



transition between low values and highest saturation is the intensity-dependent [7]. It has been shown by [7] that for higher values of the intensity, saturation of 0.2 distinguishes domination between hue and intensity. Assuming that the maximum intensity is 255, the function of threshold to determine if a pixel should be represented by its hue or its intensity is given by:

$$Th_s(V) = 1 - \frac{0.8V}{255} \quad (2)$$

which can lead to a characteristic component of both parties: the values of hue between 0 and  $2\pi$  quantified after a transformation and a set of quantified intensity values.

Abutaleb [8] introduced another type of histogram 2-D grayscale, monochrome image which differs from the co-occurrence matrix grayscale. It is formed by the Cartesian product of the grayscale classic histogram 1-D and the local average of 1-D grayscale histogram obtained by applying a local window to each pixel in the image and by calculating the average of the gray levels in the window. In fact, the change in the value of pixels in the horizontal or vertical directions appears slow and diagonal direction the continuity of gradation change appears strong. Zhang and Zhang [9] propose to use the minimum gray value in the four-neighbor and the maximum gray value in the vicinity of a window of (3 x 3) exceptions of pixels in the four-neighbor.

On the basis of the above description, we are developing a new approach for image retrieval. In the first step, we use the space color HSV to construct the histogram where each pixel is represented by its hue or its intensity.

On the basis of the threshold function (2), it determines an intermediate image (Fig. 1 (b)): for values low saturation, a color can be approximated with a gray value specified by the intensity level, while for higher saturation, the color can be approximated by its hue value.

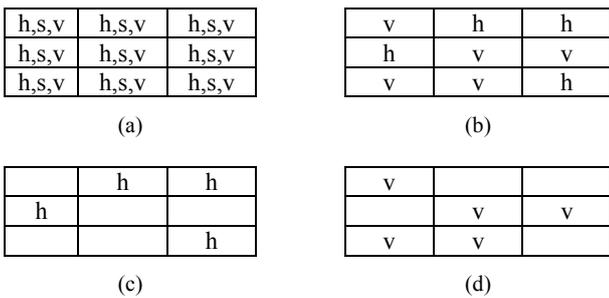


Fig. 1 (a) original image in HSV (b) Intermediate image (c) Pixels of the image represented by the hue (d) Pixels of the image represented by the intensity

Then, we build two histograms color 2-D for intermediate images (Figs. 1 (c) and (d)) as follows: for each block of pixels 3 x 3 of two intermediate images, we consider the central pixel which can be a component of intensity (Fig. 2 (a)) or hue component (Fig. 2 (c)), according to (2), and calculate the maximum value (Max) and minimum value (Min) in the

component matching (H or V) in the original image (Figs. 2 (b) and (d)). While Max is the maximum hue or intensity in the 3 x 3 block, except four-neighbor pixels and Min is the minimum hue or intensity in the four-neighbor.

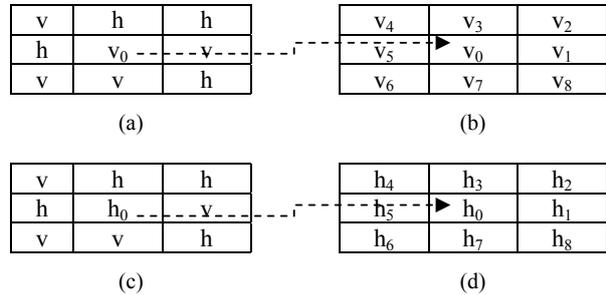


Fig. 2: 3x3 block: a) and (c) intermediate image; (b) and (d) their component in the original image

Max = maximum [(v<sub>4</sub>, v<sub>2</sub>, v<sub>6</sub>, v<sub>8</sub>) or (h<sub>4</sub>, h<sub>2</sub>, h<sub>6</sub>, h<sub>8</sub>)] and Min = minimum[(v<sub>0</sub>, v<sub>1</sub>, v<sub>3</sub>, v<sub>5</sub>, v<sub>7</sub>) or (h<sub>0</sub>, h<sub>1</sub>, h<sub>3</sub>, h<sub>5</sub>, h<sub>7</sub>)].

Then, we build each Histogram 2-D by mapping the number of pixels with the minimum and maximum (Min and Max) in the block 3 x 3 image depending on P<sub>h</sub> (Min, Max) and P<sub>v</sub>(Min, Max) for the central pixel represented by the component H or V respectively.

Finally, we will obtain a 2-D histogram in the HSV space (Fig. 3 (c)). It consists of two separate histograms. The first represents the pixels associated with the values of hue (Fig. 3 (a)) and the second with the intensities (Fig. 3 (b)).

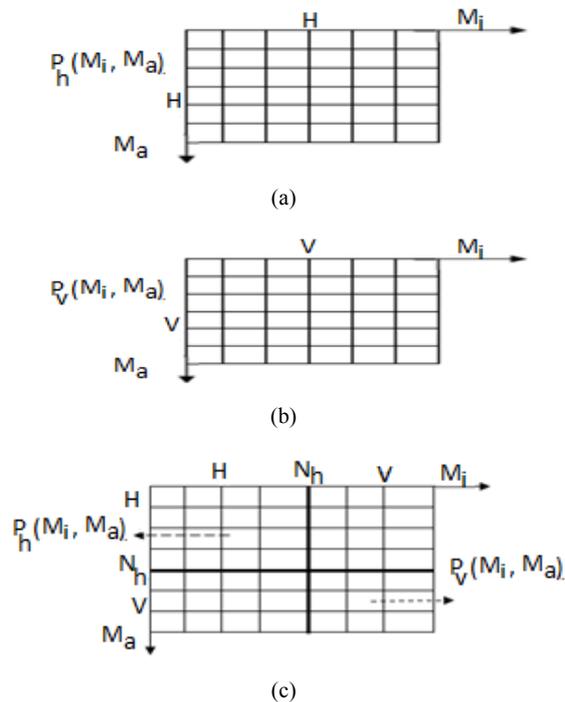


Fig. 3 (a) and (b) Histograms representing the pixels associated with the values of the hue and the intensities (c) 2-D Histogram

The proposed histogram is based on a window 3 x 3 and not on the intensity of the pixel only. This approach overcomes the disadvantage of the classical histogram, which lies in the fact that it ignores the spatial distribution of pixels in the image.

#### IV. HISTOGRAM AND HISTOBIN

A technique widely used for color is the histogram intersection. With this method, first we must compute the image histogram (histogram of an image can be presented by a vector in which each component is a number of color pixels corresponding to the index). Then a histobin will be created from the histogram as follows: each bin (hole) in the histobin is the sum of a few neighbouring elements of the histogram. The number of neighbours is determined by the number of the bin histobin. Let N the number of bins for each component, the histobin has 3N bins in total. The number of neighbours is 256 / N. One can see that the histobin is more compact than the histogram. In our system we calculate the histobin for each color component in HSV space. Having had the histobin image, the distance between two images is the distance between two histobins. This distance can be calculated by [10]:

$$D_{Hist}(Q, I) = \frac{\sum_i |Q_{Histoi} - I_{Histoi}|}{\sum_i Q_{Histoi}} \quad (3)$$

where: Q, I are two images;  $Q_{Histoi}$  is the Q histobin;  $I_{Histoi}$  is the I histobin;  $Q_{Histoi}$  is the bin i of Q histobin;  $I_{Histoi}$  is the bin i of I histobin;  $D_{Hist}(Q, I)$  is the distance between two images Q and I according to the intersection of histograms.

#### V. STATISTICAL MOMENTS

To escape the effects of the discretisation of the color space which is intrinsic to the use of histograms, another approach has been used successfully in several image retrieval systems; it is the statistical moments of colors. This approach involves calculating a weighted sum of the average and the variance and the third moment for each color channel, for providing a unique number used to index.

In [11] the authors showed that the methods of using statistical moments goes faster and give better results than the methods of histogram.

If  $p_{ij}$  is the value of the pixel j for channel i (i: HSV), N is the number of pixels in the image, these moments are defined by:

$$\mu_i = \frac{1}{N} \sum_{j=1}^N p_{ij} \quad (4)$$

$$\sigma_i = \sqrt{\frac{1}{N} \sum_{j=1}^N (p_{ij} - \mu_i)^2} \quad (5)$$

$$s_i = \left( \frac{1}{N} \sum_{j=1}^N (p_{ij} - \mu_i)^3 \right)^{1/3} \quad (6)$$

The moment  $\mu_i$  of the first order is the average color of the image. The second moment  $\sigma_i$  characterizes the contrast of an image. More variance of colors is big, more the picture is contrasted. The moment of the third order  $s_i$  characterizes the amount of light in an image. An image with a positive dissymmetry coefficient tends to appear darker and brighter than a similar image with a lower coefficient of dissymmetry.

#### A. Calculation of the Distance between two Different Images

The distance between the query image Q and an image I from the database using the statistical moments is given by:

$$D_{Mom}(Q, I) = \left( \sum_i [(\mu_i(Q) - \mu_i(I)) + (\sigma_i(Q) - \sigma_i(I)) + (s_i(Q) - s_i(I))]^2 \right)^{1/2} \quad (7)$$

#### VI. COMBINATION OF FEATURES

Since each distance between features was calculated using a different standard, prior to combining these features, we normalize the distances using a Gaussian [12]. This method ensures that the distance is normalized between 0 and 1. After the normalized distances, we used a linear combination of distances.

$$D(Q, I) = w_{Hist} * D_{Hist}(Q, I) + w_{Mom} * D_{Mom}(Q, I) \quad (8)$$

where  $w_{Hist2D}$  and  $w_{Mom}$  take a value between 0 and 1. This value can be a new program argument.

#### VII. ARCHITECTURE

To test and to implement our descriptor for color image retrieval, we develop the following algorithm step.

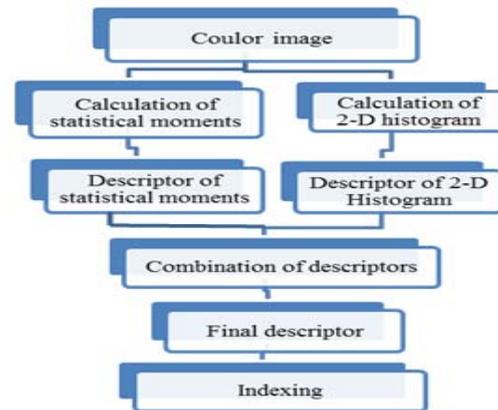


Fig. 4 Process of calculation of our descriptor

In Fig. 4, the way to create the final descriptor is performed in a distributed computing. Indeed, the input image will be indexed as follows: we first create two processes. The first process calculates histogram 2-D, and then it creates the histobin and finally the distance. While the second process calculates statistical moments to create descriptor of statistical

moments .these two processes work in parallel, and finally we combine the results to have the final descriptor. This technique accelerates our algorithm according to execution time.

VIII. EXPERIMENTAL RESULTS

In our system and to evaluate the new algorithm, we tested each feature as well as combinations of these features. We display the first 16 similar images. The results obtained on several image databases [13]-[15] will be presented in the figures (Figs. 5-8).

A. Tests with Wang Database[13]



(a)



(b)



(c)



(d)

Fig. 5 Similar images for class: Flowers [13]: (a) Results obtained by RGB algorithm using 24 bins (b) Results obtained by HSV algorithm using 24 bins (c) Results obtained by the 2-D histograms using 24 bins (d) Results obtained by the new algorithm

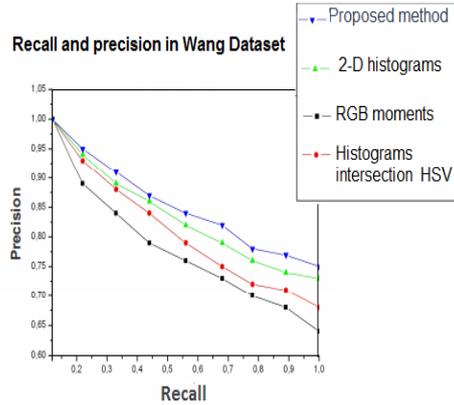


Fig. 6 Recall and precision in Wang Dataset [13]

B. Tests with Coil Database [14]

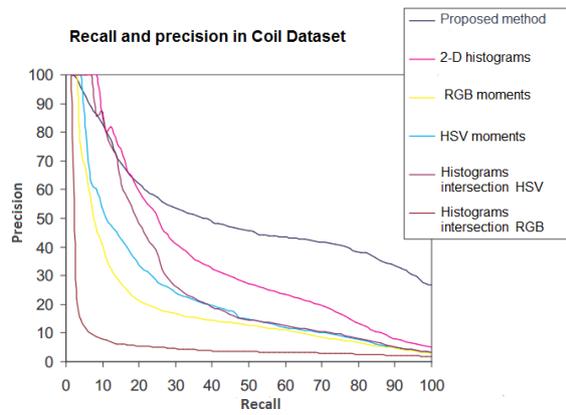


Fig. 7 Recall and precision in Coil Dataset [14]

C. Tests with FeiFei Database [15]

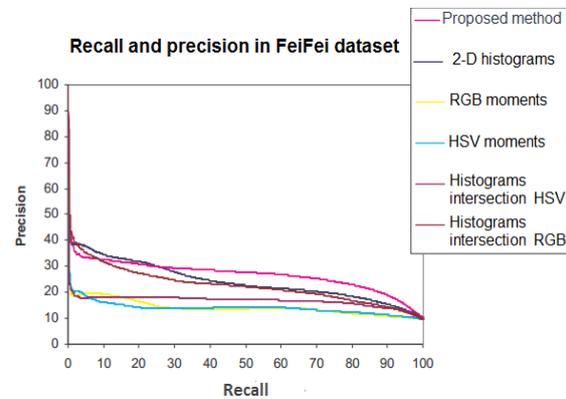


Fig. 8 Recall and precision in FeiFei Dataset [15]

D. Comments

When applying the proposed algorithm on class: Flowers [13], as shown in Fig. 5. Fig. 5 shows the top 16 retrieval results of a query (the top left image is the query image and others are retrieved images). Our system performs based on combine both 2-D histogram and statistical moments in the HSV space clearly give better performance in image retrieval than systems of indexing using Figs. 5 (a)-(c).

To evaluate the test results, we use the values precision and recall rate. These values are defined as follows:

$$\text{precision} = \frac{\text{number of relevant images selected}}{\text{total number of retrieved images}} \quad (9)$$

$$\text{recall} = \frac{\text{number of relevant images selected}}{\text{total number of similar images in the database}} \quad (10)$$

The proposed algorithm tested with many databases. Figs. 6-8 present the average precision curves of different representations and similarity measure without any indexing scheme. The performances of the proposed method using 2-D histograms and statistical moments in the HSV space are evaluated and compared with the classical histograms intersection in RGB and HSV color space and the statistical moments in RGB, HSV. From Figs. 6-8, it is clear that the proposed approach performed better than other methods mentioned above according the curve of recall and precision.

#### IX. CONCLUSION

We introduced in this study new content-based color image retrieval algorithm based on combine both 2-D histogram and statistical moments in the HSV color space. This approach is to overcome limitations of conventional 1-DHistogram. The method was applied to different real images to demonstrate the performance of the algorithm of color image retrieval. The results obtained were compared to other methods, or those obtained by the intersection of the 1-D histograms in the RGB space and HSV, or those obtained by the intersection of the 2-D histograms in the HSV space, or with those obtained by statistical moments in RGB and HSV spaces. From the figures, we see that the new method is more efficient than other methods mentioned above.

#### ACKNOWLEDGMENTS

The authors would like to thank the anonymous reviewers for their insightful comments in improving the quality of this paper.

#### REFERENCES

- [1] Niblack, W., Barber, R., Equitz, W., Flickner, M., Glasman, E., Petkovic, D., Yanker, P., Faloutsos, C., & Taubin, G. "The QBIC Project: Querying Images by Content Using 37 Color, Texture, and Shape". SPIE Int. Symp. On Electronic Imaging: Science and Technology Conf. 1908, Storage and Retrieval for Image and Video databases. (1993).
- [2] Carson, C., Thomas, M., Belongie, S., Hellerstein, J. M., & J. Malik. "Blobworld: A System for Region-based Image Indexing and Retrieval", Proc. Third Int. Conf. on Visual Information Systems, June. (1999).
- [3] Clément, A., Vigouroux, B. "Unsupervised segmentation of scenes containing vegetation (Forsythia) and soil by hierarchical analysis of bi-dimensional histograms". Patt. Recogn. Lett, Vol. 24, 1951–1957. (2003).
- [4] Masmoudi, Lh., Zennouhi, R., & EL Ansari, M. Participation with a chapter «Image Segmentation Based on Two Dimensional Histogram» in a book «image segmentation» published by InTech, (2011).
- [5] Cheng, H. D., Jiang, X. H., Sun, Y., Wang, J. "Color image segmentation: advances and prospects". Pattern Recognition, Vol.34, No.6, 2259-2281. (2001).
- [6] Zennouhi, R., Masmoudi, L. H. "A new 2D histogram scheme for color image segmentation". The Imaging Science Journal, Vol. 57, 260-265. (2009).
- [7] Sural, S., Qian, G., Pramanik, S. "Segmentation and histogram generation using the HSV color space for image retrieval", Proc. Int. Conf. on Image processing: ICIP'02, Rochester, NY, USA, IEEE, Vol. 2, pp. II589–II592. (2002).
- [8] Abutaleb, A. S. "Automatic thresholding of gray-level pictures using two-dimensional entropy". Journal of Computer Vision, Graphic and Image Process, Vol. 47, 22–32. (1989).
- [9] Zhang, Y. F., Zhang, Y. "Another Method of Building 2D Entropy to Realize Automatic Segmentation". Journal of Physics Conference Series, Vol. 48, 303–307. (2006).
- [10] Swain M.J., Ballard D.H., "Color indexing". International Journal of Computer Vision, vol. 7, no. 1, pp. 11-22. (1991).
- [11] Stricker, M., Orengo, M. "Similarity of Color Images", In Proceedings of SPIE, Vol. 2420 (Storage and Retrieval of Image and Video Databases III), SPIE Press, Feb. (1995).
- [12] Iqbal, Q., Aggarwal, J. K. (2002), "Combining structure, color and texture for image retrieval: a performance evaluation". Proc. of International Conference on Pattern Recognition (ICPR), Quebec (Canada). (2002).
- [13] Penn state university's web page for modeling objects, concepts, and aesthetics in images project. Online available at: <http://wang.ist.psu.edu/docs/related/>.
- [14] Coil (Columbia Object Image Library)dataset, Online available at: <http://www1.cs.columbia.edu/CAVE/research/softlib/>
- [15] The Fei-Fei dataset, Online available at: <http://www.vision.caltech.edu/feifeili/Datasets.htm>.



**El Asnaoui Khalid** was born in Awddim Hsiya -Morocco in 1985. Received the master's degrees in Computer Engineering from Faculty of Sciences Oujda (FSO), Mohamed First University (UMP) in 2013. He is currently working toward the PhD degree in the ASIA Team at the Faculty of Sciences and Techniques-Er Rachidia. His research areas are Image Processing and Pattern Recognition.

(Corresponding author)