

A colour space based on the image content

Javier Vazquez ^{a,1}, Maria Vanrell ^a, Anna Salvatella ^a and Eduard Vazquez ^a

^a *Computer Vision Center/Computer Science Department, Edifici O - Campus UAB*

Abstract. The main goal of this paper is to define a new colour space that represents the colour information of the image in such a way to give a more coherent spatio-chromatic representation. This space can allow to improve the performance of the algorithms of blob detection. To build the space we base colour representation on the ridges of the colour image distribution since it has been proved that they capture the essential colours of the image. Then we will define a colour space where each channel depends on one of the ridges. Finally, to select the essential channels we apply a Constraint Satisfaction algorithm that allows to get a reduced number of channels minimizing the correlation between them.

Keywords. colour space, blob detection

1. Introduction

In this work we propose a new colour space that adapts to the image content. The final goal for this new space is to achieve the best representation of colour information for a specific visual task. In particular, we focus on optimal representations for computational detection of colour blobs. Our proposal is a colour space that pursue the criterion of minimum inter-channel correlation. As a consequence of the procedure defined to build the space it also presents the property of similar entropy in all the channels.

In colour science a lot of different colour spaces have been defined [10]. Each one presenting different properties for different purposes. For instance, uniform spaces, such as, CIELAB or CIELUV, allow an Euclidean metric to represent perceptual similarity. Other spaces as RGB or CMY have been the basis for building acquisition, visualization or printing devices. In this paper we propose to define a new colour space that pursue a good representation to better extract the image content. As we already mentioned above, we will adapt the colour information to improve the detection of coloured image blobs.

Detection of coloured image blobs is a low-level visual task of a great importance in computer vision. In computer vision an image blob is a connected image region that presents an homogeneous colour. A successful extraction of image blobs can be the basis to overcome the subsequent steps in the image understanding process. Blob extraction is essential in the first steps of texture description [4], background subtraction [2],[3],

¹Correspondence to: Javier Vázquez Corral, Computer Vision Center/Computer Science Department, Edifici O - Campus UAB, 08193 Bellaterra (Barcelona), Spain Tel.: +34 935 811 828; Fax: +34 935 811 670; E-mail: jvazquez@cvc.uab.es.

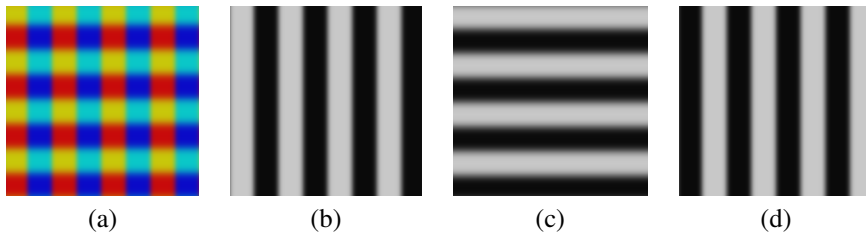


Figure 1. RGB channels of image (a), where (b) is the red channel, (c) the green channel, and (d) the blue channel

or motion analysis [5]. The computational approaches to deal with blob extraction have been essentially developed for gray-level images [6], where the laplacian filtering is the basis to extract image blobs. However, not such an effort has been done to extend this theory to colour images, and usually, the extension is done by just applying the gray-level algorithms on each colour channel separately. Hence the final detection of blobs is the combination of the blobs detected separately on each colour channel, usually red, green and blue. However, in figure 1 we can see that detecting blobs in RGB channels separately, does not assure to get the blobs we perceive in the colour image, in this case we can see small non-elongated blobs in red, yellow, light blue and dark blue, whereas in the RGB channels we have long elongated blobs in different orientations.

The results of this work are framed in the context of a project on automatic image annotation, where one of the goals is the description of textures. There are different approaches to extract and describe texture information and several works discuss how to deal with coloured texture [8]. A type of approaches are those that build the texture description based on the attribute of its blobs, following psychophysical theories [4]. These are the ones that motivates the goal of this paper, that is, to build a colour space that provide an adequate representation to detect colour image blobs as the basis for colour-texture description.

To this end, this paper has been organized as follows. In section 2 we introduce an algorithm to extract the essential colour information of an image, that is, the ridges of the colour distribution. Afterwards, in section 3 we propose a procedure to build a new colour representation based on this essential ridges. In section 4 we propose a method to reduce the number of dimensions of this new space, using a constraint satisfaction algorithm that allow to evaluate the feasibility of our proposal. Finally, we sum up the conclusions and explain different lines of research, since this is just a new research line we are just beginning to explore in this work.

2. Colour-Content Structure

In order to be able to define a colour space with the properties we have expressed above, we will need to extract the essential information of the image content. To this end we will extract this information dealing with the results of a recent work from Vazquez -et al [9] where they propose to cope the essential colour structure of an image by extracting the ridges [7] of the 3D colour distribution. In figure 2 we can see an example of the ridges of the image given in figure 1.(a). In figure 2.(a) we show a 3D coloured representation

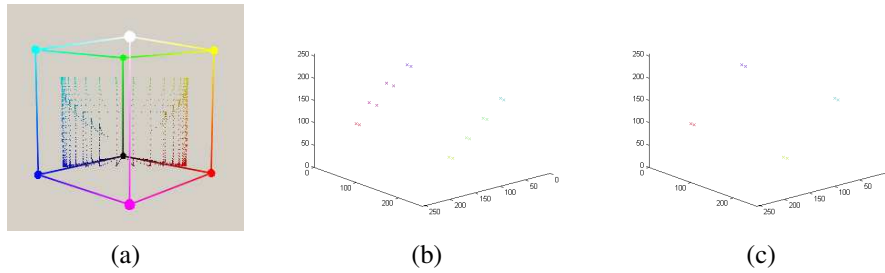


Figure 2. Explanation of the ridges of the image 1.(a). In (a) we could see a 3D coloured representation of the image, (b) are the detected ridges and (c) are the important ones

that gives an idea of its colour distribution. Figure 2.(b) gives the ridges detected by the method and in figure 2.(c) we show a simplification after removing the noisy ridges as we will explain in next section. This reduction let us to represent the four principal colours we perceive in the original image, these are red, yellow, dark blue and light blue.

It has been proved in [9] that ridges fulfill two essential properties which are those that allow them to cope the colour structure, these are:

- Connectivity of all the points in a ridge, that assure the connectivity of the influence zone.
- Peek and valley substraction, since the ridges extract all the distribution maxima plus all their nearby important colours.

Therefore, we will use this reduced representation of the essential image colour as the basis, in some sense, of our proposal for a new space. We can do it, by considering that any image point will belong to an influent zone of a ridge, that is to an essential colour. Influent zones are computed from the Voronoi diagram of the ridges on the colour distribution. The RGB values of an image point will determine to which influence zone it belongs.

3. Content-Based Colour Space

Before to define how to compute our proposal for a content-based colour space (CBCS), we will specify which are the main requirements we pursuit with it:

1. Distances in this space should correlate with perceived colour differences
2. Important blobs must maintain its perceived geometric structure
3. Each space dimension should represent a different colour property in order to cope with most of the color information.
4. All important blobs should appear at least in one of the space dimensions.

To fulfill these requirements we propose a colour space whose dimension will coincide with the number of important ridges we extract. We decide this number applying a preprocessing step on the ridges obtained with the algorithm of section 2. It is based on an iterative process that reduces those ridges containing a minimum number of pixels in its influence zone, afterwards all these points are redistributed to the remaining ridges. This is repeated iteratively until we achieve a prefixed number of ridges that will be the number of channels of our proposed colour space. The number of ridges we select will

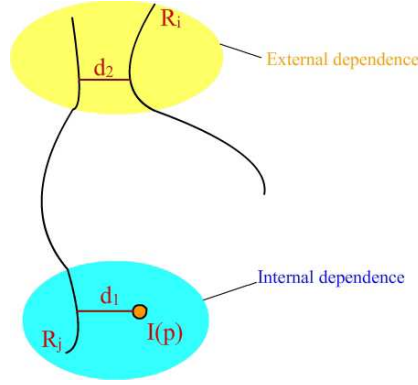


Figure 3. Scheme of ridge distances used to build the space.

determine the amount of information we are able to represent and this number will allow to fulfill the requirement 4 we have established before.

Therefore, each image channel will be related to a specific ridge. Considering that each image pixel belong to the influence zone of one specific ridge, we propose to build the components of pixel in the new space by using two distances, these are:

- The distance between the ridge of the pixel and the ridge of the channel
- The distance between the pixel and the ridge it belongs to.

This proposed distances, as it can be seen in figure 3 are representing two different dependencies, an internal dependency only depending on the pixel and its own ridge, and another external depending on the ridge of the channel. These two dependencies let the method to accomplish requirements 1 and 2.

Hence, we propose to build the i -component of a pixel p , we denote as $Ch_i(p)$ as

$$Ch_i(p) = \frac{Y_m}{\max_{q \in I} h(q)} h(p) \quad (1)$$

where p, q are pixels of the image $I : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, Y_m is the maximum intensity of the channel and

$$h(p) = \max_{q \in I} (f(q)) - f(p), \quad (2)$$

where $h(p)$ takes the 0 when p is the farthest pixel to ridge i , and $h(p)$ takes its maximum when p is the nearest, where $f(p)$ is representing the distance of p to the ridge i and its defined as:

$$f(p) = k_1 d_1(I(p), R_j) + k_2 d_2(R_i, R_j) \quad (3)$$

that weights the pixels of the new channel according to the distance to the image ridges, which are denoted as R_k , and depends on the image content. In this case, R_i is the ridge related to the i channel, and R_j is the ridge whom pixel p belongs to. Distances d_1 and d_2 are defined as

$$d_1(\vec{x}, R_j) = \inf_{\vec{y} \in R_j} d(\vec{x}, \vec{y}) \quad (4)$$

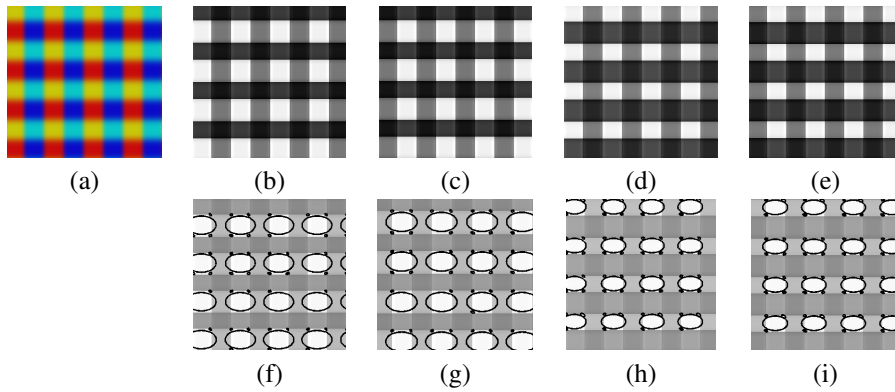


Figure 4. CBCS channels (b),(c),(d),(e) of an original image (a), and results of a blob detection (f),(g),(h),(i) in these channels where red, dark blue, yellow and light blue blobs are detected in this order

that is the infim distance between a point of the RGB space and its own ridge, and

$$d_2(R_i, R_j) = \inf_{\vec{x} \in R_i, \vec{y} \in R_j} d(\vec{x}, \vec{y}) \quad (5)$$

is the distance between two ridges, that is also computed by the infim distance between their points; $d(\vec{x}, \vec{y})$ is the euclidean distance between any two points, \vec{x}, \vec{y} of the space.

Finally, k_1 and k_2 are scalar factors that can be used to establish the entropy of the channels we have built. Depending on the distance between ridges we could serve these factors to order the ridges an increase or decrease the entropy, this would require to define k_2 as a vector of scalar factors, one for each $d_2(R_i, R_j)$.

In figure 4 we can see the results of the original image in figure 1 in our colour space at this level and the capability of our space to detect the four different types of blobs: red, yellow, light blue and dark blue.

4. Dimension reduction and evaluation

Once we have built a proposal for a content-based colour space, now we will try to see how the requirement 3 is fulfilled. That is, we propose to analyze what is the information included in each channel by the proposed space. We have proposed an algorithm based on a fixed number of channels, let say n . In this section we propose a constraint satisfaction algorithm that will allow to select a reduced version of the CBCS space based on the selection of channels that best represent all the image blobs.

The Constraint Satisfaction algorithm will allow to select a smaller number, m (where $m < n$), of channels selecting those ones that maximize the premise of better blob representation. We can see this process as searching the best m channels of the n dimensional CBCS representation fulfilling the constraint of minimum inter-channel correlation, that is to minimize:

$$F(a_1, \dots, a_m) = \sum_{i=1:m, j=1:m, i>j} |r(a_i, a_j)| \quad (6)$$

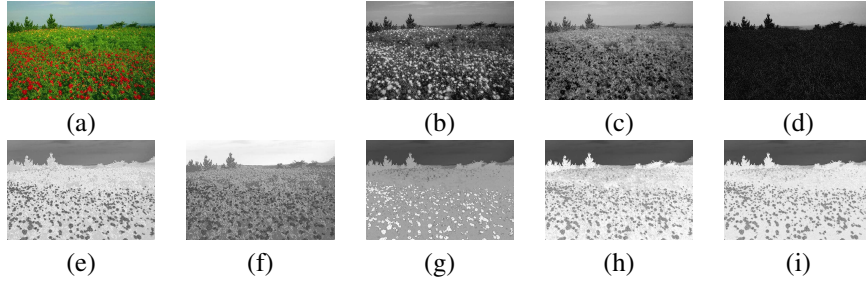


Figure 5. RGB channels and CBCS channels from an image. (b),(c),(d) are the RGB channels from image (a), and (e),(f),(g),(h),(i) are the CBCS channels where (h),(i) are the rejected ones

where $(a_1, \dots, a_m) \in A$, $A = (a_1, \dots, a_n)$ is the set of channels and r is the correlation coefficient computed as:

$$r(a_i, a_j) = \frac{\sum_{k=1:n} (x_k - \bar{x})(y_k - \bar{y})}{(n-1)s_x s_y} \quad (7)$$

where x_k, y_k are the values of the pixels in each channel, \bar{x}, \bar{y} are the mean of these values and s_x, s_y are the standard deviation of these values.

The m-dimensional representation we can derive as a solution of this constraint satisfaction problem allow to achieve requirement 3.

To evaluate the performance of this new colour space we have posed an experiment over a set of 200 images, built from the Mayang image dataset and from the Vistex image dataset. On these images we have computed the ridges of their colour distribution and we have computed their CBCS representation for $n = 5$. Afterwards we have applied a constraint satisfaction search on these channel and we have selected those dimensions minimizing interchannel correlation, we have done it for $m = 3$.

This experiment has allowed to compare this new representation to the classical RGB representation. While the RGB representation presents a mean inter-channel correlation of 2.41, the CBCS space for $m = 3$ achieve a mean inter-channel correlation of 1.29.

In figures 5 and 6 we show some examples of the 3-dimensional CBCS we have computed by applying the minimum inter-channel correlation criterion.

In fact, in figure 5 we could see from image 5.(a), their RGB channels 5.(b),5.(c),5.(d) in RGB order, and, as in this case CBCS space has been done with $n = 5$ and $m = 3$, five CBCS channels where 5.(e),5.(f),5.(g) are the channels selected by the Constraint Satisfaction algorithm and 5.(h),5.(i) are the rejected ones. We can see that in RGB channels information could be essentially found in 5.(b). On the other hand, in CBCS channels we found some regions of the plants in 5.(e), the sky in 5.(f), and the red flowers in 5.(g). In this channel we could also try to find yellow flowers with a detection of low intensity blobs. Furthermore, we can see that refused channels 5.(h),5.(i) are quite similar to 5.(e), this means they have a great correlation coefficient between them.

Finally, in figure 6 we can see the CBCS channels of some images (images on the left of each group), and, multiplying each channel by the image (images on the right), it is easy to percept what is the detected part of the image in this channel. For example, in 6.(b) it is quite clear that we detect the two different kinds of flowers in the second and the third channel and also one part of the plants in the first. It is also clear that in 6.(c)

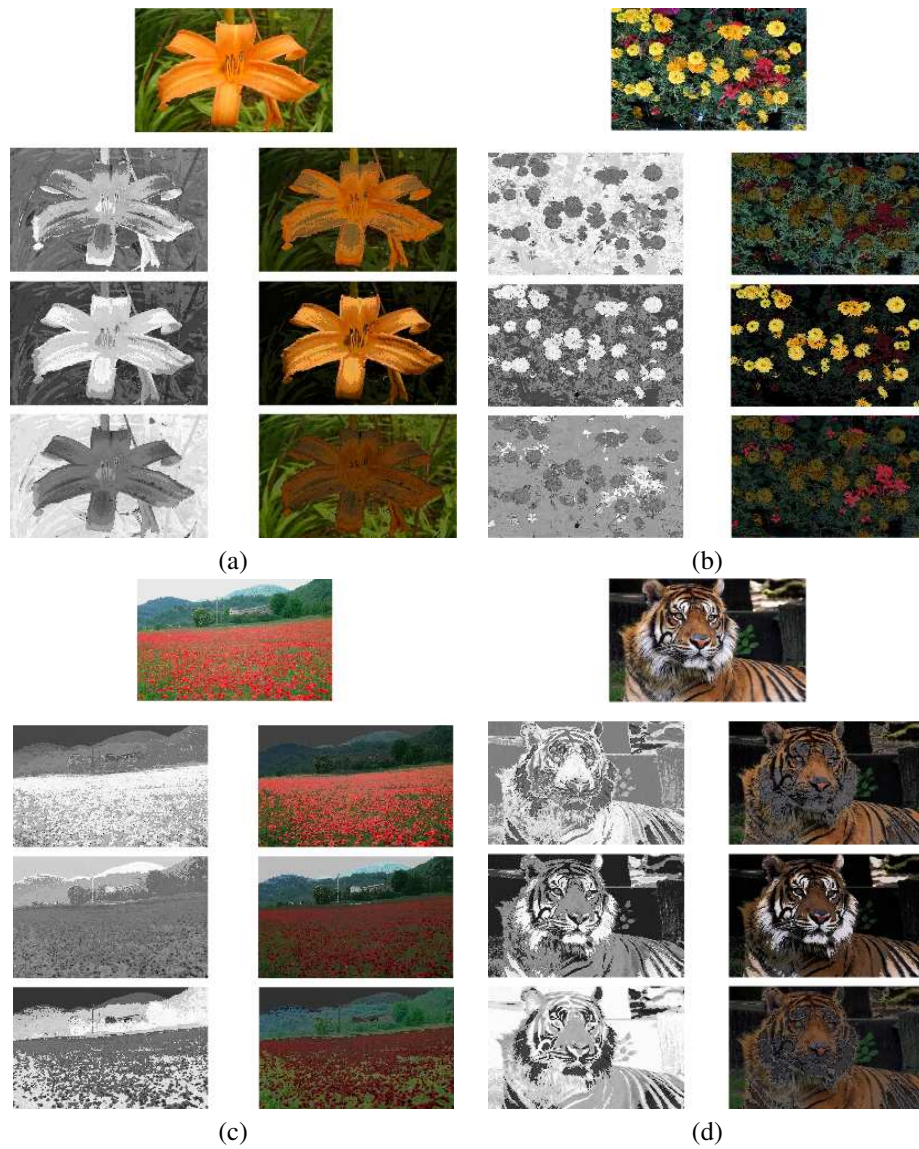


Figure 6. From each group we can see the original image (on the top), our CBCS channels (on the left) and what happens if we multiplies the original image by each CBCS channel, it means, if we stress the detected part of the image (on the right)

we are detecting red flowers in the first channel, the mountain of behind of the image in the second, and the nearest trees in the third. 6.(a) and 6.(d) show good results too.

5. Conclusions and further work

In this work, we have done a first step in our idea of adapting colour information to help texture descriptors. Results, as could be seen in section 4 are quite encouraging. But, this

is only a preliminary work that has opened some interesting ways to continue.

Firstly, considering that CBCS space is not linear some problems can appear, such as an ambiguous representation when two symmetric points in different directions of a given ridge becomes the same value in all the channels. This could be an important problem in some applications but, in the blob detecting case, as this points have different neighbours, d_1 will respect the inherent structure of the blob.

Another important research line is to make this space fully recoverable. Currently, we could do an approximation if we save the RGB values of the pixel p_{max} that has maximum intensity in the channel, and a 3D matrix, putting in its position (i, j) the values of the vector that starts in the RGB values of p_{max} and finishes in the RGB values of the pixel (i, j) .

Some other further work that could be done is to change the constraints of the CSF. This would be important if we want to extend our colour space to other different applications.

Finally, the introduction of colour names [1] could improve the method with two different objectives. On the first hand, we should discard those ridges having two different names, since they can introduce confusion. On the other hand, we should join ridges sharing the same colour name. These could allow to introduce robustness in the first steps of the algorithm.

References

- [1] R. Benavente, M. Vanrell, and R. Baldrich. Estimation of fuzzy sets for computational colour categorization. *Color Research and Application*, 29(5):342–353, 2004.
- [2] R. Collins, A. Lipton, and T. Kanade. A system for video surveillance and monitoring. In *American Nuclear Society 8th Internal Topical Meeting on Robotics and Remote Systems*, 1999.
- [3] Ismail Haritaoglu, Davis Harwood, and Larry S. David. W4: Real-time surveillance of people and their activities. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(8):809–830, 2000.
- [4] B. Julesz and J.R. Bergen. Textons, the fundamental elements in preattentive vision and perception of textures. *Bell Systems Technological Journal*, 62:1619–1645, 1983.
- [5] W. Hu L. Wang and T. Tan. Recent developments in human motion analysis. *Pattern Recognition*, 36(3):585,601, 2003.
- [6] T. Lindeberg and Jan-Olof Eklundh. On the computation of a scale-space primal sketch. *Journal of Visual Communications and Image Representation*, 2(1):55–78, 1991.
- [7] Antonio M. López, David Lloret, Joan Serrat, and Juan J. Villanueva. Multilocal creaseness based on the level-set extrinsic curvature. *Computer Vision and Image Understanding: CVIU*, 77(2):111–144, February 2000.
- [8] Pietikaainen Maempaa, T. Classification with color and texture: jointly or separately? *Pattern Recognition*, 37:1629–1640, 2004.
- [9] Eduard Vazquez, Ramon Baldrich, Javier Vazquez, and Maria Vanrell. Topological histogram reduction towards colour segmentation. *4477-0055 - Lecture Notes in Computer Science - Pattern Recognition and Image Analysis*, 2007.
- [10] G. Wyszecki and W.S. Stiles. *Color science: concepts and methods, quantitative data and formulae*. John Wiley & Sons, 2nd edition, 1982.