Automatic colour segmentation and colour palette identification of complex images

Aiman Raza[†], Sophie Jost, Marie Dubail[*] and Dominique Dumortier

Laboratoire Génie Civil et Bâtiment, Ecole Nationale des Travaux Publics de l'Etat, Université de Lyon, 69120 Vaulx en Vélin, France
[*]Essilor International, Paris, France
[†]Email: aiman.raza@entpe.fr

A fast and simple colour classification algorithm was developed based on k-means++ algorithm for quick colour analysis of complex images. The algorithm processes Gaussian blurred CIELAB version of the original image to segment the scene in clusters of six colours with the highest representation in the image. It identifies the colours with the help of the ISCC-NBS colour terminology and locates them on the scene. The accuracy of our algorithm was evaluated with a psycho-visual experiment. The results show a dominant colour classification coherent with human visual perception. The algorithm is efficient for quick analysis of complex colour images to retrieve automatically the colour composition. It can be applied directly in various domains for example: lamp testing for colour distortion, image retrieval, artwork analysis etc.

Received 22 January 2021; accepted 30 January 2021
Published online: 16 February 2021

Introduction

Context

Colour perception is a major component of human visual system and an important visual discrimination cue. Colour analysis has been mainly developed for single or uniform patches. It is used in the industry (textile – graphic arts…) to evaluate colour differences or colour shifts. Colour analysis applied on complex images allows to characterise the coloured appearance of the objects in the scene under the lit environment of the scene (natural or artificial). The assessment of the overall colour sensation of an object in a given environment can be useful in various domains, for example: to assess the reproduction of a colour on a complex object or image; to analyse an object under different lightings to assess metamerism, or under several directions of lighting to predict the impact of a specific coloured layer. A prior knowledge of the colour composition of a scene is thus important for any industry relying on colour reproduction quality, like the printing industry, the display industry or even the glazing industry.

The colour composition of a natural/artificial scene tries to describe the scene in terms of the significant colours that guide the visual perception of the observers. The colour that will dominate the visual cue of the image is the dominant colour, and along with the other significant colours, it forms the colour composition of the image. With increasingly sophisticated high-resolution cameras coupled with High Dynamic Range (HDR) screens and high colour bit depths, the amount of visible colours on displays is also increasing. This has made the autonomous identification of the colour composition of an image a challenging task. This rise in image colour quality standards also provides an impetus in developing newer methodologies for colour descriptors that are efficient, fast and accurate.

Motivation

In lighting engineering, colour content descriptors are an important tool to analyse objectively the final rendered colour of a scene under a new type of lighting system. The direction of hue change can be easily identified with the change in dominant colours of the scene. Similarly, for daylight, the final rendered colour of the outside view seen through the glazing should be known before window installation. For this, the spectral transmission of the glazing can be multiplied by the hyperspectral image of any scene seen outside to visualise how the scenes will be rendered through the glazing. A hyperspectral image captures the spectra for each pixel instead of the RGB values. Thus, with a colour descriptor, we can easily identify the colorimetric distortion of a scene with and without the glazing. The same technique could be applied on car windows, especially those tinted for sunlight protection.

In computer vision, Content Based Image Retrieval (CBIR) is the technique to browse, organise and/or search images within a large database based on their content or text [1]. Text based methods rely largely on human subjective perception and are time consuming [2]. Other methods use the image content (colour and shape) and colour composition to speed up image retrieval [3-8]. Colour content analysis becomes particularly relevant for images with a multitude of shapes where shape based analysis is difficult [9].

In cultural heritage management, for art-work restoration and conservation non-invasive methods are preferred, since works of art are delicate and cannot be submitted to invasive techniques for a colour control [10]. Whereas a non-invasive colour control through a photographic colour description at a precise location with a colour calibrated camera can make the process relatively simple and fast. Colour description is also important for studying ancient marble and stone structures [11], and an efficient colour descriptor can effectively render the process much reliable by reducing the human subjectivity of colour perception.

In photography and cinema, specific colour rendering may increase the interest in a scene but special care needs to be taken to avoid colour shifts, or disturbing the colour harmony of the scene [12]. This information can be known before hand with the help of simple colour descriptors.

Current techniques

Numerous methods exist for determining the colour composition of images, the most common being a colour histogram [9], [13-17]. A histogram is a mathematical function that counts the number of observations that fall into discrete categories [18]. An image colour histogram identifies the pixels in the image in terms of a probability density functions of the pixel information, which are the colour coordinates in a colour space like RGB, HSV, XYZ etc. [19].

The major issue with a histogram is the lack of complementary spatial information. In Figure 1, three different colour spaces (RGB, HSV and XYZ) are used to plot the colour histograms of a desert image. The histograms give an idea of the pixel colours distribution but no clues regarding their spatial

distribution, i.e. orange in the lower half and blue in the upper half. The lack of spatial information associated to pertinent colour bins makes histograms less usable for complex images [20].

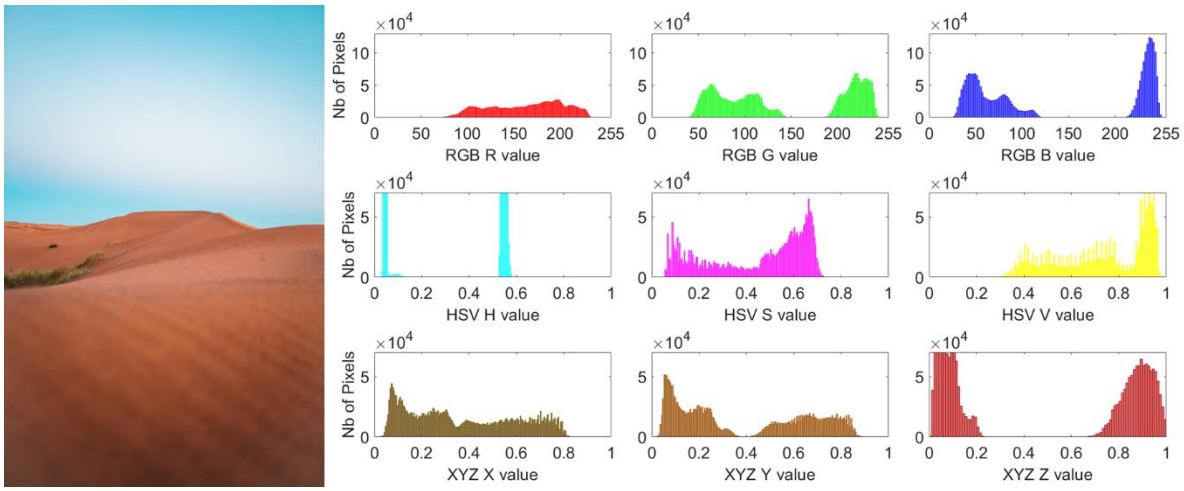

Figure 1: Image (open access) and its RGB, HSV and XYZ colour histograms.

Apart from colour histograms, many other approaches exist for colour based image description and dominant colour retrieval. A brief description of a few prominent techniques is described thereafter.

Image segmentation techniques have a large share in finding the image colour composition. For example, region growing, in which initial 'seeds' (pixels) become a region by adding similar neighbouring pixels if they clear the pre-determined threshold [21]. Studies have shown successful implementation of the region growing technique [22-24], though it suffers from a lack of global approach which poses a problem for complex scenes.

Combination of various techniques provides significant improvement in the final results by using techniques that complete each other, for example: image segmentation combined with histograms [25], or image segmentation combined with a 2-stage hierarchical artificial neural network map based on Kohonen Self organizing maps [26]. The major problem with these approaches is their complexity which hinders their adaptation to various domains.

Methodology

This article uses the kmeans++ algorithm for image segmentation to provide a simple but efficient colour descriptor of a complex image [27]. Kmeans++ algorithm determines only the first pixel by random assignment, the rest of the pixels are carefully chosen to maximise the distance between the centroids of each cluster. Even though, this approach takes longer in initialising, the clustering process has been shown to be much faster than the original kmeans clustering, thus globally reducing the time taken to converge the k clusters [28].

The first step of the algorithm is to undo any gamma correction on the input image, thus the sRGB image is converted to a linear version. This linear image is further transformed into the CIELAB colour space. CIELAB colour space was chosen as the base colour space not only because of simple, homogenous, and uniform colour distribution but also because of its effectiveness in describing perceived differences between colours. The LAB colour space requires the knowledge of the white point of the illuminant which creates a problem for images with unknown illuminant. Illuminant estimation methods do exist and provide an approximation of the illuminant white-point from an image. The more

common ones being the White Patch Retinex algorithm and the Gray world algorithm [29-30]. Both algorithms, though quite effective, are prone to large estimation errors [31]. Another method uses the Principal Component Analysis (PCA) to rank the intensity values of the bright and dark pixels in the scene to determine their deviation from the statistical mean of the data. The PCA based illuminant estimation method has been shown to produce significantly better results than many other methodologies [32].

Knowing the illuminant white-point, the linear RGB image is converted into a LAB image which is then treated with a Gaussian blur (sigma, σ= 8). A blurring is essential to our approach to facilitate faster convergence of the k-means++ algorithm. Blurring with a relatively high standard deviation gives less importance to the edges and local differences and bring out the global colour tendencies of the image. To visually interpret the impact of blurring and the eventual reduction in the spatial frequency of the image, a Discrete Fourier transform is applied on an image while progressively increasing the sigma value. Figure 2 shows the effective reduction in the high frequency components of an image as the filter size increases (a large spread of the grey points indicates the presence of high frequency components).

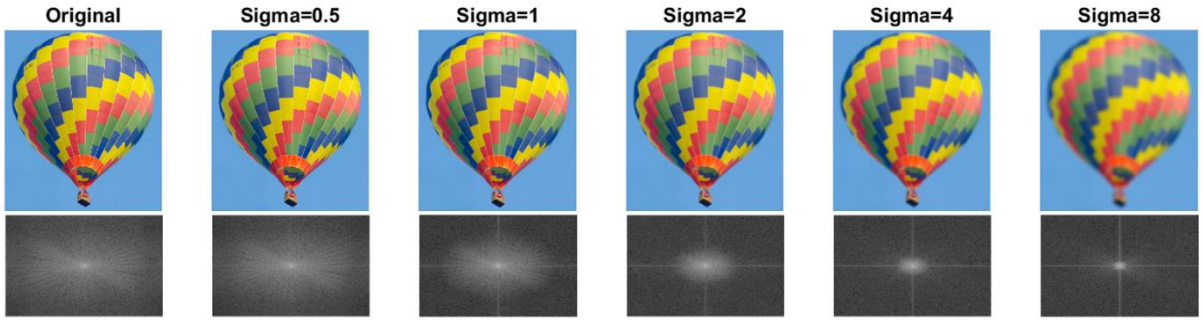

Figure 2: Gaussian blurs on the original image and their corresponding spatial frequency maps.

The image is then ready for the clustering process and kmeans++ with six initial seeds is applied on the blurred LAB image. The resulting six colour clusters are used to identify the spatial location of the untouched RGB colour clusters and to calculate the pixel percentage distribution of each colour cluster. A median sRGB triplet is calculated for each cluster to visualise the colour and then converted to LAB values (without any Gaussian blurring). These LAB triplets are then compared with the ISCC-NBS colour palette to find the closest colour triplet and its name through the CIEDE2000 colour difference formula [33-34]. The ISCC-NBS colour palette uses a three level colour naming system, where 13 basic colour names form the first level, 29 intermediate colour categories form a finer second level and 20 adjectives like vivid, dull, bright, moderate etc. form the finest third level. Each combination of basic colour, intermediate colour category complemented with adjectives for the hue and colour distinguishes between various categories of colourfulness.

Once the colour names are identified, they are checked for redundancy, i.e. colour names repeated more than once. This may happen in images which have an abundance of a particular hue and thus forms less than six actual clusters. The repeating clusters, if any, are merged to form a new cluster and the proportion of colour distribution is updated. Figure 3 summarises the entire pipeline of the process in a flowchart. As output, the program gives: (1) a labelled image as per the colour clusters, (2) a bar plot with the colour name and proportions, (3) a CIE ab chromaticity plot with colour proportions and an MS Excel file with the CIE LAB 1931 observer values, the colour name and its proportions, see Figure 4.

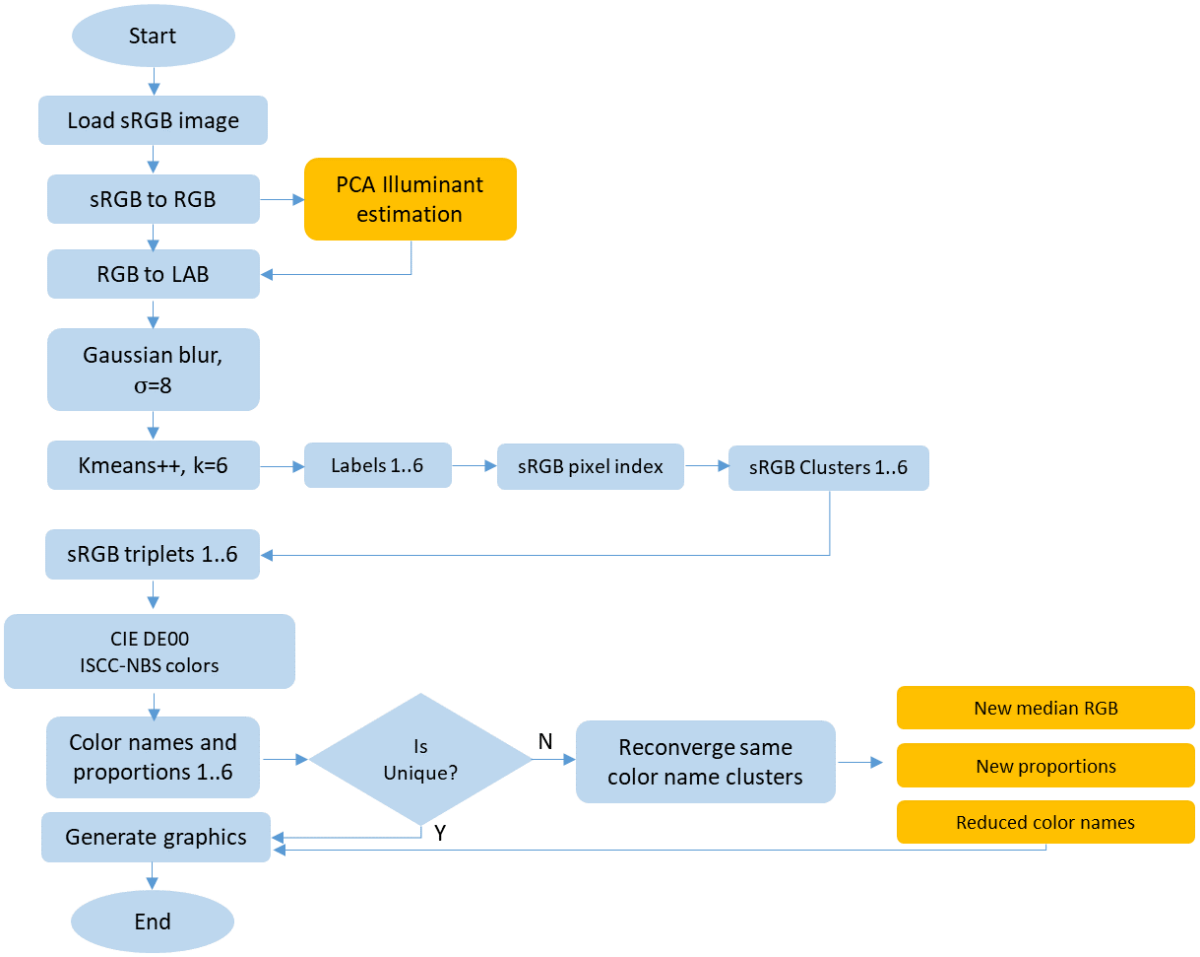

Figure 3: Flowchart of the colour distribution tool.

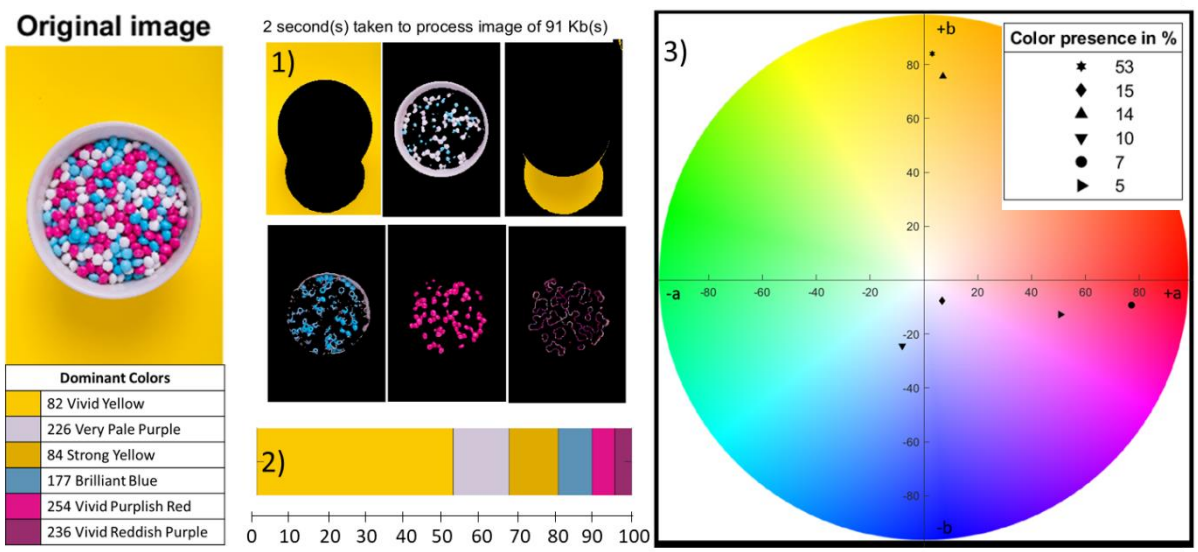

*Figure 4: Image segmentation results and CIE a*b* chromaticity plot of a tested image.*

Validation

The colour descriptor was tested on 50 open source no attribution required photos, and 50 photos from our own database. Figure 5 shows the colour distribution obtained through our method on various indoor/outdoor images.

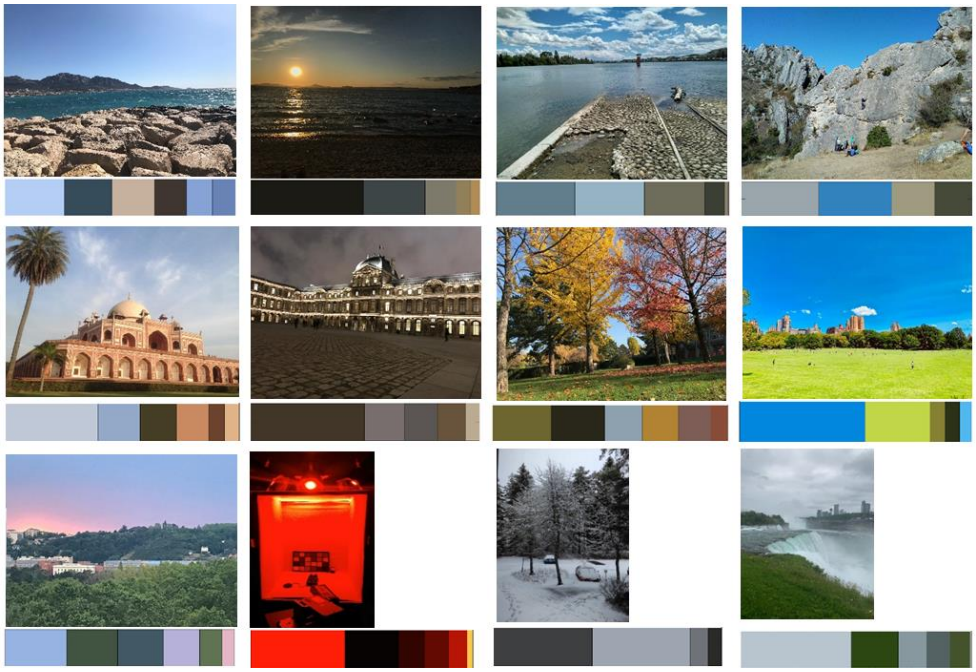

Figure 5: Colour distribution results of various indoor/outdoor images.

The processing time taken by the algorithm for 100 photos had a median of 1,16 seconds. The time taken and the size of the images is shown in Figure 6. The algorithm was developed on a Dell Precision 7520 computer, equipped with an Intel Xeon ER-1535M v6 processor and 32 Gb RAM.

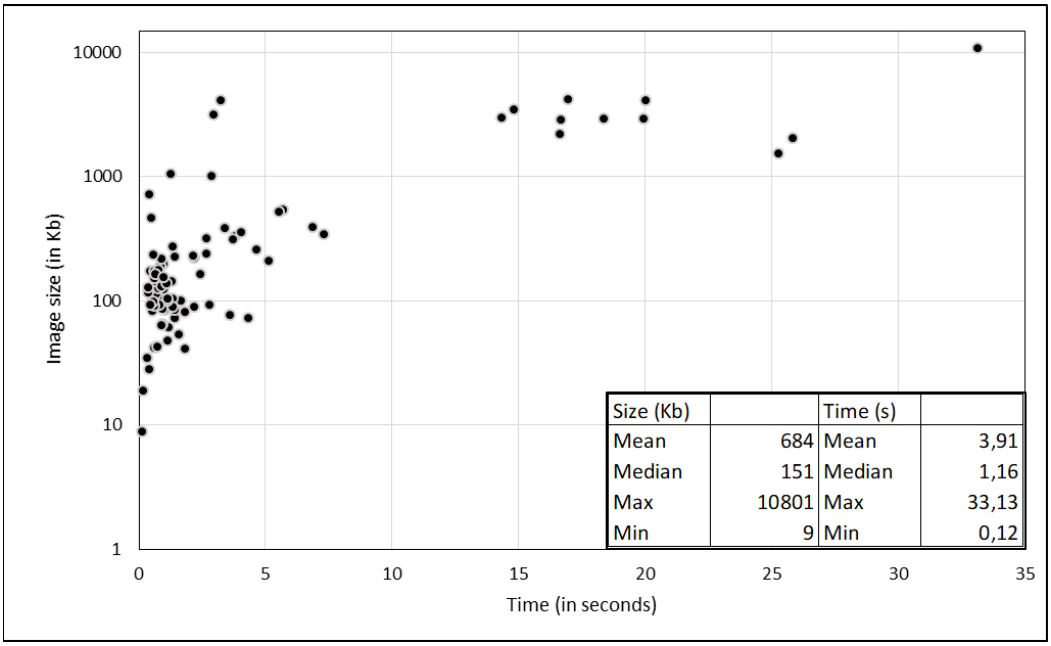

Figure 6: Image size vs processing time for 100 photos.

Experimental procedure

To further validate the results in terms of visual perception, the data from a previously conducted psycho-visual experiment was analysed [35]. Twelve people (age ∈ [23,62]) with normal colour vision, and normal-to-corrected vision performed this experiment in a dark room. Six images (five natural and one urban scene, Figure 7) were shortlisted and presented on a colour calibrated monitor (EIZO ColorEdge CG277) in a controlled Latin square sequence. These images were acquired from a hyperspectral camera and converted to Adobe sRGB for 2D visualisation. The participants were asked to name the dominant colours in each image and list their proportions in order. They were further asked to allocate each colour to a specific colour category: Red, Green, Blue (historical primary colours), Purple, Yellow, Orange (associated secondary colours) and Black, White and Gray (neutral colours). The colours retrieved through our methodology were also similarly categorised.

Figure 7: Images used for the psycho-visual experiment.

Results

The objective and subjective results show coherence in the overall colour distribution, and particularly in the dominant colours for all the six images with a 93.2% correlation. Figure 8 illustrates the similarity/dissimilarity between the objective and the subjective results for each colour category and image. One of the inconsistency in the results is the distribution of white and yellow colours for image A. The algorithm detects no pure white tones and many yellow tones while the psychophysical experiment shows a larger amount of white tones and less amount of yellow. This supposedly erroneous classification can be attributed to the effect of memory colours where the participants recognise the two cars in the centre of the scene to be white. Even though the cars would be characterised as white in a neutral environment, the large amount of sunlight in the scene has added a yellow hue to the cars. Another problem of memory colours is in the image E, where the participants identify the semi-dried grass as pure green while our method detects it as a combination of yellow and green.

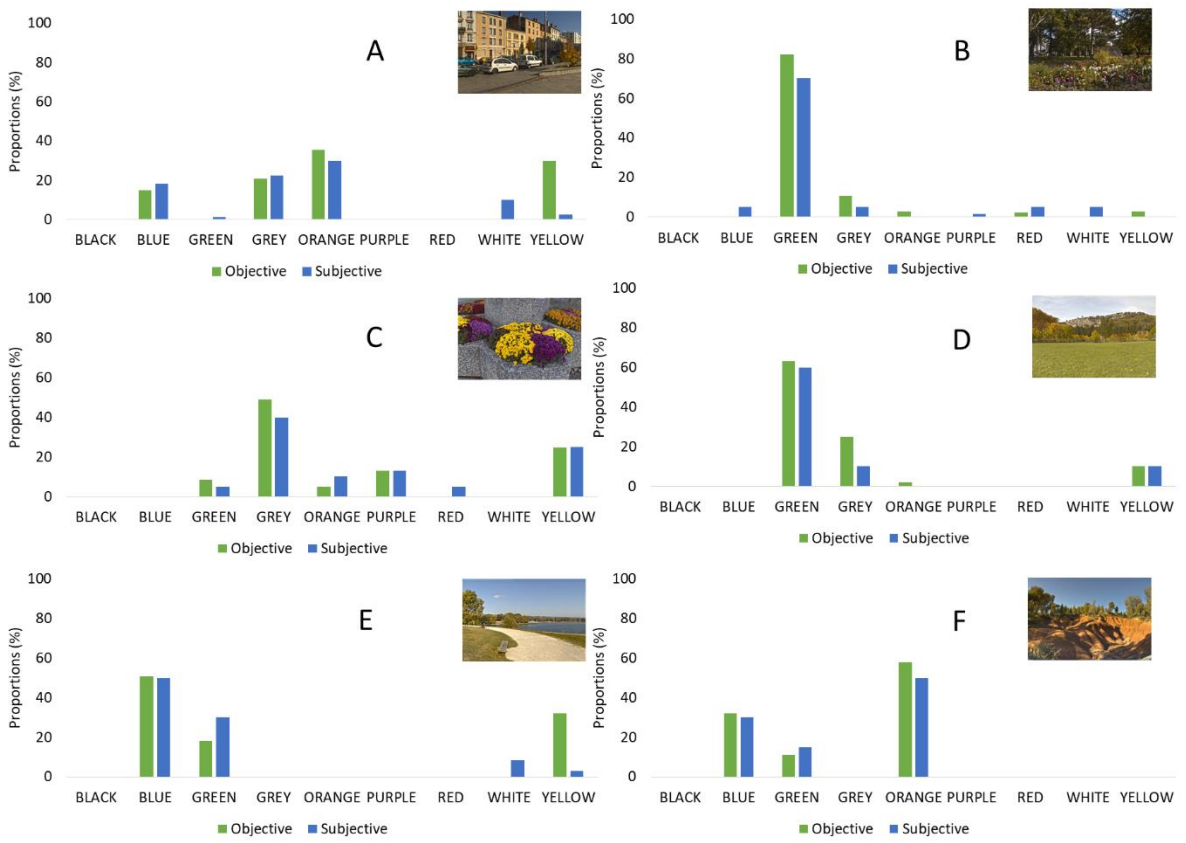

Figure 8: Objective vs subjective results on colour distribution of complex scenes.

Discussion and conclusions

The aim of the developed algorithm is to propose a simplified methodology to automatically determine the colour distribution of images. Our tool achieves a reasonable precision with a method of clustering and associated parameters that directly influence the clustering results. In particular, the choice of CIELAB colour space influences the results positively by determining clusters closer to human visual perception. A Gaussian blur improves the processing time and the accuracy of the clustering process. It removes the highest spatial frequencies and thus removes the details that do not represent the global colour composition of the image. The results from the psycho-visual experiment are in agreement with the results predicted by the algorithm. The k-means++ algorithm guarantees that the results are repeatable and that the image will be segmented in the same manner provided that no parameter is changed.

A concern we have in our algorithm is the inability to determine the number of clusters as per the content of the image. We determined k=6 as an optimum seed value after various pre-tests. Evaluation methods like Silhouette index, Davies-Boudin, Calinski Harabasz etc. exist to check if the number of seeds (k) is appropriate for the data sample or not [36-37]. But the validation process is long for images, it takes more than 10 minutes for a Calinski Harsbasz Cluster Evaluation (system information). The possible k values tested $\in$ [3,8], and the optimum k from Calinski Harsbasz Cluster Evaluation method was found to be 6, thus our empirically determined value was validated a posteriori. Silhouette Index failed to converge even after 10 minutes of evaluation and 100 iterations, and thus was ruled out as well.

The algorithm provides a good solution for dominant colour retrieval and image segmentation at the same time. Nevertheless, there are still about 5-10% of images that had different coloured objects

clustered together. This happens when the image has a lot of overlapping objects of different colour. This leads to inaccurate segmentation, even if the sRGB triplets correspond to the dominant visual perception. For example, in Figure 9, there is a coherence in the six dominant colours and the visual perception but it would have been preferred to have green and yellow in two separate clusters. With higher values of k (as high as 10), similar blue and red coloured objects, are segmented into more than one cluster, producing even more incoherent results.

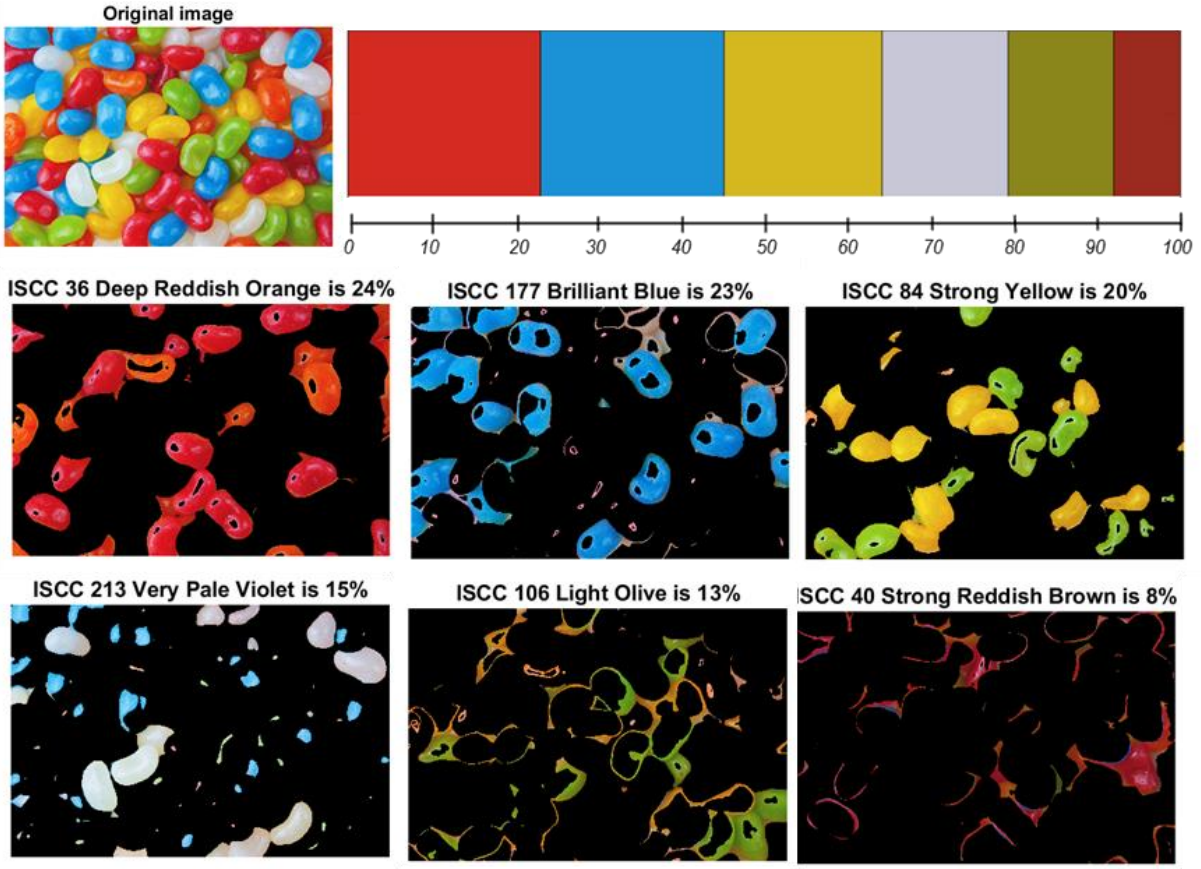

Figure 9: Example of inaccurate clustering but accurate colour distribution.

Nevertheless, the main objective, which is to describe automatically the colour composition of an image has been achieved despite the above problems. Overall, the algorithm works efficiently and rapidly to analyse colours for any image, with or without complementary metadata and can be applied across various domains.

References

1. Datta R, Joshi D, Li J and Wang JZ (2008), Image retrieval: Ideas, influences, and trends of the new age, ACM Computing Surveys. **40** (2), 1-60.
2. Jeyakumar V and Kanagaraj B (2019), A medical image retrieval system in PACS environment for clinical decision making, in *Intelligent Data Analysis for Biomedical Applications: Challenges and Solutions*, Hemanth J, Gupta D and Balas V (eds.), 121-146, Elsevier.

3. Belongie S, Carson C, Greenspan C and Malik J (1998), Color- and texture-based image segmentation using EM and its application to content-based image retrieval, *Proceedings of the 6th International Conference on Computer Vision (IEEE Cat. No.98CH36271)*, 675-682, Bombay (India).
4. Gevers T and Smeulders AWM (2000), PicToSeek: combining color and shape invariant features for image retrieval, *IEEE Transactions on Image Processing*, **9** (1), 102-119.
5. Yu H, Li M, Zhang H-J and Feng J (2002), Color texture moments for content-based image retrieval, *Proceedings of the International Conference on Image Processing*, **3**, 929-932, Rochester (USA).
6. Jain AK and Vailaya A (1996), Image retrieval using color and shape, *Pattern Recognition*, **29** (8), 1233-1244.
7. Mehtre BM, Kankanhalli MS, Desai Narasimhalu A and Chang Man G (1995), Color matching for image retrieval, *Pattern Recognition Letters*, **16** (3), 325-331.
8. Smith JR and Chang S-F (1996), Tools and techniques for color image retrieval, *Proceedings of the SPIE 2670: Storage and Retrieval for Image and Video Databases IV*, 426-437.
9. Lu H, Ooi B-C and Tan K-L (1994), Efficient image retrieval by color contents, in *Applications of Databases*, Litwin W and Risch T (eds.), 95-108, Springer-Verlag Berlin Heidelberg.
10. Bacci M (1996), Non-invasive instrumentation for detection and colour control of paintings and artworks, *Archeometriai Műhely*, 46-50.
11. Bradley M (2009), The importance of colour on ancient marble sculpture, *Art History*, **32** (3), 427-457.
12. Kalmus NM (1935), Color consciousness, *Journal of the Society of Motion Picture Engineers*, **25** (2), 139-147.
13. Gong Y, Chuan CH and Xiaoyi G (1996), Image indexing and retrieval based on color histograms, Multimedia Tools and Applications, **2** (2), 133-156.
14. Sethi IK, Coman IL, Day B, Jiang F, Li D, Segovia-Juarez J, Wei G and You B-M (1997), Color-WISE: a system for image similarity retrieval using color, *Proceedings of the SPIE 3312: Storage and Retrieval for Image and Video Databases VI*,140-149.
15. Stricker M and Orengo M (1995), Similarity of color images, *Proceedings of the SPIE 2420: Storage and Retrieval for Image and Video Databases III*, 381-392.
16. Stricker M and Swain M (1994), The capacity of color histogram indexing, *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 704-708, Seattle (USA).
17. Yihong G, Hongjiang Z and Chuan CH (1994), An image database system with fast image indexing capability based on color histograms, *Proceedings of TENCON'94 - 1994 IEEE Region 10's 9th Annual International Conference on: Frontiers of Computer Technology*, **1**, 407-411, Singapore.
18. Delon J, Desolneux A, Lisani J-L and Petro A-B (2007), A nonparametric approach for histogram segmentation, *IEEE Transactions on Image Processing*, **16** (1), 253-261.
19. Worring M and Gevers T (2001), Interactive retrieval of color images, *International Journal of Image and Graphics*, **1** (3), 387-414.
20. Talib A, Mahmuddin M, Husni H and George L (2003), Efficient, compact, and dominant color correlogram descriptors for content-based image retrieval, *Proceedings of the 5th International Conferences on Advances in Multimedia*, 52-61, Venice (Italy).
21. Adams R and Bischof L (1994), Seeded region growing, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **16** (6), 641-647.
22. Fan J, Yau DKY, Elmagarmid AK and Aref WG (2001), Automatic image segmentation by integrating color-edge extraction and seeded region growing, *IEEE Transactions on Image Processing*, **10** (10), 1454-1466.
23. Merzougui M and Allaoui AE (2019), Region growing segmentation optimized by evolutionary approach and maximum entropy, *Procedia Computer Science*, **151**, 1046-1051.
24. Tremeau A and Borel N (1997), A region growing and merging algorithm to color segmentation, *Pattern Recognition*, **30** (7), 1191-1203.

25. Sural S, Qian G and Pramanik S (2002), Segmentation and histogram generation using the HSV color space for image retrieval, *Proceedings of the International Conference on Image Processing*, **2**, II-589-II–592, Rochester (USA).
26. Ong SH, Yeo NC, Lee KH, Venkatesh YV and Cao DM (2002), Segmentation of color images using a two-stage self-organizing network, *Image and Vision Computing*, **20** (4), 279-289.
27. Arthur D and Vassilvitskii S (2007), k-means++: The advantages of careful seeding, *Proceedings of the 18th annual ACM-SIAM symposium on Discrete algorithms*, 1027-1035.
28. Aubaidan B, Mohd M and Albared M (2014), Comparative study of k-means and k-means++ clustering algorithms on crime domain, *Journal of Computer Science*, **10** (7), 1197–1206.
29. Buchsbaum G (1980), A spatial processor model for object colour perception, *Journal of the Franklin Institute*, **310** (1). 1-26.
30. Land EH (1977), The retinex theory of color vision, *Scientific American*, **237** (6), 108-129.
31. Hordley SD (2006), Scene illuminant estimation: Past, present, and future, Color Research and Application, **31** (4), 303-314.
32. Cheng D, Prasad DK and Brown MS (2014), Illuminant estimation for color constancy: why spatial-domain methods work and the role of the color distribution, *Journal of the Optical Society of America A*, **31** (5), 1049-1058.
33. Judd DB and Kelly KL (1939), Method of designating colors, *Contributions to Color Science*, 219–230.
34. Cobeldick S (2019), Convert between RGB and Color Names [last accessed 22 January 2020].
35. Cauwerts C and Jost S (2018), A color graphic informing on the impact of electric lighting and coated glazing in complex architectural scenes, *Color Research and Application*, **43** (6), 885-898.
36. Caliński T and Harabasz J (1974), A dendrite method for cluster analysis, *Communications in Statistics*, **3** (1), 1-27,
37. Desgraupes B (2013), Clustering indices, *University of Paris Ouest-Lab Modal'X,* **1**, 34.