb{R})$ as a real valued function $h : f(x) \to \mathbb{R}$ for which there exists a weight function $\xi : GL(3,\mathbb{R}) \to \mathbb{R}$ such that

$$h(Af) = \xi(A)h(f) \tag{4}$$

where $A \in GL(3, \mathbb{R})$. In the special case of $\xi = 1$, h(f) is called an absolute invariant.

From the definition of the group action, we know that h((AB)f) = h(A(Bf)) for all $A, B \in GL(3, \mathbb{R})$. This implies that for non-trivial invariants $(h \neq 0)$, this relation will hold only if the weight function satisfies the property $\xi(AB) = \xi(A)\xi(B)$ for all $A, B \in GL(3, \mathbb{R})$. It is a standard result in Lie group theory [15] that for the group $GL(3, \mathbb{R})$, the weight function must take the form

$$\xi(A) = [\det(A)]^{\beta} \tag{5}$$

for some $\beta \in \mathbb{R}$. We will consider the case of $\beta = 1$ since invariants for other values may be reduced to this case.

2.2.1 Case of 1-D signals

We first analyze the case of a 1-D signal, i.e. $f(x) : \mathbb{R} \to \mathbb{R}^3$. Consider the function

$$h(f) = \det \left(\begin{bmatrix} r(x) & r_x(x) & r_{xx}(x) \\ g(x) & g_x(x) & g_{xx}(x) \\ b(x) & b_x(x) & b_{xx}(x) \end{bmatrix} \right)$$
(6)



Figure 2: Interest points matched between the (a) reference *halogen* lamp and the (b-d) the *mb*-5000+3202 Macbeth 5000 tube with Roscolux #3202 full blue filter illuminants and rotation by 0° , 45° and 90° . Top row uses the invariant h_{full} and bottom row uses the diagonal invariant h_{diag} . Presented images have been contrast-enhanced for clarity.

where the subscripts denote first and second order derivatives of the signals. By the linearity of the differentiation operator, the effect of a change in illuminant modeled by a group element $A \in GL(3,\mathbb{R})$ can be expressed as

$$h(Af) = \det\left(\begin{bmatrix} A \begin{bmatrix} r(x)\\ g(x)\\ b(x) \end{bmatrix} & A \begin{bmatrix} r_x(x)\\ g_x(x)\\ b_x(x) \end{bmatrix} & A \begin{bmatrix} r_{xx}(x)\\ g_{xx}(x)\\ b_{xx}(x) \end{bmatrix} \end{bmatrix}\right) = \det(A) h(f)$$
(7)

using the property of determinants. Hence by our definition in Eqn. (5), h(f) is a relative invariant to the group $GL(3,\mathbb{R})$. The assumption of non-singularity of the transformation matrices avoids reduction to the trivial case of h(Af) = 0. Physically this corresponds to the absence of degenerate lighting such as monochromatic light sources that would elicit sensor response only in a small part of the spectrum.

In practice, differentiation is not a well-defined operator in the Hadamard sense, and we are forced to look for locally regularized alternatives for computing Eqn. (6). The pioneering work by Iijima [9] followed by Witkin [20] and others introduced the idea of constructing a regularized multi-resolution representations of a signal through a one-parameter family of functions. The functions were obtained by convolution with Gaussian kernels $G(x, \sigma)$ of increasing width parameter σ , also termed as the *scale*. Later work by Lindeberg [11] showed that by a clever change of variables, one can obtain normalized derivatives whose extrema positions are covariant to spatial scaling of the signal.

Following [11], we normalize the terms of the relative invariant in Eqn. (6) to yield

$$h\left(f(x),\sigma^{2}\right) = \det\left(\begin{bmatrix} G(x,\sigma^{2})*r(x) & \sigma G_{x}(x,\sigma^{2})*r(x) & \sigma^{2} G_{xx}(x,\sigma^{2})*r(x) \\ G(x,\sigma^{2})*g(x) & \sigma G_{x}(x,\sigma^{2})*g(x) & \sigma^{2} G_{xx}(x,\sigma^{2})*g(x) \\ G(x,\sigma^{2})*b(x) & \sigma G_{x}(x,\sigma^{2})*b(x) & \sigma^{2} G_{xx}(x,\sigma^{2})*b(x) \end{bmatrix} \right)$$

$$= \sigma^{3} \det \begin{pmatrix} \begin{bmatrix} G(x,\sigma^{2}) * r(x) & G_{x}(x,\sigma^{2}) * r(x) & G_{xx}(x,\sigma^{2}) * r(x) \\ G(x,\sigma^{2}) * g(x) & G_{x}(x,\sigma^{2}) * g(x) & G_{xx}(x,\sigma^{2}) * g(x) \\ G(x,\sigma^{2}) * b(x) & G_{x}(x,\sigma^{2}) * b(x) & G_{xx}(x,\sigma^{2}) * b(x) \end{bmatrix}$$
(8)

where the '*' denotes convolution and the subscripts denote the appropriate order of Gaussian derivative.

Combined with Eqn. (7), the scale invariance property of each term in the 3rd order polynomial formed by the determinant Eqn. (8) implies the property

$$h(Af(\gamma x), \sigma^2) = \det(A)h(f(x), \sigma^2/\gamma)$$
(9)

for a spatial scaling factor $\gamma \in \mathbb{R}$ and $A \in GL(3,\mathbb{R})$. Hence the set of 2-D extrema of the function $h(f, \sigma^2)$ in *x* and σ is invariant to illumination change and relative invariant to spatial scaling of the signal.

Note that because the effect of the illuminant is factored out in the invariant through the scalar det(A), the detection of extrema does not require the actual estimation of matrix A. Also, the effect of the illuminant through A is treated *locally* though the particular value of scale. Hence we may expect that the detection of extrema will be robust to gradual changes in the value of A across the image.

2.2.2 Case of 2-D signals

For 2-D images, we desire the additional property of invariance to spatial rotation. This is done by choosing the appropriate rotationally invariant 2-D form of the terms in Eqn. (8). We henceforth use $G(\mathbf{x}, \sigma^2)$ to denote a circular 2-D Gaussian.

The zero-th order term $G(\mathbf{x}, \sigma^2) * r(\mathbf{x})$ requires no modification as it is already rotationally invariant. Also the second order terms in the Eqn. (8) can be easily corrected by replacing G_{xx} with the rotationally symmetric Laplacian of Gaussian $\nabla G(\mathbf{x}, \sigma^2) = G_{xx}(\mathbf{x}, \sigma^2) + G_{yy}(\mathbf{x}, \sigma^2)$.

The first order derivative terms, however, cannot be corrected easily. Because the Gaussian function is radially symmetric, it is not possible to generate a linear combination of odd-order spatial derivatives that is invariant to rotation. Some remedies include (a) using a different family of scale-space generators such as the Poisson kernel [16] and its derivatives to replace the odd-order Gaussian derivatives, and (b) using regularized higher order even derivatives of $G(\mathbf{x}, \sigma^2)$. However the former requires more computation and the latter risks susceptibility to noise.

We opt to compute a local invariant frame $\{u, v\}$ at each pixel with its *u* axis aligned with the direction of maximal norm in intensity change in all 3 channels. The value of $G_x(\mathbf{x}, \sigma^2) * r(\mathbf{x})$ is then replaced by $G_u(\mathbf{x}, \sigma^2) * r(\mathbf{x})$. This is a computationally efficient alternative but with the drawback that it is only invariant to a subgroup of GL(3, \mathbb{R}). However, we show experimentally that it works well over a broad range of illuminants.

Another alternative is to settle for the diagonal illumination model which, as outlined in Section 2.1, will work well for narrow band sensors. This reduction in the number of model parameters vastly increases the space of available relative invariants. We will consider the invariant formed by a simple modification to the LOG detector as

$$h_{\text{diag}}(f(x), \sigma^2) = \sigma^6 \prod_{c = \{r, g, b\}} \nabla G(\mathbf{x}, \sigma^2) * c(\mathbf{x})$$
(10)



Figure 3: Scatter plot of LOG repeatability against the (a) full and (b) diagonal matrix illumination models between halogen and 10 other illuminants.

The final algorithm then proceeds as follows in a manner much similar to traditional interest point detection:

- 1. Construct pyramids of Gaussian blurred and 1st order derivatives of the color image following the procedure of [13].
- 2. Compute the invariant

 $h_{\text{full}}\left(f(x),\sigma^2\right) = \sigma^3 \det\left(\begin{bmatrix}G(\mathbf{x},\sigma^2) * f(x) & G_u(\mathbf{x},\sigma^2) * f(x) & \nabla G(\mathbf{x},\sigma^2) * f(x)\end{bmatrix}\right)$ (11)

(or h_{diag}) using the image channels and their derivatives at each pyramid level.

3. Find the extrema in the invariant scale-space pyramid and their corresponding scales.

3 Experiments

In this section we evaluate the repeatability of the detected regions across illuminants. We use images from two online color constancy datasets from SFU. The older dataset, which we label *objects98*, is associated with work in [5] and contains static household objects under 5 different illuminants. The second, labeled *mondrian*, is more recent [2] and contains scenes captured under as many as 11 illuminants. Images in both datasets were taken with a camera having narrow-band sensors [19].

We adopt the error metric of area repeatability used in the survey of [14] that compared affine-invariant detectors. The repeatability score between a pair of images is defined as the ratio of the number of pairs of interest points matched between the images to the minimum of the number of detected points in the pair. Two detected keypoints are determined to overlap if the ratio of their area of intersection to their union exceeds 60%.

Figure 2 shows an example of matched interest points between the reference *tide* image in *objects98* and rotated images under the harsh blue *mb*-5000+3202 illuminant. The top and bottom rows shows keypoints detected using the full and diagonal illuminant model respectively.



Figure 4: Variation in area repeatability with in-plane rotation between *halogen* and the *ph-ulm* and *syl-cwf* illuminants for the full matrix illumination model.

For each object in *objects98*, we pick its image under the *halogen* illuminant as the fixed reference image and record its pairwise repeatability rate with test images taken under the other illuminants. We also evaluate repeatability when the test images are transformed by in-plane rotation (Figure 4) and spatial scaling (Figure 5). Due to space limitations, we show only a subset of the objects for two test illuminants for only the full invariant h_{full} . It may be seen from the plots that the detection rate is quite stable to geometric transformations.

Figure 3 shows a scatter plot of repeatability rates for images from the *mondrian* dataset. Each point in the plot represents the repeatability rate using a luminance-based LOG detector and that with the full (h_{full}) or diagonal (h_{diag}) invariants for the indicated test illuminant. It can be seen that both invariants tend to either appreciably increase the repeatability rate or leave it relatively unchanged.

We consider the null hypothesis that the median difference in repeatability between the LOG and $h_{\text{full}}/h_{\text{diag}}$ detectors is zero. The Wilcoxon two-sided paired sign test convincingly rejects the null hypothesis for the h_{full} detector with a *p*-value of 1.21×10^{-5} at a 5% significance level, and similarly rejects the same for the h_{diag} detector with a *p*-value of 1.73×10^{-11} . Thus the use of illuminant invariants has a statistically significant and favorable influence on the repeatability rate.

4 Related Work

The goal of scale and illuminant invariance from color images relates to two kinds of endeavors in the literature, and to the authors' knowledge they have only been pursued independently. The first is that of the extraction of scale invariant features from scalar images [11], popularized by [13] and recently surveyed in [14] for the affine invariance setting. The detectors were designed to be at most invariant to affine transformations of the luminance values. Recent work in [18] independently formed color invariants and concatenated them to the chosen descriptor only after the initial detection of keypoints from luminance.

The second is that of color constancy and extracting features that are illumination invariant. The main challenge here is how to correctly combine information from each channel of a vector-valued image. Work in [7] constructed color invariants using the Kubelka-



Figure 5: Variation in area repeatability with spatial scaling between *halogen* and the *ph-ulm* and *syl-cwf* illuminants for the full matrix illumination model.

Munk model and was based on approximating the Gaussian function and its derivatives by linear combinations of implicitly known sensor profiles. Lenz et al [10] constructed semi-differential invariants for 2×2 transformations of color values, but only for a fixed pair of points having the same material properties. An early paper by Di Zenzo [21] proposed a modified structure tensor, termed the color tensor, that has been adapted in [3, 17] and others for detecting edges and corners from an illuminant invariant representation of the image. However the non-linearities normally associated with the formation of these invariants do not preserve the relationships required for a linear scale-space representation. The target illuminant invariance is also traditionally restricted to the diagonal model for simplicity [3, 18] or to a rigid rotation of the color coordinate axis [17].

In contrast to the above work, this paper focused on the use and construction of *joint* invariants that are covariant with spatial scaling while being robust to the traditional $GD(3,\mathbb{R})$ and the larger $GL(3,\mathbb{R})$ group of transformations.

5 Conclusions

The use of grayscale to represent a color image can have adverse effects depending on the distribution of colors in the image. While scale invariant detectors have reached a sufficient level of maturity, the next step, as also concluded in [14], consists of generalizing them to other representations such as RGB.

In this paper, we proposed to complement the existing class of interest region detectors by using an alternate intermediate representation. The constructed invariants enjoy robustness to a larger class of illuminant transformations that is traditionally addressed, while retaining the much desirable property of scale invariance.

Qualitatively we have observed that the detector using the full model finds fewer keypoints that its diagonal model counterpart. By our choice of h_{diag} , the latter also yields many keypoints at locations common to those found using the traditional LOG detector. As seen in Figure 3, the influence of using a full invariant is either large, particularly in scenes where the repeatability with the traditional luminance-based LOG detector is low, or negligible. We hypothesize that invariants to the larger GL(3, \mathbb{R}) group tend to form smoother functions with extrema that are not as well localized in some scenes as extrema of invariants to the diagonal subgroup. Hence, an appropriate practical strategy would be to combine the detector outputs using both the full and diagonal models. Future directions include a more rigorous characterization of possible joint invariants, experiments using images acquired with broad-band sensors, as well as the design of better metrics to compare detectors in a manner that is independent of the scheme for threshold selection.

References

- K. Barnard, F. Ciurea, and B. Funt. Sensor sharpening for computational color constancy. *Journal of the Optical Society of America*, 18(11):2728–2743, 2001. 3, 4
- [2] K. Barnard, L. Martin, B. Funt, and A. Coath. A data set for colour research. *Color Research and Application*, 27(3):147–151, 2002.
- [3] F. Faille. Stable interest point detection under illumination changes using colour invariants. In British Machine Vision Conference, 2005. 3, 9
- [4] G. Finlayson, M. Drew, and B. Funt. Diagonal transforms suffice for color constancy. In ICCV93, pages 164–171, 1993. 3
- [5] B. Funt, K. Barnard, and L. Martin. Is color constancy good enough? In *European Conference* on Computer Vision, pages 445–459, 1998. 2, 7
- [6] B. Funt and H. Jiang. Non-diagonal color correction. In Proc. Intl. Conference on Image Processing, 2003. 3
- [7] J. M. Geusebroek, R. van den Boomgaard, A. W. M. Smeulders, and H. Geerts. Color invariance. *IEEE Trans. Pattern Anal. Machine Intell.*, 23(12):1338–1350, 2001.
- [8] A. A. Gooch, S. C. Olsen, J. Tumblin, and B. Gooch. Color2Gray: salience-preserving color removal. ACM Trans. on Graphics, 24(3):634–639, 2005. 2
- [9] T. Iijima. Basic theory of pattern observation. Technical report, IECE, 1959. 5
- [10] R. Lenz, L. V. Tran, and T. H. Bui. Group theoretical invariants in color image processing. In 11th Color Imaging Conference, pages 212–216, 2003. 9
- [11] T. Lindeberg. Feature detection with automatic scale selection. Intl. Journal of Computer Vision (IJCV '98), 30(2):77–116, 1998. 1, 3, 5, 8
- [12] M. Livingstone. Vision and Art: The Biology of Seeing. Harry N. Abrams, Inc., 2002. 2
- [13] D. G. Lowe. Object recognition from local scale-invariant features. In International Conference on Computer Vision (ICCV '99), pages 1150–1157, 1999. 1, 7, 8
- [14] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. V. Gool. A comparison of affine region detectors. *Intl. Journal of Computer Vision*, 65(1/2):43–72, 2005. 1, 7, 8, 9
- [15] T. Moons, E. J. Pauwels, L. J. V. Gool, and A.Oosterlinck. Foundations of semi-differential invariants. *Intl. Journal of Computer Vision*, 14(1):25–47, 1995. 4
- [16] E. J. Pauwels, L. J. V. Gool, P. Fiddelaers, and T. Moons. An extended class of scale-invariant and recursive scale space filters. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 17(7):691–701, 1995. 6
- [17] J. van de Weijer, T. Gevers, and A. Smeulders. Robust photometric invariant features from the color tensor. *IEEE Trans. on Image Processing*, 15(1), 2006. 3, 9
- [18] J. van de Weijer and C. Schmid. Coloring local feature extraction. In European Conference on Computer Vision (ECCV), 2006. 2, 8, 9
- [19] P. Vora, J. E. Farrell, J. D. Tietz, and D. H. Brainard. Digital color cameras spectral response. Technical Report HPL-97-54, Hewlett Packard Laboratories, 1997. 3, 7
- [20] A. Witkin. Scale-space filtering. In IJCAI, volume 2, pages 1019–1022, 1983. 5
- [21] S. D. Zenzo. A note on the gradient of a multi-image. Computer Vision, Graphics, and Image Processing, 1986. 9