

Colour Constancy: Biologically-inspired Contrast Variant Pooling Mechanism

Arash Akbarinia¹
www.cvc.uab.es/people/sakbarinia/

Raquel Gil Rodríguez²
<http://ip4ec.upf.edu/user/53>

C. Alejandro Parraga¹
www.cvc.uab.es/people/aparraga/

¹ Centre de Visió per Computador
Universitat Autònoma de Barcelona
Barcelona, Spain

² Department of Information and
Communication Technologies
Universitat Pompeu Fabra
Barcelona, Spain

Abstract

Pooling is a ubiquitous operation in image processing algorithms that allows for higher-level processes to collect relevant low-level features from a region of interest. Currently, max-pooling is one of the most commonly used operators in the computational literature. However, it can lack robustness to outliers due to the fact that it relies merely on the peak of a function. Pooling mechanisms are also present in the primate visual cortex where neurons of higher cortical areas pool signals from lower ones. The receptive fields of these neurons have been shown to vary according to the contrast by aggregating signals over a larger region in the presence of low contrast stimuli. We hypothesise that this contrast-variant-pooling mechanism can address some of the shortcomings of max-pooling. We modelled this contrast variation through a histogram clipping in which the percentage of pooled signal is inversely proportional to the local contrast of an image. We tested our hypothesis by applying it to the phenomenon of colour constancy where a number of popular algorithms utilise a max-pooling step (*e.g.* White-Patch, Grey-Edge and Double-Opponency). For each of these methods, we investigated the consequences of replacing their original max-pooling by the proposed contrast-variant-pooling. Our experiments on three colour constancy benchmark datasets suggest that previous results can significantly improve by adopting a contrast-variant-pooling mechanism.

1 Introduction

Many computer vision frameworks contain a pooling stage that combines local responses at different spatial locations [0]. This operation is often a sum- or max-pooling mechanism implemented by a wide range of algorithms, *e.g.* in feature descriptors (such as SIFT [16] and HOG [10]) or convolutional neural networks [35, 12]. Choosing the correct pooling operator can make a great difference in the performance of a method [0]. Current standard pooling mechanisms lack the desired generalisation to find equilibrium between frequently-occurring and rare but informative descriptors [33]. Therefore, many computer vision applications can benefit from a more dynamic pooling solution that takes into account the content of pooled signals.

Such pooling operators are commonly used in modelling the phenomenon of colour constancy (*i.e.* a visual effect that makes the perceived colour of surfaces remain approximately constant under changes in illumination [18, 21]) both in biological [9] and computational [8] solutions. Despite decades of research, colour constancy still remains as an open question [18], and solutions to this phenomenon are important from a practical point of view, *e.g.* camera manufacturers need to produce images of objects that appear the same as the actual objects in order to satisfy customers. Motivated by above, in this article we propose a contrast-variant-pooling mechanism and investigate its feasibility in the context of computational colour constancy.

1.1 Computational Models of Colour Constancy

Mathematically, the recovery of spectral reflectance from a scene illuminated by light of unknown spectral irradiance is an ill-posed problem (it has infinite possible solutions). The simplest and most popular solution has been to impose some arbitrary assumptions regarding the scene illuminant or its chromatic content. Broadly speaking colour constancy algorithms can be divided into two categories: (i) low-level driven, which reduce the problem to solving a set of non-linear mathematical equations [22, 33, 45], and (ii) learning-based, which train machine learning techniques on relevant image features [10, 17, 20]. Learning-based approaches may offer the highest performance results, however, they have major setbacks which make them unsuitable in certain conditions: (a) they rely heavily on training data that is not easy to obtain for all possible situations, and (b) they are likely to be slow [23] and unsuitable for deployment inside limited hardware. A large portion of low-level driven models can be summarised using a general Minkowski expression [14, 45] such as:

$$L_c(p) = \left(\int_{\Omega} [f_c(x)]^p dx \right)^{\frac{1}{p}} = ke_c, \quad (1)$$

where L represents the illuminant at colour channel $c \in \{R, G, B\}$; $f(x)$ is the image's pixel value at the spatial coordinate $x \in \Omega$; p is the Minkowski norm; and k is a multiplicative constant chosen such that the illuminant colour, e , is a unit vector.

Distinct values of the Minkowski norm p results into different pooling mechanisms. Setting $p = 1$ in Eq. 1 reproduces the well known Grey-World algorithm (*i.e.* sum-pooling), in which it is assumed that all colours in the scene average to grey [8]. Setting $p = \infty$ replicates the White-Patch algorithm (*i.e.* max-pooling), which assumes the brightest patch in the image corresponds to the scene illuminant [33]. In general, it is challenging to automatically tune p for every image and dataset. At the same time inaccurate p values may corrupt the results noticeably [23].

The Minkowski framework can be generalised further by replacing $f(x)$ in Eq. 1 with its higher-order derivatives [45]. These non-linear solutions are analogous to centre-surround mechanisms of visual cortex [34], which is typically modelled by a Difference-of-Gaussians (DoG) operators where a narrower, positive Gaussian plays the role of the "centre" and a broader, negative Gaussian plays the role of the "surround" [13, 37]. Recently, a few biologically-inspired models of colour constancy grounded on DoG have offered promising results [22, 40], however their efficiency largely depends on finding an optimum pooling strategy for higher cortical areas. In short, pooling is a crucial component of many colour constancy models driven by low-level features (or even in deep-learning solutions [6, 19]). In the primate visual systems the size of the receptive field varies according to the local

contrast of the light falling on it [4, 24] presenting a dynamic solution dependent on the region of interest.

The low-level models mentioned above are related to the early stages of visual processing, *i.e.* the primary visual cortex (area V1), that are likely to be involved in colour constancy. Physiological measures suggest that although receptive fields triple in size from area V1 to area V2 [7], their basic circuitry with respect to surround modulation is similar, *i.e.* keeping the same size dependency with contrast properties found in V1 [24]. This is consistent with the large body of physiological and psychophysical literature highlighting the significance of contrast in the visual cortex. In computer vision, contrast-dependent models have also shown encouraging results in various applications such as visual attention [28], tone mapping [41], and boundary detection [2], to name a few. From these we can hypothesise the convenience and explanatory value of various “pooling strategies” such as those proposed by previous colour constancy methods. In the rest of this work we will explore the advantages of replacing the different feed-forward (pooling) configurations of some successful colour constancy models [22, 33, 45] by that of the visual system (as described by current neurophysiological literature [4, 24]). Our aim in doing so is dual, on the one hand we want to explore the technological possibilities of creating a more efficient algorithm and on the other hand we would like to test the idea that contrast-variant-pooling might play an important role in colour constancy.

1.2 Summary of the Contributions

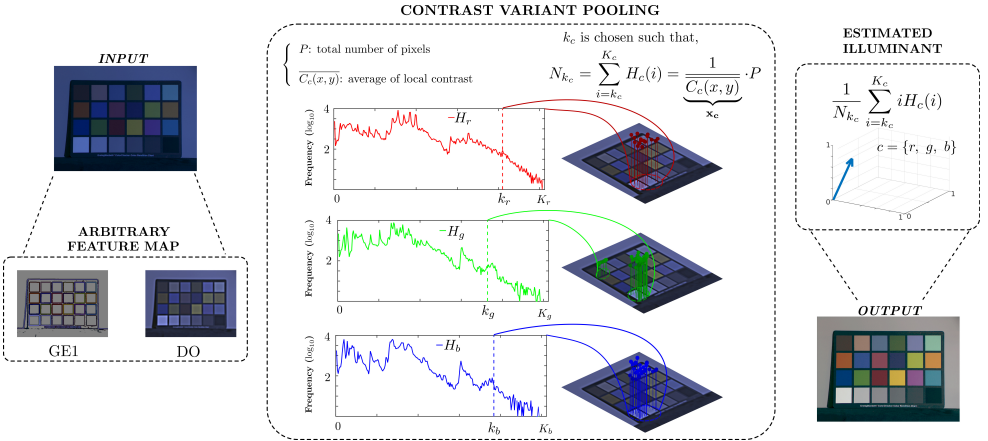


Figure 1: Flowchart of the proposed contrast-variant-pooling (CVP) mechanism in the context of colour constancy. We implemented CVP through a *top-x-percentage-pooling*. Given an input image or a feature map: (i) value of the x is computed according to the inverse of local contrast for each channel, and (ii) we estimate the scene illuminant as the average value of the *top-x-percentage* of pooled pixels (those on the right side of the depicted dashed lines).

In the present article we propose a generic contrast-variant-pooling (CVP) mechanism that can replace standard sum- and max-pooling operators in a wide range of computer vision applications. Figure 1 illustrates the flowchart of CVP, which is based on local contrast and therefore it offers a dynamic solution that adapts the pooling mechanism accord-

ing to the content of region of interest. We tested the feasibility of CVP in the context of colour constancy by substituting the pooling operation of four algorithms: (a) White-Patch [33], (b) first-order Grey-Edge [45], (c) second-order Grey-Edge [45], and (d) Double-Opponency [22] in three benchmark datasets. Results of our experiments show the quantitative and qualitative benefits of CVP over max-pooling.

2 Model

2.1 Max-pooling Colour Constancy

One of the earliest computational models of colour constancy (White-Patch) is grounded on the assumption that the brightest pixel in an image corresponds to a bright spot or specular reflection containing all necessary information about the scene illuminant [33]. Mathematically this is equivalent to a max-pooling operation over the intensity of all pixels:

$$L_c = \arg \max_{x,y} I_c(x,y), \quad (2)$$

where L represents the estimated illuminant at each chromatic channel $c \in \{R, G, B\}$; I is the original image and (x, y) are the spatial coordinates in the image domain.

One important flaw of this simple approach is that a single bright pixel can misrepresent the whole illuminant. Furthermore, the White-Patch algorithm may fail in the presence of noise or clipped pixels in the image due to the limitations of the max-pooling operator [22]. One approach to address these issues is to account for a larger set of “white” points by pooling a small percentage of the brightest pixels (*e.g.* the top 1%) [22], an operation referred as *top- x -percentage-pooling*. In this manner, the pooling mechanism is collectively computed considering a group of pixels rather than a single one. This small variant might be a crucial factor in the estimation of the scene illuminant [33]. A similar mechanism has also been deployed successfully in other applications such as shadow removal [46].

In practice, given the chosen *x*-percentage, the *top- x -percentage-pooling* can be implemented through a histogram-based clipping mechanism [22, 46]. Let H be the histogram of the input image I , and let $H_c(i)$ represents number of pixels in colour channel c with intensity $i \in [0, \dots, K_c]$ (histogram’s domain). The scene illuminant L_c is computed as:

$$L_c = \frac{1}{N_{k_c}} \sum_{i=k_c}^{K_c} i \cdot H_c(i), \quad (3)$$

where N_{k_c} is the total number of pixels within the intensity range k_c to K_c . The values of $k_{c \in \{R, G, B\}}$ are determined from the chosen *x*-percentage such that:

$$N_{k_c} = \sum_{i=k_c}^{K_c} H_c(i) = x_c \cdot P, \quad (4)$$

where P is the total number of pixels in the image and x_c is the chosen percentage for colour channel c . Within this formulation it is very cumbersome to define a universal, optimally-fixed percentage of pixels to be pooled [22] and consequently the free variable x requires specific tuning for each image or dataset individually. Naturally, this limitation restricts the usability of the *top- x -percentage-pooling* operator. In the following sections we show how to automatically compute this percentage, based on the local contrast of each image.

2.2 Pooling Mechanisms in the Visual Cortex

We know from the physiology of the cerebral cortex that neurons in higher cortical areas pool information from lower areas over increasingly larger image regions. Although the exact pooling mechanism is yet to be discovered, “winner-takes-all” and “sparse coding” kurtotical behaviour are common to many groups of neurons all over the visual cortex [9, 69], and it is conceivable that a mechanism analogous to max-pooling might be present within the cortical layers. Indeed such behaviour has been discovered across a population of cells in the cat visual cortex [32] and the activation level of cells with max-like behaviour was reported to vary depending on the contrast of visual stimuli.

Results reported by [32] suggest an inverse relationship between the contrast of a stimulus and the percentage of the signal pooled. When pooling neurons were exposed to low contrast stimuli, their responses shifted slightly away from pure max-pooling (selecting the highest activation response within a region) towards integrating over a larger number of highly activated neurons. In the language of computer vision, this can be regarded as *top-x-percentage-pooling*, where x assumes a smaller value in high contrast and a larger value in low contrast. Interestingly, the pooling of those neurons remained always much closer to max-pooling than to the linear integration of all neurons (sum-pooling) [32]. Mathematically, this can be interpreted as having a very small (*top-x-percentage-pooling*) x value.

It does not come as a great surprise that the pooling mechanism in the visual cortex depends on the stimulus’ contrast. There is a large body of physiological studies showing that receptive fields (RF) of neurons are contrast variant (for a comprehensive review refer to [9]). Quantitative results suggest that RFs in visual area one (V1) of awake macaques double their size when measured at low contrast [44]. Similar expansions have also been reported for the RFs of neurons in extrastriate areas, such as V2 and V4. This implies that a typical neuron in higher cortical areas that normally pool responses from its preceding areas over about three neighbouring spatial locations [47] can access a substantially larger region to pool from in the presence of low contrast stimuli. This is in line with the reported pooling mechanism in the cat visual cortex [32].

2.3 Contrast Variant Pooling

In order to model this contrast-variant-pooling mechanism, we first computed local contrast C of the input image I at every pixel location by means of its local standard deviation defined as:

$$C_c(x, y; \sigma) = \sqrt{\frac{1}{\#\mathcal{N}_\sigma(x, y)} \sum_{(x', y') \in \mathcal{N}_\sigma(x, y)} (I_c(x', y') - \mu(x, y))^2} \quad (5)$$

where c indexes each colour channel; (x, y) are the spatial coordinates of a pixel; μ is an isotropic average kernel with size σ ; and $\mathcal{N}_\sigma(x, y)$ represents the neighbourhood centred on pixel (x, y) of radius σ .

To simulate this inverse relation between stimulus contrast and percentage of signal pooled [32] in the *top-x-percentage-pooling* operator, we determined the percentage x_c in Eq. 4 as the average of inverse local contrast signals:

$$x_c = \frac{1}{P} \sum_{x, y} \frac{1}{C_c(x, y; \sigma)}, \quad \forall (x, y) \in \Omega, \quad (6)$$

where C_c is computed from Eq. 5, and Ω is the spatial image domain. In this fashion, instead of defining a fix percentage of signal (pixels) to be pooled (as in [46]), we chose an adaptive

percentage according to the contrast of image. In terms of colour constancy, this effectively relates the number of pooled pixels to compute the scene illuminant to the average contrast of an image. We illustrated this mechanism of contrast-variant-pooling in the central panel of Figure 1, where red, green and blue signals correspond to the histogram of each chromatic channel. Pixels on the right side of the dashed lines (k_c) are pooled. In this example, contrast is higher for the red signal and therefore a smaller percentage of cells are pooled in the red channel.

Bearing in mind that “contrast” is just a fraction in the range $[0, 1]$ – with 0 characterising an absolutely uniform area and 1 representing points with the highest contrast, *e.g.* edges – we can predict that the percentage x will be a very small number for natural images where homogeneous regions are likely to form the majority of the scene. Consequently in our *top-x-percentage-pooling* operator we always pool a small percentage. This is in agreement with observations of [52] which indicate that such pooling mechanism is always much closer to max-pooling than to sum-pooling.

2.4 Generalisation to Other Colour Constancy Models

There is a number of colour constancy models in the literature which are driven by low-level features that require a pooling mechanism on top of their computed feature maps in order to estimate the scene illuminant. In the Double-Opponency [22] algorithm this feature map is computed by convolving a colour-opponent representation of the image with a DoG kernel followed by a max-pooling operation. In the Grey-Edge [45] model, higher-order derivatives of the image are calculated through its convolution with the first- and second-order derivative of a Gaussian kernel. This is complemented by a Minkowski summation, which oscillates between sum- and max-pooling depending on its norm.

Similar to the White-Patch algorithm, the pooling mechanism of these models can also be replaced with our *top-x-percentage-pooling* operator, where x is computed according to the local contrast of image as explained above. The only difference is that instead of pooling from an intensity image (as in case of the White-Patch algorithm), Double-Opponency and Grey-Edge pool over their respective feature maps. This means that Eq. 3 receives a feature map M instead of an intensity image I as input.

3 Experiments and Results

In order to investigate the efficiency of our model, we applied the proposed contrast-variant-pooling (CVP) mechanism to four different colour constancy algorithms whose source code were publicly available: White-Patch [53], first-order Grey-Edge [45], second-order Grey-Edge [45], and Double-Opponency [22]. We simply replaced their max-pooling operator with our proposed pooling mechanism. To evaluate each method we used the recovery angular error defined as:

$$\varepsilon^\circ(e_e, e_t) = \arccos\left(\frac{\langle e_e, e_t \rangle}{\|e_e\| \|e_t\|}\right), \quad (7)$$

where $\langle . \rangle$ represents the dot product of the estimated illuminant e_e and the ground truth e_t ; and $\|.\|$ stands for the Euclidean norm of a vector. It is worth mentioning that this error measure might not correspond precisely to observers’ preferences [46], however, it is the most commonly used comparative measure in the literature. We also computed the reproduction

angular error [15] in all experiments (due to lack of space these results are not reported here). Readers are encouraged to check the accompanying supplementary materials.

We conducted experiments on three benchmark datasets¹, (i) SFU Lab (321 images) [9], (ii) Colour Checker (568 images) [43], and (iii) Grey Ball (11,346 images) [10]. In Table 1 we have reported the best median and trimean angular errors for each of the considered methods (these metrics were proposed by [26] and [24] respectively to evaluate colour constancy algorithms since they are robust to outliers). Mean angular errors are reported in the supplementary materials.

Method	SFU Lab [9]		Colour Checker [43]		Grey Ball [10]	
	Median	Trimean	Median	Trimean	Median	Trimean
White-Patch [33]	6.44	7.51	5.68	6.35	6.00	6.50
Grey-Edge 1 st -order [45]	3.52	4.02	3.72	4.76	5.01	5.80
Grey-Edge 2 nd -order [45]	3.22	3.65	4.27	5.19	5.72	6.39
Double-Opponency [27]	2.37	3.27	3.46	4.38	4.62	5.28
CVP White-Patch	2.99	3.42	2.97	3.45	5.02	5.15
CVP Grey-Edge 1st-order	3.29	3.72	2.48	2.79	4.70	5.17
CVP Grey-Edge 2nd-order	2.85	3.13	2.59	2.93	4.65	5.05
CVP Double-Opponency	2.02	2.56	2.39	2.84	4.00	4.24

Table 1: Recovery angular errors of four colour constancy methods under max- and contrast-variant-pooling (CVP) on three benchmark datasets. Lower figures indicate better performance.

Figure 2 illustrates three exemplary results obtained by the proposed contrast-variant-pooling (CVP) operator for two colour constancy models: White-Patch and the first-order Grey-Edge. Qualitatively, we can observe that CVP does a better job than max-pooling in estimating the scene illuminant. This is also confirmed quantitatively for the angular errors, shown at the right bottom side of each computed output.

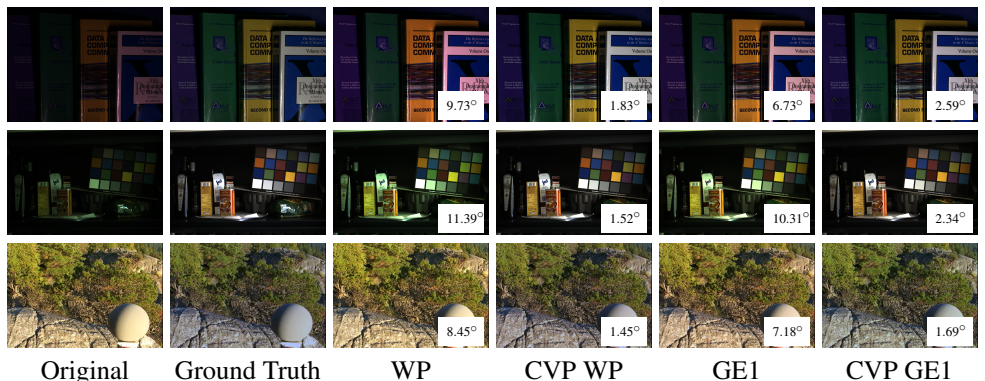


Figure 2: Qualitative results of White-Patch (WP) and the first-order Grey-Edge (GE1) under max- and contrast-variant-pooling (CVP). Angular errors are indicated on the bottom right corner of each panel. Images are from the SFU Lab, Colour Checker and Grey Ball dataset respectively.

¹All source code and materials are available in the supplementary submission.

3.1 Influence of the Free Parameters

For each free variable of the tested models we compared the performance of max- to contrast-variant-pooling. White-Patch does not have any free variable, therefore it is exempted from this analysis. In Figure 3 we have reported the impact of different σ s (receptive field size) on Double-Opponency algorithm for the best and the worst results obtained by free variable k in each dataset (results for all k s are available in the supplementary material). We can observe that almost in all cases contrast-variant-pooling outperforms max-pooling. The improvement is more tangible for the Colour Checker and Grey Ball datasets and in low σ s.

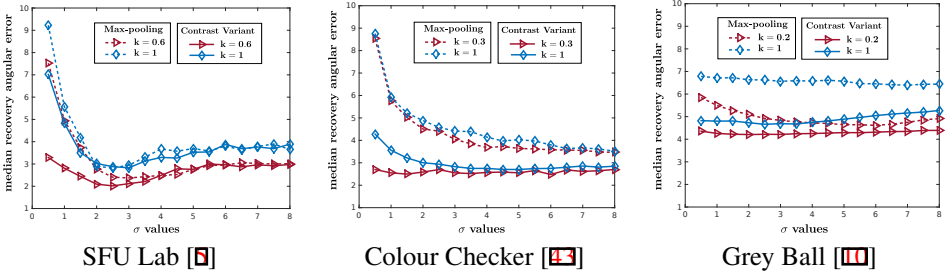


Figure 3: The best and the worst results obtained by max- and contrast-variant-pooling for the free variables of Double-Opponency algorithm (k and σ).

Figure 4 illustrates the impact of different σ s (Gaussian size) on the first- and second-order Grey-Edge algorithm. We can observe similar patterns as with Double-Opponency (contrast-variant-pooling outperforms max-pooling practically in all cases). This improvement is more significant for low σ s, for the Colour Checker dataset and for the second-order derivative. It must be noted that the objective of this article was merely to study the performance of max-pooling and CVP on top of the Grey-Edge algorithm. However, the angular errors of our *CVP Grey-Edge* happen to be in par with the best results reported for Grey-Edge (obtained by using the optimum Minkowski norm for each dataset [45]), with the important caveat that CVP has no extra variables to be tuned, whereas in the Minkowski norm optimisation the value of p must be hand-picked for each dataset.

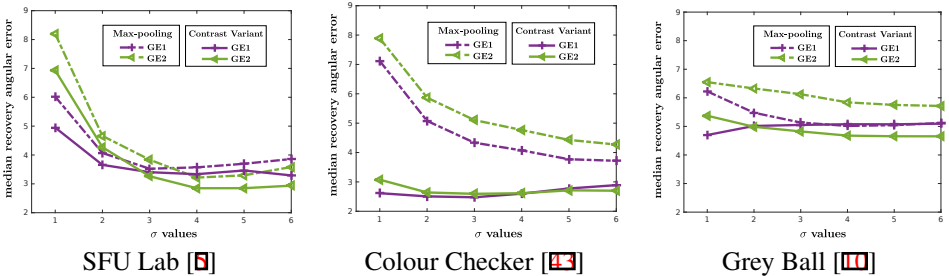


Figure 4: Comparison of max- and contrast-variant-pooling for the free variable σ of Grey-Edge algorithm (both the first- and second-order derivatives).

From Figures 3 and 4 we can observe that the greatest improvement occurs in the Colour Checker dataset. We speculate that one of the reasons for this is the larger range of intensity values in the Colour Checker dataset (16-bit) in comparison to the other two datasets that contain 8-bit images, therefore, an inaccurate max-pooling is more severely penalised.

3.2 Discussion

We would like to emphasise that the objective of this article is not to improve state-of-the-art in colour constancy, but to show that contrast-variant-pooling (CVP) almost always produces improvements over max-pooling. Surprisingly, the results we obtained are even competitive with the state-of-the-art. For instance, in the SFU Lab dataset, the lowest reported angular error is 2.1 (obtained by an Intersection-based Gamut algorithm [4]). This means that our *CVP Double-Opponency* with an angular error of 2.0 outperforms the state-of-the-art in this dataset. In the Colour Checker and Grey Ball datasets there are a few learning-based models (e.g. Exemplar-based method [5]) that obtain lower angular errors in comparison to *CVP Double-Opponency*, nevertheless our results are comparable with theirs.

Physiological evidence besides, the better performance of CVP can be explained intuitively by the fact that max-pooling relies merely on the peak of a function (or a region of interest), whereas in our model, pooling is defined collectively based on a number of elements near the maximum. Consequently, those peaks that are outliers and likely caused by noise get normalised by other pooled elements. The rationale within our model is to pool a larger percentage at low contrast since in those conditions, peaks are not informative on their own, whereas at high contrast peaks are likely to be more informative and other irrelevant details must be removed (therefore a smaller percentage is pooled).

Although, the importance of choosing an appropriate pooling type has been demonstrated both experimentally [29, 48], and theoretically [7], current standard pooling mechanisms lack the desired generalisation [68]. We believe that contrast-variant-pooling can offer a more dynamic and general solution. In this article, we evaluated CVP on the colour constancy phenomenon as a proof-of-concept, however our formulation of CVP is generic (and based on local contrast) and in principle it can be applied to a wider range of computer vision algorithms, such as deep-learning, where pooling is a decisive factor [42].

In our implementation of CVP, we approximated local contrast through local standard deviation (see Eq. 5). There are at least two factors that require a more profound analysis: (i) incorporating more sophisticated models of contrast perception [25] accounting for extrema in human contrast sensitivity; and (ii) analysing the role of kernel size in the computation of local contrast.

4 Conclusion

In this article, we presented a novel biologically-inspired contrast-variant-pooling (CVP) mechanism grounded in the physiology of the visual cortex. Our main contribution can be summarised as linking the percentage of pooled signal to the local contrast of the stimuli, *i.e.* pooling a larger percentage at low contrast and a smaller percentage at high contrast. Our CVP operator remains always closer to max-pooling rather than to sum-pooling since natural images generally contain more homogeneous areas than abrupt discontinuities.

We tested the efficiency of our CVP model in the context of colour constancy by replacing the max-pooling operator of four algorithms with the proposed pooling. After that, we conducted experiments in three benchmark datasets. Our results show that contrast-variant-pooling outperforms the commonly used max-pooling operator nearly in all cases. This can be explained by the fact that our model allows for more informative peaks to be pooled while suppressing less informative peaks and outliers.

We further argued that the proposed CVP is a generic operator, thus its application can

be extended to a wider range of computer vision algorithms by offering a dynamic (and automatic) framework that is based on the local contrast of an image or a pixel. This opens a multitude of possibilities for future lines of research and it remains an open question whether our model can reproduce its excellent results in other domains as well. Therefore, it certainly is interesting to investigate whether our CVP can improve convolutional neural networks.

Acknowledgements

This work was funded by the Spanish Secretary of Research and Innovation (TIN2013-41751-P and TIN2013-49982-EXP).

References

- [1] Vivek Agarwal, Andrei V Gribok, and Mongi A Abidi. Machine learning approach to color constancy. *Neural Networks*, 20(5):559–563, 2007.
- [2] Arash Akbarinia and C. Alejandro Parraga. Biologically plausible boundary detection. In *Proceedings of the British Machine Vision Conference (BMVC)*, September 2016.
- [3] Alessandra Angelucci and S Shushruth. Beyond the classical receptive field: Surround modulation in primary visual cortex. *The new visual neurosciences*, pages 425–444, 2013.
- [4] Kobus Barnard. Improvements to gamut mapping colour constancy algorithms. In *Computer Vision—ECCV 2000*, pages 390–403. Springer, 2000.
- [5] Kobus Barnard, Lindsay Martin, Brian Funt, and Adam Coath. A data set for color research. *Color Research & Application*, 27(3):147–151, 2002.
- [6] Jonathan T Barron. Convolutional color constancy. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 379–387, 2015.
- [7] Y-Lan Boureau, Jean Ponce, and Yann LeCun. A theoretical analysis of feature pooling in visual recognition. In *Proceedings of the international conference on machine learning (ICML)*, pages 111–118, 2010.
- [8] Gershon Buchsbaum. A spatial processor model for object colour perception. *Journal of the Franklin institute*, 310(1):1–26, 1980.
- [9] Matteo Carandini and David J Heeger. Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, 13(1):51–62, 2012.
- [10] Florian Ciurea and Brian Funt. A large image database for color constancy research. In *Color and Imaging Conference*, volume 2003, pages 160–164, 2003.
- [11] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 886–893, 2005.
- [12] Marc Ebner. *Color constancy*, volume 6. John Wiley & Sons, 2007.

- [13] Christina Enroth-Cugell and John G Robson. The contrast sensitivity of retinal ganglion cells of the cat. *The Journal of physiology*, 187(3):517–552, 1966.
- [14] Graham D Finlayson and Elisabetta Trezzi. Shades of gray and colour constancy. In *Color and Imaging Conference*, volume 2004, pages 37–41, 2004.
- [15] Graham D Finlayson and Roshanak Zakizadeh. Reproduction angular error: An improved performance metric for illuminant estimation. *perception*, 310(1):1–26, 2014.
- [16] Graham D Finlayson, Steven D Hordley, and Mark S Drew. Removing shadows from images. In *Computer Vision–ECCV 2002*, pages 823–836. Springer, 2002.
- [17] David A Forsyth. A novel algorithm for color constancy. *International Journal of Computer Vision*, 5(1):5–35, 1990.
- [18] David H Foster. Color constancy. *Vision research*, 51(7):674–700, 2011.
- [19] Damien Fourure, Rémi Emonet, Elisa Fromont, Damien Muselet, Alain Trémeau, and Christian Wolf. Mixed pooling neural networks for color constancy. In *Image Processing (ICIP), 2016 IEEE International Conference on*, pages 3997–4001. IEEE, 2016.
- [20] Brian Funt and Weihua Xiong. Estimating illumination chromaticity via support vector regression. In *Color and Imaging Conference*, volume 2004, pages 47–52. Society for Imaging Science and Technology, 2004.
- [21] Brian Funt, Kobus Barnard, and Lindsay Martin. Is machine colour constancy good enough? In *Computer Vision–ECCV’98*, pages 445–459. Springer, 1998.
- [22] Shao-Bing Gao, Kai-Fu Yang, Chao-Yi Li, and Yong-Jie Li. Color constancy using double-opponency. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 37(10):1973–1985, 2015.
- [23] Arjan Gijsenij and Theo Gevers. Color constancy using natural image statistics and scene semantics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(4):687–698, 2011.
- [24] Arjan Gijsenij, Theo Gevers, and Marcel P Lucassen. Perceptual analysis of distance measures for color constancy algorithms. *JOSA A*, 26(10):2243–2256, 2009.
- [25] Andrew M Haun and Eli Peli. Perceived contrast in complex images. *Journal of vision*, 13(13):3–3, 2013.
- [26] Steven D Hordley and Graham D Finlayson. Reevaluation of color constancy algorithm performance. *JOSA A*, 23(5):1008–1020, 2006.
- [27] Paul M Hubel. The perception of color at dawn and dusk. *Journal of Imaging Science and Technology*, 44(4):371–375, 2000.
- [28] Laurent Itti and Christof Koch. Computational modelling of visual attention. *Nature reviews neuroscience*, 2(3):194–203, 2001.
- [29] Kevin Jarrett, Koray Kavukcuoglu, Yann LeCun, et al. What is the best multi-stage architecture for object recognition? In *International Conference on Computer Vision (ICCV)*, pages 2146–2153, 2009.

- [30] Hamid Reza Vaezi Joze and Mark S Drew. Exemplar-based color constancy and multiple illumination. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 36(5):860–873, 2014.
- [31] Hamid Reza Vaezi Joze, Mark S Drew, Graham D Finlayson, and Perla Aurora Troncoso Rey. The role of bright pixels in illumination estimation. In *Color and Imaging Conference*, volume 2012, pages 41–46, 2012.
- [32] Ilan Lampl, David Ferster, Tomaso Poggio, and Maximilian Riesenhuber. Intracellular measurements of spatial integration and the max operation in complex cells of the cat primary visual cortex. *Journal of neurophysiology*, 92(5):2704–2713, 2004.
- [33] Edwin H Land. *The Retinex Theory of Color Vision*. Scientific American offprints. W.H. Freeman Company.
- [34] Edwin H Land. An alternative technique for the computation of the designator in the retinex theory of color vision. *Proceedings of the National Academy of Sciences*, 83(10):3078–3080, 1986.
- [35] Yann LeCun, Bernhard E Boser, John S Denker, Donnie Henderson, Richard E Howard, Wayne E Hubbard, and Lawrence D Jackel. Handwritten digit recognition with a back-propagation network. In *Advances in neural information processing systems*, pages 396–404, 1990.
- [36] David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.
- [37] David Marr and Ellen Hildreth. Theory of edge detection. *Proceedings of the Royal Society of London B: Biological Sciences*, 207(1167):187–217, 1980.
- [38] Naila Murray and Florent Perronnin. Generalized max pooling. In *Computer Vision and Pattern Recognition (CVPR)*, pages 2473–2480, 2014.
- [39] Bruno A Olshausen et al. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381(6583):607–609, 1996.
- [40] C. Alejandro Parraga and Arash Akbarinia. Colour constancy as a product of dynamic centre-surround adaptation. *Journal of Vision*, 16(12):214–214, 2016.
- [41] Erik Reinhard, Michael Stark, Peter Shirley, and James Ferwerda. Photographic tone reproduction for digital images. In *ACM Transactions on Graphics (TOG)*, volume 21, pages 267–276. ACM, 2002.
- [42] Dominik Scherer, Andreas Müller, and Sven Behnke. Evaluation of pooling operations in convolutional architectures for object recognition. *Artificial Neural Networks–ICANN 2010*, pages 92–101, 2010.
- [43] Lilong Shi and Brian Funt. Re-processed version of the gehler color constancy dataset of 568 images. <http://www.cs.sfu.ca/~colour/data/>.
- [44] S Shushruth, Jennifer M Ichida, Jonathan B Levitt, and Alessandra Angelucci. Comparison of spatial summation properties of neurons in macaque v1 and v2. *Journal of neurophysiology*, 102(4):2069–2083, 2009.

-
- [45] Joost Van De Weijer, Theo Gevers, and Arjan Gijsenij. Edge-based color constancy. *IEEE Transactions on image processing*, 16(9):2207–2214, 2007.
- [46] Javier Vazquez-Corral, C. Alejandro Parraga, Ramon Baldrich, and Maria Vanrell. Color constancy algorithms: Psychophysical evaluation on a new dataset. *Journal of Imaging Science and Technology*, 53(3):31105–1, 2009.
- [47] Hugh R Wilson and Frances Wilkinson. Configural pooling in the ventral pathway. *The new visual neurosciences*, pages 617–626, 2014.
- [48] Jianchao Yang, Kai Yu, Yihong Gong, and Thomas Huang. Linear spatial pyramid matching using sparse coding for image classification. In *Computer Vision and Pattern Recognition (CVPR)*, pages 1794–1801, 2009.