On Comparing Colour Spaces From a Performance Perspective: Application to Automated Classification of Polished Natural Stones

Francesco Bianconi[1](✉), Raquel Bello[2], Antonio Fernández[2], and Elena González[2]

[1] Department of Engineering, Università degli Studi di Perugia, Via G. Duranti, 93, 06125 Perugia, Italy
bianco@ieee.org

[2] School of Industrial Engineering, Universidade de Vigo, Campus Universitario, Rúa Maxwell s/n, 36310 Vigo, Spain
bellocerezo@gmail.com, {antfdez,elena}@uvigo.es

Abstract. In this paper we investigate the problem of choosing the adequate colour representation for automated surface grading. Specifically, we discuss the pros and cons of different colour spaces, point out some common misconceptions about their use, and propose a number of 'best practices' for colour conversion. To put the discussion into practice we generated a new dataset of 25 classes of natural stone products which we used to systematically compare and evaluate the performance of seven device-dependent and three device-independent colour spaces through two classification strategies. With the nearest neighbour classifiers no significant difference emerged among the colour spaces considered, whereas with the linear classifier it was found that device-independent Lab and Luv spaces performed significantly better than the others.

Keywords: Soft colour descriptors · Colour spaces · Visual appearance · Natural stones

1 Introduction

The evaluation and comparison of the visual appearance of real-world materials plays a pivotal role in many applications such as defect detection, surface grading, creation of batches of similar appearance and image-based material recognition. For a wide range of products including natural stones, ceramic tiles, parquet, leather and the like such tasks are of primary importance. Among them natural stones play an important part with a net worldwide production of finished products in excess of 76.000 tons/year [12]. Since natural stones are typically used for cladding and tiling areas which are supposed to be uniform in appearance, one of the major problems is to guarantee the uniformity of the visual aspect within

V. Murino et al. (Eds.): ICIAP 2015 Workshops, LNCS 9281, pp. 71–78, 2015.
DOI: 10.1007/978-3-319-23222-5_9

the same lot. In the stone industry this task is usually carried out manually by skilled and experienced workers. Consequently, the process is intrinsically subjective, and the results may suffer from significant intra- and inter-observer variability. The use of automated computer vision systems either in replacement of or as an aid for human operators could in principle offer higher quality standards, better reproducibility, and more reliable product records. In this context the objective of this work is to propose a theoretically well-founded procedure to compare colour spaces and determine the most appropriate for this task.

In the remainder of the paper, after a brief review of the relevant literature (Sec. 2), we describe the materials (Sec. 3) and methods (Sec. 4) used in this study. We present the experimental set-up and the results in Sec. 5 and conclude the paper with final considerations and ideas for future studies in Sec. 6.

2 Related Research

Performance analysis of colour spaces in computer vision has generated considerable research interest since early on. One could reasonably expect that device-dependent, uniform colour spaces should be superior to the others, but the results available in the literature are rather inconclusive in this regard. On the one side Adel *et al.* [1] compared five colour spaces for wood defect detection and found that device-independent spaces (CIE Luv and CIE Lab) outperformed device-dependent spaces (Ohta's I1I2I3, HSI and RGB). Likewise, Paschos [15] determined that in most cases perceptually uniform/approximately uniform colour spaces (CIE Lab and HSV) outperformed RGB. Similar findings were later on reported also by Rajadell and García-Sevilla [18]. On the other hand, Drimbarean and Whelan [7] found that none of the RGB, HSI, CIE XYZ, CIE Lab and YIQ spaces was significantly superior in classifying colour texture, whereas Qazi *et al.* [17] observed that for pure colour features RGB and IHLS colour spaces gave slightly better results than CIE Lab, but the trend was just the opposite for colour textures. More recently Bianconi *et al.* [6] compared the performance of colour descriptors and spaces for automated sorting of parquet slabs and noticed very little difference among the performance of RGB, HSV and CIE Lab.

The lack of agreement among the results available in the literature also parallels with the fact that in most cases data and/or code are not available to the public, making it difficult (if not impossible) to reproduce the experiments and carry on further comparisons and evaluations. It is also important to point out that, regrettably, in most studies the experimental configuration was far from optimal: it is not uncommon that conversion from device-dependent to device-independent colour spaces be performed in a sloppy way using pre-defined formulas regardless of the acquisition system and the lighting conditions, an approach which is likely to produce biased results [16, Chap. 7.3].

3 Materials

One hundred tiles representing 25 classes (four tiles for each class) of polished natural stone products (marble and granite) were kindly provided to the authors

Fig. 1. The 25 classes of MondialMarmi 2.0. The images are column-wise ordered by decreasing average luminance (brightest → top-left, darkest → bottom-right).

by Mondial Marmi S.r.l., a stone manufacturing company based near Perugia, Italy. The dataset, which we refer to as MondialMarmi 2.0 (Fig. 1), extends the previous version 1.1 (see also [4]) by adding 13 more classes; it also provides a higher image resolution and avoids the JPEG compression artifacts that affected version 1.1 [9].

Image acquisition was based on the system described in [5], which is composed of a dome-shaped illuminator, a slot to accommodate the object to be acquired and a camera mounted on a rotatable device placed at the top of the dome. The acquisition process was carried out as follows: each tile was placed in the image acquisition slot and 10 images were taken at different rotation angles from $0°$ to $90°$ by steps of $10°$. Diffuse illumination with an average level of $\approx$600lx measured at the top surface of the tile was guaranteed by white leds with colour temperature in the range 5700K-6000K. The illumination conditions were kept constant throughout the whole acquisition process. The images were captured through a CMOS camera (Edmund Optics EO-5012C LE) equipped with a 6mm fixed-focal lens (Pentax H614-MQ-KP) at the natural resolution of the sensor (2560px × 1920px), and were finally cropped to a central part of size 1500px × 1500px corresponding to an area of ≈21cm × 21cm. Image encoding was linear with no gamma correction. An X-Rite Digital SG colour checker (140 colour targets) was also put in place with the aim of enabling the colour calibration procedure as described in Sec. 4.5. The RGB values of the 140 colour targets were acquired using the same settings adopted for the tiles; the corresponding device-independent CIE XYZ coordinates (under illuminant D65) were measured through a Minolta CR-200 Chroma Meter. The whole dataset is freely available to the public for future evaluations and comparisons [11].

4 Methods

4.1 Colour Descriptors

Soft colour descriptors as defined by López *et al.* [10] are sets of global statistics describing colour and texture properties of the material. Soft colour descriptors are particularly appealing since they are easy to implement, computationally fast, and produce low-dimensional feature vectors. Based on the good results provided in previous studies [6,10] herein we considered the following five descriptors: mean, mean + standard deviation, mean + moments from 2^{nd} to 5^{th}, quartiles and quintiles of each colour channel.

4.2 Colour Spaces

Colour spaces may be either device-dependent or device-independent. The data encoded by device-dependent spaces such as RGB and HSV are device-specific, since they depend on the characteristics of the imaging system. Among them RGB plays a pivotal role, for most imaging devices provide their output in this space. By contrast, data encoded by device independent spaces such as CIE XYZ (also referred to as *colorimetric* spaces) represent the response of an ideal standard observer, hence they do not depend on the acquisition device. The kernel of the device-independent spaces is CIE XYZ: any other device-independent space is a space which can be converted to or obtained from CIE XYZ with no additional information. Perceptual uniformity is another important feature of colour spaces: in a perceptually uniform space the difference between two colour stimuli as perceived by the human eye is expected to be proportional to the Euclidean distance between the corresponding points in the space. In this sense RGB is not perceptually uniform; CIE Lab and CIE Luv are locally uniform, whereas HSV is approximately uniform.

4.3 Device-Dependent Spaces

In the experiments we considered the following device-dependent spaces: RGB, HSV, YUV, YIQ, YCbCr, Ohta's I1I2I3 and the opponent colour space RG-YeB-WhBl. Conversion from the original RGB data to the other device-dependent spaces was carried out using the standard formulas for which references are given hereafter.

HSV is a colour space in which colour information is decomposed into three intuitive and perceptually relevant coordinates: hue (H), saturation (S) and intensity (V). Unfortunately, conversion from RGB to HSV requires a non-linear transform [8, Sec. 7.2] which generates non-removable singularities. YIQ, YUV and YCbCr are obtained from RGB through simple linear transforms of the colour channels [14, Secs. 4.3.1–4.3.3]. Their common feature is to separate luminance (Y) from chrominance; they are also known as *television* colour spaces, since they have been used for encoding and broadcasting television signals. Ohta's I1I2I3 colour space aims at separating colour information into three approximately orthogonal (decorrelated) components [13]. Conversion from RGB is carried out through a simple linear transform [14, Sec. 4.5]. Finally, the RG-YeB-WhBl space derives from Hering's theory of opponent colour vision: the conversion formula [14, Sec. 4.4] is again a linear transform of the RGB space.

4.4 Device-Independent Spaces

The following device-independent spaces were included in this study: CIE XYZ, CIE Lab and CIE Luv. The original RGB data were first converted to CIE XYZ trough colour calibration (see Sec. 4.5), then from CIE XYZ to CIE Lab and CIE Luv through the standard non-linear transforms [8, Secs. 5.2–5.3].

4.5 Colour Calibration and Gamut Mapping

Conversion from RGB to CIE XYZ was performed under the assumption of a linear transformation model between device-dependent and device-independent data. Thought other approaches are also possible [2], the linear model was chosen on the basis on the ease of implementation, low computational demand and the good results provided in previous work [3]. In formulas we have:

$$\left\{\begin{matrix} X \\ Y \\ Z \end{matrix}\right\} = \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \left\{\begin{matrix} T_1 \\ T_2 \\ T_3 \end{matrix}\right\} \tag{1}$$

The conversion requires estimating the 12 parameters of the model (i.e.: the $M_{i,j}$ and the T_i; where $i, j \in \{1, \ldots, 3\}$), operation which was carried out by solving the resulting overdetermined linear system (12 unknowns and $3 \times 140 = 420$ equations) through least-square estimation based on the 140 colour targets of the calibration rig (Sec. 3). Gamut mapping was obtained by projecting the eight vertices of the RGB cube into each of the other colour spaces considered in the experiments. The transformed gamuts were used to normalise the data in each of the destination colour spaces. For comparison purposes we also performed the transformation from RGB to XYZ using the approximate method based on the general-purpose approximated formula [10] instead of colour calibration. In the remainder we indicate as Lab_f and Luv_f the corresponding Lab and Luv coordinates obtained this way.

Table 1. Overall accuracy (1-NN and linear classifier).

Col. space	Mean	Mean+Std	Mean+Mom.	Quartiles	Quintiles
	Classifier: *1-NN*				
CIE XYZ	82.3	90.0	85.7	90.9	91.1
CIE Lab	89.8	93.5	91.8	93.5	93.6
CIE Luv	**90.4**	94.0	**92.4**	93.9	94.0
CIE Lab_f	88.0	93.8	91.0	93.0	93.3
CIE Luv_f	89.1	94.1	91.8	93.8	94.1
RGB	89.1	**94.9**	91.2	94.4	**94.7**
HSV	90.0	94.1	90.0	91.7	91.3
YUV	89.1	94.1	91.6	93.5	93.7
YIQ	88.5	93.8	91.2	93.1	93.4
YCbCr	89.1	94.0	91.6	93.5	93.7
I1I2I3	89.8	**94.9**	92.3	**94.2**	94.5
RG-YeB-WhBl	89.8	**94.9**	89.8	94.1	94.4

Col. space	Mean	Mean+Std	Mean+Mom.	Quartiles	Quintiles
	Classifier: *Linear*				
CIE XYZ	82.3	92.6	91.3	88.6	89.0
CIE Lab	**94.8**	**98.6**	**98.3**	**95.7**	**96.0**
CIE Luv	**94.8**	98.1	97.6	**95.7**	95.8
CIE Lab_f	92.4	95.7	96.5	93.8	94.4
CIE Luv_f	88.9	92.4	89.7	94.7	95.0
RGB	85.2	94.9	93.7	90.4	90.8
HSV	91.3	97.2	93.7	91.7	91.6
YUV	93.0	95.6	96.4	94.5	94.3
YIQ	93.3	95.6	96.5	94.5	95.0
YCbCr	93.0	95.6	96.3	94.2	94.4
I1I2I3	93.1	95.9	96.8	94.5	95.2
RG-YeB-WhBl	93.4	96.1	95.5	94.8	95.2

5 Experiments and Results

To evaluate the performance of the colour spaces presented in Sec. 4.2 we run a supervised classification experiment using two different classification strategies: nearest-neighbour (1-NN) rule with Euclidean distance and linear classifier with diagonal covariance matrix. The experimental basis included the 25 classes described in Sec. 3 where each image was subdivided into four non-overlapping sub-images giving 16 samples for each class. Accuracy estimation was based on 100-fold validation with stratified sampling and a train ratio of 1/4. To assess the robustness against rotation of the descriptors the train images were always picked from the 0°-group; the test images from each of the θ°-group (including the 0°-group). For each rotation angle the accuracy was the percentage of test images correctly classified; finally, the overall accuracy values were averaged over the rotation angles of each dataset to give a global accuracy measure (Tab. 1). To estimate the effect of the colour space we computed, for each of them, the average accuracy over all the colour descriptors considered in the experiment and estimated the 95% confidence interval assuming a normal distribution (Tab. 2).

Table 2. Average overall accuracy by colour space (confidence intervals for the mean).

Colour space	C.I. for the mean
Classifier: *1-NN*	
CIE XYZ	87.8 - 88.1
CIE Lab	92.4 - 92.5
CIE Luv	92.9 - 93.0
CIE Lab_f	91.7 - 91.9
CIE Luv_f	92.5 - 92.7
RGB	92.8 - 93.0
HSV	91.3 - 91.5
YUV	92.3 - 92.5
YIQ	91.9 - 92.1
YCbCr	92.3 - 92.5
I1I2I3	93.0 - 93.2
RG-YeB-WhBl	92.5 - 92.7

Colour space	C.I. for the mean
Classifier: *Linear*	
CIE XYZ	88.6 - 88.9
CIE Lab	**96.6 - 96.7**
CIE Luv	**96.4 - 96.5**
CIE Lab_f	94.5 - 94.6
CIE Luv_f	92.0 - 92.2
RGB	90.8 - 91.1
HSV	93.0 - 93.2
YUV	94.7 - 94.8
YIQ	94.9 - 95.0
YCbCr	94.6 - 94.8
I1I2I3	95.0 - 95.2
RG-YeB-WhBl	95.0 - 95.1

6 Conclusions and Future Work

In this paper we have presented a rigorous procedure for comparing colour spaces from a performance perspective, and have applied it to the problem of automated grading polished natural stones. The results show that under controlled illumination conditions simple and computationally cheap colour descriptors are quite effective for the task, with overall accuracy beyond 97%. With the nearest neighbour classifier no colour space emerged as significantly superior to the others. This finding parallels with the results reported by Drimbarean and Whelan [7] and by Bianconi *et al.* [5]. With the linear classifier, however, the perceptually-uniform spaces CIE Lab and CIE Luv outperformed the others. With both classifiers the conversion from device-dependent to device-independent spaces gave better results when performed through colour calibration than through the approximated formula. In future it would be interesting to validate these results in a larger cohort of materials. Another important question for future research is the extension of this study to the case of variable illumination conditions.

Acknowledgments. This work was partially supported by the European Commission under project LIFE12 ENV/IT/000411 and by the Spanish Government under project AGL2014-56017-R.

References

1. Adel, M., Wolf, D., Vogrig, R., Husson, R.: Evaluation of colour spaces in computer vision. In: Proc. of the International Conference on Systems, Man and Cybernetics, vol. 2, pp. 499–504. Le Touquet, France, October 1993
2. Bianco, S., Gasparini, F., Russo, A., Schettini, R.: A new method for RGB to XYZ transformation based on pattern search optimization. IEEE Transactions on Consumer Electronics **53**(3), 1020–1028 (2007)
3. Bianconi, F., Saetta, S.A., Sacchi, G., Asdrubali, F., Baldinelli, G.: Colour calibration of an artificial vision system for industrial applications: comparison of different polynomial models. In: Rossi, M. (ed.) Colour and Colorimetry Multidisciplinary Contributions. Optics and Photonics Series Notebooks, no. 21, pp. 18–25. Maggioli Editore (2011)

4. Bianconi, F., González, E., Fernández, A., Saetta, S.A.: Automatic classification of granite tiles through colour and texture features. Expert Systems with Applications **39**(12), 11212–11218 (2012)
5. Bianconi, F., González, E., Fernández, A., Saetta, S.A.: Apparato per acquisire una pluralità di immagini di almeno un corpo e relativo metodo (Apparatus to acquire a plurality of superficial images of at least one body and related method), 2015. IT patent no. 0001413266. Filed on July 25, 2012; granted on January 16, 2015
6. Bianconi, F., Fernández, A., González, E., Saetta, S.A.: Performance analysis of colour descriptors for parquet sorting. Expert Systems with Applications **40**(5), 1636–1644 (2013)
7. Drimbarean, A., Whelan, P.F.: Experiments in colour texture analysis. Pattern Recognition Letters **22**(10), 1161–1167 (2001)
8. Kang, H.R.: Computational Color Technology. Spie Press (2006)
9. Kylberg, G., Sintorn, I.-M.: Evaluation of noise robustness for local binary pattern descriptors in texture classification. EURASIP Journal on Image and Video Processing **2013**(17) (2013)
10. López, F., Valiente, J.M., Prats, J.M., Ferrer, A.: Performance evaluation of soft color texture descriptors for surface grading using experimental design and logistic regression. Pattern Recognition **41**(5), 1744–1755 (2008)
11. Marmi, M.: A collection of images of polished natural stones for colour and texture analysis. version 2.0 (2015). http://dismac.dii.unipg.it/mm. (last accessed on May 7, 2015)
12. Montani, C.: XXV World Marble and Stone Report. Aldus Casa di Edizioni, Carrara (2014)
13. Ohta, Y., Kanade, T., Sakai, T.: Color information for region segmentation. Computer Graphics and Image Processing **13**(3), 222–241 (1980)
14. Palus, H.: Representations of colour images in different colour spaces. In: Sangwine, S.J., Horne, R.E.N. (eds.) The Colour Image Processing Handbook, pp. 67–90. Springer (1998)
15. Paschos, G.: Perceptually uniform color spaces for color texture analysis: An empirical evaluation. IEEE transactions on Image Processing **10**(6), 932–937 (2001)
16. Petrou, M., Petrou, C.: Image Processing: The Fundamentals. John Wiley & Sons Ltd (2010)
17. Qazi, I.U.H., Alata, O., Burie, J.C., Moussa, A., Fernández Maloigne, C.: Choice of a pertinent color space for color texture characterization using parametric spectral analysis. Pattern Recognition **44**(1), 16–31 (2011)
18. Rajadell, O., García-Sevilla, P.: Influence of color spaces over texture characterization. In: Medina Barrera, M.G., Ramírez Cruz, J.F., Sossa Azuela, J.H. (eds.) Advances in Intelligent and Information Technologies. Research in Computing Science, vol. 38, pp. 273–281. Instituto Politécnico Nacional, Centro de Investigación en Computación, México (2008)