EVALUATING THE PERFORMANCE AND CORRELATION OF COLOUR INVARIANT LOCAL IMAGE FEATURE DETECTORS

Tony Marrero Barroso *Ovidiu Ghita* *Paul F. Whelan*

Centre for Image Processing and Analysis
Dublin City University, Dublin 9, Ireland.
www.cipa.dcu.ie
tony.marrero@eeng.dcu.ie, ghitao@eeng.dcu.ie, paul.whelan@dcu.ie

ABSTRACT

This paper evaluates the performance of state of the art colour invariants for the purposes of local image feature detection. We adapt the Harris-Laplace detector for colour invariance and test it under general image distortions. A second investigation examines the correlation between the colour invariants where the number of correctly detected unique points are analysed. This paper aims to answer if colour invariants should be used for feature detection purposes, and if they could be jointly used by feature fusion techniques to augment the performance of intensity-based detectors.

***Index Terms*—** colour detector, Harris-Laplace, local feature detection, photometric invariance, colour invariants.

1. INTRODUCTION

In the last two decades local invariant interest point detection has established itself as one of the most important research areas of computer vision. Local image features have proven successful in their tasks, as they can be made robust to varying viewing conditions such as scale, rotation and perspective changes. The majority of local feature detectors are still intensity-based only, in spite of substantial developments in various colour invariant models.

In this study, we implement scale-invariant colour local feature detectors and aim to answer two questions that will facilitate the integration of colour information into mainstream modern local feature detection. Firstly, this study evaluates the most promising colour gradient photometric invariants on two natural scene datasets, one with only illumination variations and the other is a dataset widely used in the evaluation of intensity-based detectors containing a large range of imaging distortions. Secondly, we investigate the correlation between the tested colour invariants on actual experimental results, which to our knowledge is here performed for the first time. Our study uses colour invariants to create a colour version of the Harris-Laplace [1] detector. For the purposes of feature detection (without a descriptor) most colour invariants have predominantly been evaluated using datasets with only lighting variations and containing images of individual objects like the Amsterdam Library of Object Images[1]. In the context of feature detection, the literature lacks an evaluation and comparison of the most prominent colour invariants using image data that contain image distortions other than illumination.

The first part of our study addresses the limitations of previous works, by evaluating colour features more rigorously using the distortion criteria that were employed in the testing of state of the art intensity-based detectors (ie. scale, viewpoint, blurring, JPEG compression and illumination). The second part is related to the complementarity between the tested colour invariants. Our correlation analysis reports the number of unique correct points that each gradient type can generate. It also identifies which gradient types could be used conjointly, to obtain a better overall performance.

The overall contributions of this paper can be summarised as follows: The robust evaluation of state of the art colour invariants in the context of feature detection under the presence of typical image distortions. Secondly, investigating their correlation to intensity and identifying if they can be utilised to augment intensity-based detectors. The results will help to maximise the distinctiveness and robustness of local image feature detection.

1.1. Previous Work

The most successful intensity-based local image features are gradient-based, and rely on scale-invariant corner and blob detection, like the well known Harris-Laplace [1], and LoG (Laplacian of Gaussian) [2] detectors. For this reason we focus our colour feature detection study on gradient-based approaches. In the case of colour-based detectors, the most stable and robust to illumination variations as was shown in [3], have been based on the colour Harris introduced by Montesinos *et* al. [4]. Van de Weijer *et* al. [5, 6] extended the

This research has been funded by the Irish Research Council for Science, Engineering and Technology: funded by the NDP.

[1] http://staff.science.uva.nl/ aloi

colour Harris by proposing a set of photometric variants and quasi-invariants. Their evaluation was limited however, to non-scale-invariant corner detection. Geusebroek *et* al. [7] propose a set of photometric colour invariants using the Gaussian colour model. The performance of local image descriptors created from these colour invariants were evaluated by Burghouts and Geusebroek [8] and some proved to be superior to gray-scale intensity. Stöttinger *et* al. [9] propose a light-invariant Harris-Laplace feature detector in the context of image retrieval, though their focus was not to fully compare their colour invariant gradient with others from the literature. Faille [10] proposes a colour Harris corner detector which is invariant to specularities, shading and colour illumination. However the method uses fixed scales for matching images under illumination distortions. Unnikrishnan and Herbert [11] evaluate two illuminant invariant functions on the RGB space, they detect scale and rotation invariant points using the LoG operator and obtain better results than using only the intensity under illumination variations. One limitation of their study is the lack of an evaluation of their detector under viewpoint changes.

2. COLOUR FEATURE POINT DETECTION

Our study adapts the Harris-Laplace [1] (HL) detector for colour feature detection using colour invariants. It has been one the most widely used gradient-based detectors, and was shown to be reliable under rotation, scale and illumination changes along with limited perspective deformations [1]. Due to space restrictions we briefly refer here to the relevant mathematical formulas required to explain the HL's adaptation for colour detection. The Harris detector is based on the second moment matrix, that is often used to describe local image gradient distributions. For an image *I*, the scale-adapted structure tensor at position $\mathbf{x}$ is given by Eqn.1:

$$H(\mathbf{x},\sigma_I,\sigma_D) = \sigma_D^2 G(\sigma_I) \otimes \begin{bmatrix} L_x^2(\sigma_D) & L_xL_y(\sigma_D) \\ L_xL_y(\sigma_D) & L_y^2(\sigma_D) \end{bmatrix} \quad (1)$$

The image gradients (L_x, L_y) are computed by convolution with the first derivatives of the Gaussian kernel with standard deviation σ_D. These derivatives are then convolved with G(σ_I), the Gaussian kernel with standard deviation σ_I. The other relevant expression is the operator that provides the scale-invariance; the scale normalised Laplacian (Eqn. 2). $L_{xx}(\mathbf{x},\sigma_n)$ denotes the response at image location $\mathbf{x}$ of the convolution of the second derivative of the Gaussian with the original input image (*in the x-direction, with std. dev.* σ_n). The LoG response is indicative of the similarity between the LoG kernel and the local image structure on which it is being convolved with. To achieve scale-invariance, a scale-space image stack is constructed by convolving the input image and the LoG with increasing σ_n. When the LoG response results in a local 3D maxima across the scale-space image stack, then a characteristic scale for that local structure exists at that location in the scale-space representation. To summarise the HL detector, Eqn.1 is used to detect corners of various sizes and Eqn. 2 allows for a characteristic scale to be estimated for those corners.

$$|LoG(\mathbf{x},\sigma_n)| = \sigma_n^2 |L_{xx}(\mathbf{x},\sigma_n) + L_{yy}(\mathbf{x},\sigma_n)| \quad (2)$$

2.1. Color Invariant Gradients

For our colour detectors, we utilise Eqns. 1 and 2 but colour invariants are used instead of the standard gray-scale intensity gradients. These invariants are formed from four colour spaces: The Opponent Colour Space $(O1, O2, O3)$ [5]. Hue Saturation and Intensity (H, S, I) [5]. The Spherical Colour Space (r, θ, φ) [5]. The Gaussian Colour Model $(\hat{E}, \hat{E}_\lambda, \hat{E}_{\lambda\lambda})$ [7], shown in Eqn. 3. Due to space restrictions we can only show the formula for the Gaussian model as it is a more uncommon colour space, please refer to the literature [5] for the expressions of the other colour spaces.

Our evaluation uses 9 different colour gradients: The Light-Invariant $(Light_{INV})$ gradient [9]. The specular-shadow-shading quasi-invariant $(SPSS_{INV})$ [5]. The specular quasi-invariant (SP_{INV}) [5]. The shadow-shading quasi-invariant (SS_{INV}) [5]. The shadow-shading full invariant (SS_{F-INV}) [6]. The specular-shadow-shading variant $(SPSS_{VAR})$ [5] is not an invariant gradient but it is included for a more complete evaluation and comparison of the invariants alongside the grayscale intensity. The last three invariants (C_{INV}, H_{INV} and W_{INV}) are proposed by Geusebroek *et* al. [7] and they use the Gaussian Color Model.

$$\begin{pmatrix} \hat{E} \\ \hat{E}_\lambda \\ \hat{E}_{\lambda\lambda} \end{pmatrix} = \begin{pmatrix} 0.06 & 0.63 & 0.27 \\ 0.3 & 0.04 & -0.35 \\ 0.34 & -0.6 & 0.17 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (3)$$

To summarise the implementation of all the colour invariants, we show in Table 1 how they are obtained from their respective colour spaces. The invariants of Van de Weijer *et* al. [5, 6] were adapted from their released code[2] (*Color Feature Detection I & II*) . Our notations are the same used in the authors original works, and the subcripts (λ) of $\hat{E}_\lambda$ and $\hat{E}_{\lambda\lambda}$ refer to the first and second colour channels of the Gaussian model. The term L_x in Table 1 refers to the L_x of the structure tensor (Eqn. 1), the x subscript denotes the first order derivatives in the x-direction. To obtain the second order L_{xx} of the LoG operator, the second derivatives of the colour channels are used instead for the expressions of Table 1.

3. EXPERIMENTS

The first part of these experiments evaluates the performance of the aforementioned colour invariants when used in local feature detection of natural scene images using the HL detector. The invariants are also compared with the standard

[2]http://cat.cvc.uab.es/ joost/software

Table 1: Summary of the implementation of the colour invariants.

Method	L_{INV}	SPSS_{INV}	SP_{INV}
L_x	$\sqrt{(H_x S)^2 + (S_x)^2}$	$H_x S$	$\sqrt{(O1_x)^2 + (O2_x)^2}$
Method	SS_{INV}	SPSS_{VAR}	SS_{F-INV}
L_x	$r\sqrt{(\varphi_x)^2 + (\sin(\varphi)\theta_x)^2}$	$\sqrt{(I_x)^2 + (S_x)^2}$	$\sqrt{(\varphi_x)^2 + (\sin(\varphi)\theta_x)^2}$
Method	C_{INV}	H_{INV}	W_{INV}
L_x	$\sqrt{\left(\frac{\hat{E}\hat{E}_{\lambda x} - \hat{E}_\lambda \hat{E}_x}{\hat{E}^2}\right)^2 + \left(\frac{\hat{E}\hat{E}_{\lambda\lambda x} - \hat{E}_{\lambda\lambda} \hat{E}_x}{\hat{E}^2}\right)^2}$	$\sqrt{\left(\frac{\hat{E}_{\lambda\lambda}\hat{E}_{\lambda x} - \hat{E}_\lambda \hat{E}_{\lambda\lambda x}}{\hat{E}_\lambda^2 + \hat{E}_{\lambda\lambda}^2}\right)^2}$	$\sqrt{\left(\frac{\hat{E}_x}{\hat{E}}\right)^2 + \left(\frac{\hat{E}_{\lambda x}}{\hat{E}}\right)^2 + \left(\frac{\hat{E}_{\lambda\lambda x}}{\hat{E}}\right)^2}$

intensity gradients (I_x). The second experiment is a correlation analysis to investigate which colour gradient generates the most unique number of correctly matched points. A correlation matrix shows the complementarity between all the 10 tested gradient types and is useful when considering the fusion of multiple gradients to achieve a higher performance than using a single gradient type on its own.

Our evaluations are carried out on two datasets: the Oxford[3] dataset which has become the de facto database to evaluate local features, and the Middlebury Stereo[4] dataset. Mikolajczyk's Oxford dataset consists of image sets with various distortions: blurring, zoom and rotation, JPEG compression, illumination and viewpoint changes. We used all sets (7 colour sets) except the black and white one which is not relevant to this study. The Middlebury Stereo dataset provided by Scharstein and Pal [12], consists of multiple sets of stereo images of natural scenes. These image sets vary in illumination conditions and we utilise a set of images (from 5 different scenes) that contain varying illumination but no viewpoint changes. Examples of these images are shown in Fig. 1.

3.1. Feature Detection Experiments

Mikolajczyk and Schmid [13] propose an approach to evaluate the quality of local interest point detection using robust metrics. We follow the same evaluation method in this study to provide standardized results. In [13] two features are considered correctly matched if the detected areas overlap by more than 40%. We set a more strict threshold of 10% to evaluate the localisation stability of the colour gradients more robustly. See Fig. 2 (a,d) for a summary of the results, showing the mean number of correct correspondances across all distortion levels for both datasets.

These detection results show that the intensity is the overall best performer. Only one invariant (W_{INV}) is superior to the intensity when applied to the illumination varying Middlebury dataset. The colour invariants were clearly inferior when applied to the Oxford dataset as they prove to be less robust to imaging distortions, namely scale and viewpoint changes. These results prove the necessity to evaluate colour invariants under a more general set of imaging conditions other than just illumination. To conclude, the overall performance of the tested colour invariants is insufficient to merit their independent adoption for local feature detection tasks. The following correlation study therefore investigates their potential for complementing the intensity channel.

3.2. Correlation of Color Gradients

Results for the correlation of the gradients are experimentally obtained from the feature detection results. We calculate the correlation between two gradient types by the percentage of correctly matched points that are common between them, out of the total number of points that can potentially be matched. The unique correct matches for a gradient type are denoted as the matched points that only that particular gradient type extracts (amongst all the 10 tested gradients). The unique matches results are shown in Fig. 2 (b, e) and the average correlation matrix plots are shown in Fig. 2 (c, f).

[3] www.robots.ox.ac.uk/ vgg/research/affine

[4] http://vision.middlebury.edu/stereo/data

Fig. 1: Top row: Oxford dataset examples. Bottom row: Samples of the Middlebury '*Art*' image set.

In Fig. 2 (b) it can be seen that with the Middlebury data the top 2 colour invariants (W_{INV} and $Light_{INV}$) generate comparable number of correct unique matches to the intensity. The Oxford dataset is again more problematic for the colour gradients and their relative performance to the intensity is here poorer. In spite of this, the W_{INV} and $Light_{INV}$ invariants on the Oxford set still generate in total 105% of the number of unique points that the intensity gradients are able to match (at the lowest distortion level), and 113% at the highest distortion. For the Middlebury set, these figures are 180% and 166% respectively. The capacity to incorporate more unique matches into the feature detection step using colour invariants is thus significant. This feature fusion could result in a final set of detected features that are more robust to distortions and more unique, benefiting subsequent feature description and matching tasks.

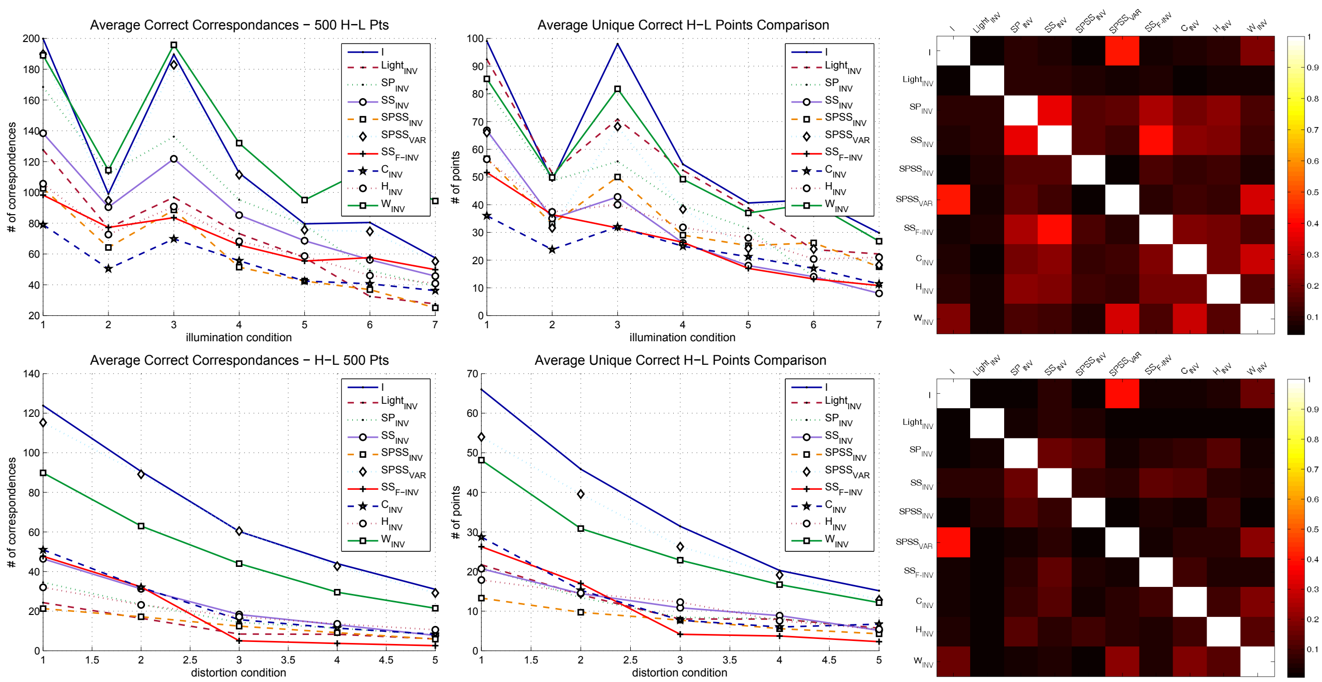
The correlation matrices indicate that all of the colour invariants are in fact highly uncorrelated with the intensity and are thus capable of positively influencing the performance of intensity if used conjointly. The summary of this correlation study is that the top 2 colour invariants provide a considerable number of unique points. A further feature fusion study is needed to properly evaluate the overall gain in performance, but the results from our study provide clear indications that colour invariants can indeed augment the performance of intensity-based feature detection.

4. CONCLUSION

This study evaluated the performance of colour photometric invariants in the context of local feature detection using natural scenes containing the typical set of imaging distortions that state of the art intensity-based detectors have been evaluated with. We found that intensity is the overall top performer, it also performs comparatively to the colour invariants under illumination variations. Colour invariants proved to have poorer robustness to the general imaging distortions (ie. scale and viewpoint) inherent in the Oxford dataset. The dominance of intensity-based detectors, can be largely attributed to the luminance axis containing the majority of the variation in the RGB-cube [14], and the stability of the localisation of its gradients. For these reasons the intensity performs best overall for general imaging conditions, and only under more severe illumination conditions should colour gradients be considered for local feature detection.

Despite their overall inferiority when utilised individually, our correlation study obtained promising results and suggests that colour invariants should indeed be considered for feature fusion as they are uncorrelated to the intensity and generate considerable number of unique matches. The colour invariants that obtained the optimal balance of uniqueness and are least correlated to intensity, are W_{INV} [7] and $Light_{INV}$ [9]. These have the biggest potential to be used in conjuction with intensity in future local image feature fusion studies.

(a) Avg. Correct pts. (Middlebury). (b) Avg. Unique pts. (Middlebury). (c) Avg. Correlation (Middlebury).



(d) Avg. Correct pts. (Oxford). (e) Avg. Unique pts. (Oxford). (f) Avg. Correlation (Oxford).

Fig. 2: Summary of Correct Correspondances and Correlation Analysis.

5. REFERENCES

[1] K. Mikolajczyk and C. Schmid, "Scale & affine invariant interest point detectors," *International Journal of Computer Vision*, vol. 60, no. 1, pp. 63–86, 2004.

[2] T. Lindeberg, "Feature detection with automatic scale selection," *International Journal of Computer Vision*, vol. 30, no. 2, pp. 79–116, 1998.

[3] V. Gouet and N. Boujemaa, "Object-based queries using color points of interest," in *IEEE CVPR Workshop on Content-Based Access of Image and Video Libraries (CBAIVL)*, 2001, pp. 30–36.

[4] P. Montesinos, V. Gouet, and R. Deriche, "Differential invariants for color images," in *Proc. IEEE International Conference on Pattern Recognition*. IEEE, 1998, vol. 1, pp. 838–840.

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

[6] J. Van De Weijer, T. Gevers, and A. W. Smeulders, "Robust photometric invariant features from the color tensor," *IEEE Trans. on Image Processing*, vol. 15, no. 1, pp. 118–127, 2006.

[7] J-M. Geusebroek, R. van den Boomgaard, A. W. M. Smeulders, and H. Geerts, "Color invariance," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23, no. 12, pp. 1338–1350, 2001.

[8] G. J. Burghouts and J-M. Geusebroek, "Performance evaluation of local colour invariants," *Computer Vision and Image Understanding*, vol. 113, no. 1, pp. 48–62, 2009.

[9] J. Stottinger, A. Hanbury, N. Sebe, and T. Gevers, "Sparse color interest points for image retrieval and object categorization," *IEEE Trans. on Image Processing*, vol. 21, no. 5, pp. 2681–2692, 2012.

[10] F. Faille, "Stable interest point detection under illumination changes using colour invariants," in *Proc. of the British Machine Vision Conference*, 2005, pp. 19–1.

[11] R. Unnikrishnan and M. Hebert, "Extracting scale and illuminant invariant regions through color," in *Proc. of the British Machine Vision Conference*, 2006, pp. 52–1.

[12] D. Scharstein and C. Pal, "Learning conditional random fields for stereo," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2007, pp. 1–8.

[13] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, pp. 1615–1630, 2005.

[14] J. Van De Weijer, T. Gevers, and A. D. Bagdanov, "Boosting color saliency in image feature detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 28, no. 1, pp. 150–156, 2006.