HYBRID COLOR SPACE CHOICE: AN OPTIMISATION REVIEW FOR COST/EFFICIENCY TRADE-OFF

Enguerran Grandchamp and Mohamed Abadi

Laboratoire GRIMAAG - Université Antilles Guyane - Campus de Fouillole - 97157 Pointe-à-Pitre Guadeloupe France
{egrandch|mabadi}@univ-ag.fr

ABSTRACT

This paper deals with image representation improvement using hybrid color spaces. This representation is important because it influences segmentation and classification results. We present two improvements of an existing supervised algorithm to obtain the most adapted hybrid color space for a given image. These improvements are based on a multi-objective optimization leading to a cost-efficiency trade-off, and have a theoretical justification. A comparison of the different approaches shows that the most adapted hybrid color space is reached with our algorithm and improves classification results.

Index Terms— Image color analysis, Optimization methods, Hybrid color space

1. BACKGROUND

About 20 color spaces ([1-5]) were defined since 1931 to reply to application necessity. These spaces are composed of three components having a different meaning. As example, some spaces have relations with human vision (color spaces based on Intensity hue and Saturation: HSV, IHS, IST, …) or were defined for specific applications (XYZ, Lab, $I_1I_2I_3$, …). There are mathematical relations (linear or not) to pass from a color space to another.
A color space can be seen as a way to represent an image in a three (or more) dimensional manner.
The choice of the color space depends on the images characteristics and then there isn't a unique color space adapted for all image types. The two main criterions are the correlation and the Discriminating Power (DP).
The correlation is computed between the different components. We often search a color space which have a low correlation, in other words with information distributed over all components.
The Discriminating Power (DP) [9-10] represents the ability of the color space to separate the information. We have to define classes and to localize samples to compute this criterion. It can be computed by different ways (Lawley-Hottelings or Pillai's criterion for example) by combining the inter class variance (maximization) and the intra class variance (minimization).
Some researches [6] have proved that classical color spaces are not the most adapted for classification or segmentation problems. First of all, all classical color spaces have three components which is not necessary or on contrary sufficient for some classification problems. Secondly, depending on the underlying classes, the component combination of classical color spaces are not the less correlated and/or the most discriminating.
For these reasons, some authors introduce hybrid color spaces [11-13]. These spaces are composed of an arbitrary selection of components taken from classical color spaces. As an example $RIXI_1$ is a hybrid color space having 4 components (R from RGB, I from HIS, X from XYZ and I_1 from $I_1I_2I_3$).

2. PROBLEMATIC

Because there is different ways to compute some components (especially non linear color space based on Intensity, Hue and Saturation) we can find more than 30 different color components. Let us note *N* this number.
The number of hybrid color spaces having *n* components is: $h(n) = \prod_{i=0}^{n-1}(N-i)$
So the total number of hybrid color spaces is: $H(N) = \sum_{i=1}^{N} h(i)$

For each image to process the main problem is now to select the best hybrid color space among this wide number of spaces regarding to a defined criterion.
We refer in section 3 to an already existing algorithm and show its limits.
Section 4 and 5 present two optimization ways to improve the research of the best hybrid color space.
Section 6 presents the main results and a comparison between the different approaches and section 7 presents the conclusion of this study.

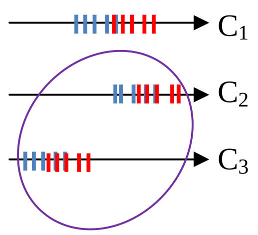
3. EXISTING ALGORITHM

3.1. Principle

An existing way ([11]) to build a hybrid color space is to iteratively add a new color component in two successive steps *(i)* a minimization of the correlation of the hybrid color space *(ii)* a maximization of the discriminating power (DP) among selecting spaces.
Step one is done by computing the correlation between the current hybrid color space (the already selected components) and all the remaining ones and by rejecting components having a correlation with the current hybrid color space higher than a given threshold.
Step two is done by selecting among the remaining components the one that maximize the discriminating power.

3.2. Limits

This process looks like a local greedy mono-objective optimization algorithm with a neighborhood restriction.
The main problem is that this algorithm doesn't systematically converge to the best hybrid color space regarding the two criterions (correlation and discriminating power). Moreover, with such an algorithm, the authors implicitly suppose the separability of the two criterions which is wrong. Figure 1 illustrates these points.
The example is built as follow: we use one dimension data and three components (C_1, C_2, C_3). Samples are chosen in order to have a discriminating power greater for C_1 than for C_2 or C_3. C_2 is chosen to be correlated with C_1 and uncorrelated with C_3. The discriminating power is lower for $C_1 \cup C_3$ than for $C_2 \cup C_3$.
The algorithm [11] return $C_1 \cup C_3$ as being the best hybrid color space but $C_2 \cup C_3$ has a lower correlation and a greater discriminating power than $C_1 \cup C_3$.



Step 1: C_1 selected ($DP(C_1) > DP(C_3) > DP(C_2)$)
Step 2: C_3 selected ($Corr(C_1 \cup C_3) < Corr(C_1 \cup C_2)$)
Resulting hybrid space: $C_1 \cup C_3$
Best hybrid space regarding DP and Corr : $C_2 \cup C_3$

Figure 1. existing algorithm limits

Finally, this algorithm requires a threshold for correlation restriction which can be hard to fix.

4. FIRST IMPROVEMENT: LOCAL GREEDY MULTI-OBJECTIVE OPTIMIZATION

4.1. Principle

This first improvement consist on changing a mono-objective optimization to a multi-objective one (we search the hybrid color space having the best *(Corr, DP)* couple instead of alternate Steps 1 and 2).
Because correlation is defined with at least two components, we start the algorithm by examining all accessible couples of components. The existence of non dominated solutions [14] leads to keep at the end of each iteration a set of solutions instead of a unique one. This set is called the Pareto set [14] of visited solutions.
At the next iteration, we start from the previous Pareto set and add a new component among the remaining ones.
The end criterion is a given number of components reached or else the stability of the Pareto set.
The selection of the final hybrid color space is made among the final Pareto set depending on the application. This selection translates the Corr-DP trade-off which can be automatically set by founding the inflection point of the Pareto curve.

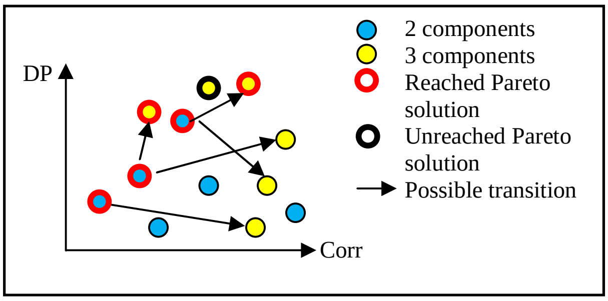


Figure 2. First optimization

Figure 2 illustrates this algorithm and plots the hybrid clor spaces in the (Corr, DP) plan. Blue discs represent all two components hybrid color spaces, yellow ones represent three components ones and red circle all Pareto solutions. We can note that the Pareto solutions are non dominated (no other solution has a lower correlation and a higher DP).

4.2. Limits

This algorithm doesn't allow eliminating a previously added component. This is the greedy aspect of the method.
Because we always start from the previous Pareto set (we don't explore every possible solution), some interesting hybrid color spaces are perhaps not reachable with this algorithm. (Black circle in Figure 2).

5. SECOND IMPROVEMENT: LOCAL MULTI-OBJECTIVE OPTIMIZATION

5.1. Principle

In this second approach we extend the neighborhood of the current solution to unvisited hybrid color spaces having *(i)* one more component (by adding one), *(ii)* the same number

of components (by replacing one), *(iii)* a fewer number of components (by deleting one).
The algorithm only manages one solution. We don't build a Pareto set at each iteration of the algorithm but we memorize all non dominated solutions visited during the research in order to return a global Pareto set.

The algorithm is the following *(i)* start from one Pareto optimal hybrid color space with two components *(ii)* explore the neighborhood according to the three kind of transitions *(iii)* choose among non dominated neighbors the next retained solution or if it doesn't exist among one previously visited.

The research stops when it reaches an empty neighborhood.

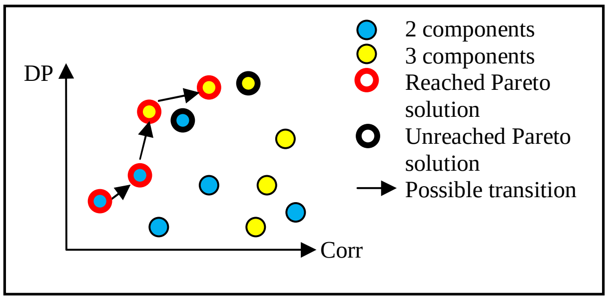


Figure 3. Second optimization

Figure 3. shows an illustration of the algorithm. The visited solutions compose as one goes along the Pareto set (Red circles).

5.2. Limits

Because of the local aspect of the algorithm and the random choice of the next visited solution the resulting Pareto set could slightly differs from one run to another. As example, Figure 3 shows some unreached Pareto solutions (Black circles).

6. COMPARAISON

To compare the three algorithms, we use about 20 images and found the most adapted hybrid color space for each images using each algorithm.
We illustrate obtained results with the image presents in figure 4. This image is composed of 6 textures and we are trying to classify it into 6 classes only by using the color information.

Figure 5. shows the results by plotting the correlation, Discriminating Power couples *(Corr, DP)* of the hybrid color spaces.
To value the quality of the results regarding to all possible hybrid color space, we exceptionally compute every hybrid color spaces (black points in figure 5.).

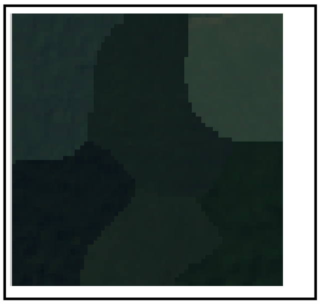


Figure 4. Illustration image

The black line is the Pareto set of all possible hybrid color space. The blue line is the Pareto set obtained using the first improvement and the red line is the Pareto set obtained using the second one.

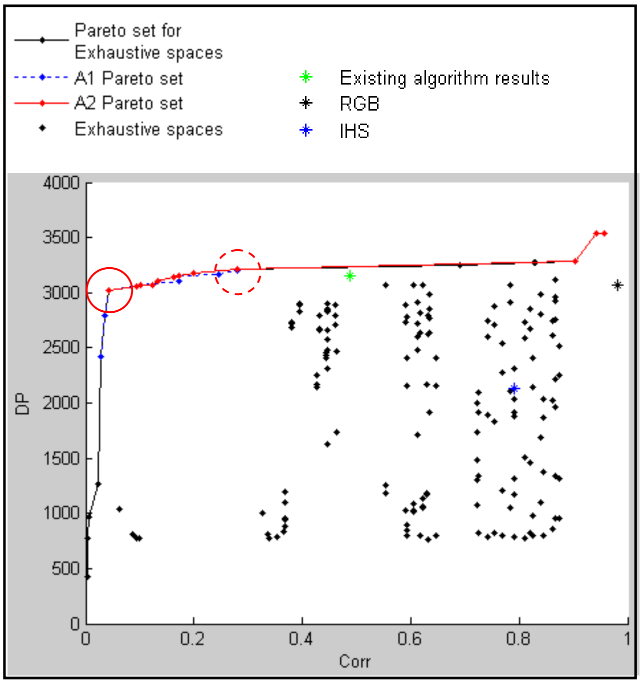


Figure 5. Results

First of all, we can see that the Pareto sets obtained by our improvements are included into the global Pareto set. This confirms that we obtain non dominated solutions. Moreover, this Pareto sets are very close (Figure 6).

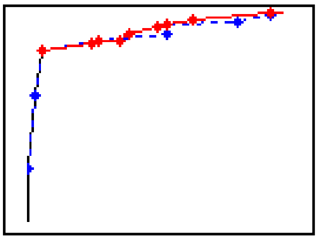
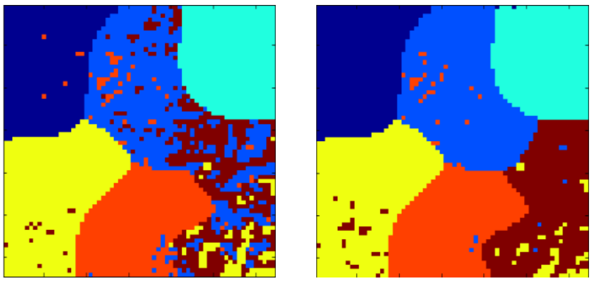


Figure 6. Zoom of Pareto set

The green star represents the *(Corr, DP)* couple for the hybrid color space obtained using the existing algorithm. Even if it is close to the Pareto set, this space is dominated by the two Pareto sets (A1 and A2 in Figure 5.). As example, the red dotted circle shows a hybrid color space which dominates it. This illustrates that our algorithms reach better solutions. The red circle is the best correlation-discriminating power trade-off.

We also plot in figure 5. two classical color spaces (RGB: black star and HIS: blue star) to show the improvement of hybrid color spaces.

Now we compare the classification results obtained using a k-means algorithm apply to *(i)* the hybrid color space components resulting from the existing algorithm (Figure 7.a) *(ii)* the hybrid color space components resulting from the best correlation-discriminating power trade-off (red circle in Figure 5) (Figure 7.b).



a) existing algorithm b) best corr-DP trade-off

Figure 7. Classification results

We can see in Figure 7 that the classification is better using the best Corr-DP trade-off than using the hybrid color space returned by the existing algorithm.

7. CONCLUSION

The search of an optimal color space to represent an image leads to explore hybrid color spaces. The way to find the most adapted one must ensure to reach it in an acceptable time. The wide search space doesn't allow exploring every possible hybrid color spaces. An existing algorithm allows reaching one but with unfounded principles. We present two improvements to reach non dominating solutions (Pareto set) which is a necessary condition to reach the most adapted hybrid color space. These algorithms differ from their principle and computation time but lead to very close Pareto sets. Considering all realized tests, our approaches always lead to better solutions (at least equivalent ones) than the existing algorithm.

Time computation could be reduced by restricting hybrid color spaces to those build from non linear classical color spaces, that is to say built from Intensity, Hue and saturation components. This reduction is made according to applications considerations. The main difficulty is to prove that the corresponding Pareto set is at least included into the global Pareto set in order to ensure reaching the most adapted hybrid color space for a given image.

8. REFERENCES

[1] Macadam D. L. “Color Measurement, Theme and Variations”, *2nd ed. Springer-Verlag*, 1985.

[2] Travis, D. “Effective colour displays : Theory and Practice”, *1ed ed. Academic Press*, 1991.

[3] Carron, T. "Segmentation d’images couleur dans la base Teinte-Luminance-Saturation: approche numérique et symbolique". *PhD thesis*, Université de Savoie, France, 1995.

[4] Lefèvre, S., AND Vincent, N. Apport de l’espace teinte-saturation-luminance pour la segmentation spatiale et temporelle. Traitement du signal 23, 1 (2006), 59–77.

[5] Levkowitz, H., AND Herman, G. T. “Glhs : a generalized lightness, hue, and saturation color model.” *CVGIP : Graphical Models and Images Processing* 55, 4 (1993), 271–285.

[6] Guerin, D., Cointault, F., Gée, C., and Guillemin, J.P. "Etude de faisabilité d’un système de comptage d’épis de blé par vision." *Revue Traitement du Signal, Session Spéciale Imagerie Couleur* 21, 5 (2005), 549–560.

[7] Ohta, Y.I., Kanade, T., and Sakai, T. “Color information for region segmentation”. *Computer Graphics and Image Processing* 13, 3 (1980), 222–241.

[8] Tominaga, S. “Color classification of natural color images”. *Color Research and Application* 17, 4 (1992), 230–239.

[9] Cocquerez, J. P., and Phillip, S. Analyse d’images : Filtrage et segmentation. *Masson*, 1995.

[10] Gonzalez, R. C., AND Woods, R. E. “Digital Image Processing, 3rd ed”. *Prentice-Hall*, Inc., Upper Saddle River, NJ, USA, 2006.

[11] Vandenbroucke, N., Macaire, L., and Postaire, J.G., “Color image segmentation by pixel classification in an adapted hybrid color space. application to soccer image analysis.”, *Computer Vision and Image Understanding 90*, 2 (2003), 190–216.

[12] Macaire, L. "Exploitation de la couleur pour la segmentation et l’analyse d’images". *PhD thesis*, Université des sciences et technologies de Lille 1, 2004.

[13] Vandenbroucke, N. "Segmentation d’images couleur par classification de pixels dans des espaces d’attributs colorimétriques adaptés. Application à l’analyse d’image de football."
PhD thesis, Université de Lille 1, France, 2000.

[14] J. G. Lin. “Three Methods for Determining Pareto-Optimal Solutions of Multiple-Objective Problems”. *Directions in Large-Scale Systems*, pp. 117-138. Edited by Y. C. Ho and S. K. Mitter. New York, Plenum Press, 1975.