INSTITUTO DE COMPUTAÇÃO
UNIVERSIDADE ESTADUAL DE CAMPINAS

Evaluating the Potential of Texture and Color Descriptors for Remote Sensing Image Retrieval and Classification

J. A. dos Santos *O. A. B. Penatti*
R. da S. Torres

Technical Report - IC-09-47 - Relatório Técnico

December - 2009 - Dezembro

Evaluating the Potential of Texture and Color Descriptors for Remote Sensing Image Retrieval and Classification

Jefersson Alex dos Santos Otávio Augusto Bizetto Penatti
Ricardo da Silva Torres*

Abstract

Classifying Remote Sensing Images (RSI) is a hard task. There are automatic approaches whose results normally need to be revised. The identification and polygon extraction tasks usually rely on applying classification strategies that exploit visual aspects related to spectral and texture patterns identified in RSI regions. There are a lot of image descriptors proposed in the literature for content-based image retrieval purposes that can be useful for RSI classification. This paper presents a comparative study to evaluate the potential of using successful color and texture image descriptors for remote sensing retrieval and classification. Seven descriptors that encode texture information and twelve color descriptors that can be used to encode spectral information were selected. We highlight the main characteristics and perform experiments to evaluate the effectiveness of these descriptors. To evaluate descriptors in classification tasks, we also proposed a methodology based on KNN classifier. Experiments demonstrate that Joint Auto-Correlogram (JAC), Color Bitmap, Invariant Steerable Pyramid Decomposition (SID) and Quantized Compound Change Histogram (QCCH) yield the best results.

1 INTRODUCTION

Agriculture has an important role in the economy of several countries. The results of agricultural activities are directly linked to the productivity. Therefore, many researches have been investigating new ways to improve the agricultural practices and, consequently, to increase the quantity and quality of what is produced. In this scenario, crop monitoring is a fundamental activity and using Geographic Information Systems (GIS) has made it easier.

Some of the main issues related to crop monitoring are: How is the land occupation? What is cultivated in a given region? Where are some cultures cultivated?

Remote Sensing Images (RSIs) provide the basis for the creation of information systems that support the decision-making process based on land cover changes. Using RSI in crop monitoring requires the recognition of the regions of interest and the extraction of the polygons around these regions.

The identification and polygon extraction tasks usually rely on applying classification strategies that exploit visual aspects related to spectral and texture patterns identified in RSI regions. These tasks can be performed automatically or manually. The “manual” approach is based on image editors where users can define or draw polygons that represent regions of interest using the raster image as background. In general, automatic approaches use classification strategies based on pixel

*Authors are grateful to FAPESP, CAPES, CNPq and Microsoft Research for financial support.

information. The most used pixel classification algorithm, MaxVer [1], however, is not very effective. Several new methods have been proposed to improve the performance of MaxVer-based techniques. In [2], a new method considering image segmentation, GIS, and data mining algorithms was presented. Compared with pixel-based classification, the results showed best agreement with visual interpretation. The work proposed in [3] applied a morphological filter in an image which was classified by MaxVer algorithm. The results were compared with the other classification algorithms (Fisher linear likelihood, minimum Euclidean distance and ECHO). In [4], three Land Cover Classification Algorithms are compared for monitoring North Korea using multi-temporal data.

The main drawback of automatic approaches is concerned with its sensitivity to image noises (e.g., for example, distortions that can be found in mountainous regions). Another important problem in the automatic approaches is concerned with the fact that they usually fail to correctly identify borders between distinct regions within the same image. Thus, in practical situations, the results obtained need to be revised. As these revisions take a lot of time, it is sometimes more convenient to the user to perform recognition manually.

Content-based Image Retrieval (CBIR) systems are developed to provide efficient and effective means to retrieve images. In these systems, the searching process consists of, for a given image, computing the most similar images stored in the database, considering only image properties, like color and texture, for instance. The searching process relies on the use of image *descriptors*. A descriptor can be characterized by two functions: *feature vector extraction* and *similarity computation*. The similarity between two images is computed as a function of their feature vectors distance.

Recently, some descriptors for RSI purposes has been proposed. Tusk et. al. [5] presented algorithms that allow automatic selection of features for region and tile similarity searches applying relevance feedback. Samal et. al. [6] proposed an RSI descriptor, called SIMR (Satellite Image Matching and Retrieval). SIMR computes spectral and spatial attributes of the images using a hierarchical representation. A unique aspect of this descriptor are the couples of second-level spatial autocorrelation with quad tree structure.

There is a large number of image descriptors proposed in the literature for CBIR that can be useful to classify and recognize RSI regions. Using descriptors, systems can calculate how similar regions of an image are when compared to a spectral or texture pattern in which users are interested. This information can, therefore, be used to classify the whole image. Santos et. al. [7] presented a semi-automatic method to vectorize regions from remote sensing images using relevance feedback based on genetic programming (GP) combining image descriptors. The solution consists in using image descriptors to encode texture and spectral features from the images, applying relevance feedback based on GP to combine these features with information obtained from the users interactions and, finally, segment the image. At the end, segmented image (raster) is converted into a vector representation.

Descriptors effectiveness can vary from one application to another. This fact shows the importance of evaluating descriptors considering specific applications. A comparative study of color descriptors for Web image retrieval is presented in [8]. However, to the best of our knowledge no study has been conducted to evaluate the performance and effectiveness of image descriptors in RSI retrieval and classification tasks.

This paper presents an evaluation of image descriptors for RSI retrieval and classification. Seven descriptors that encode texture information and twelve color descriptors that can be used to encode spectral information were selected. We highlight the main characteristics and perform experiments to evaluate the effectiveness of these descriptors in retrieval sessions and classification tasks.

2 IMAGE DESCRIPTORS

The descriptors chosen for the evaluation are important descriptors from the literature and recently proposed descriptors. This section brief introduces each of the them.

2.1 Color descriptors

The color descriptors evaluated in this work are: GCH [9], CGCH [10], LCH [9], CCV [11], ACC [12], JAC [13], BIC [14], CBC [15], Color Bitmap [16], CSD [17], CW-HSV [18] and CM [19].

Global Color Histogram (GCH) [9] is the most popular color descriptor from the literature. It is commonly used as baseline for comparisons of color descriptors. Its extraction algorithm quantizes the color space uniformly and scans the image counting the number of pixels that belong to each bin. The feature vector size depends on the quantization used. In our experiments, the RGB color space was divided into 64 bins, leading to a feature vector with 64 values. The distance function used was the L1 distance.

Cumulative Global Color Histogram (CGCH) [10] is also very popular in the literature and is very similar to the GCH descriptor. The main difference in the extraction algorithm is that the value of each bin is cumulated in the next bin. This makes the last bin have the sum of all the previous bins plus the actual bin. In our experiments the color space was quantized into 64 bins and L1 distance function was used.

Local Color Histogram (LCH) [9] is one of the most popular descriptors that is based on fixed size regions to describe image properties. Its extraction algorithm splits the image into fixed size regions and computes a color histogram for each region. After that, the histograms of each region are concatenated to compose one single histogram. The implemented version splitted the image in 16 regions (4x4 grid) and quantized the RGB color space into 64 bins. This generated feature vectors with 1024 values. The L1 distance function was used.

Color Coherence Vector (CCV) [11] is also very popular in the literature. Its extraction algorithm classifies the image pixels in coherent or incoherent pixels. This classification considers if the pixel belongs or not to a region with similar colors, called coherent region. After the classification, two color histograms are computed: one for coherent pixels and other for incoherent pixels. Both histograms are concatenated to compose the feature vector. In our experiments, the RGB color space was quantized into 64 bins and L1 distance function was used.

Color Autocorrelogram (ACC) [12] maps the spatial information of colors by pixels correlations in different distances. The autocorrelogram computes the probability of finding in the image two pixels with color C in a distance d from each other. After the autocorrelogram computation, there are m probability values for each distance d considered, where m stands for the number of colors in the quantized space. The implemented version quantized the RGB color space into 64 bins and considered 4 distance values (1, 3, 5, and 7). L1 distance function was used.

Joint Auto-Correlogram (JAC) [13] follows the same principle used by ACC. However, its extraction algorithm computes the autocorrelogram for more than one image property. The properties considered are: color, gradient magnitude, *rank* and *texturedness*. Color is extracted in RGB color space and the other properties are extracted from the gray level image. The joint autocorrelogram indicates, for each distance considered, the probability of simultaneously occurring the four properties considered. The implemented version used the HSV color space quantized into 64 bins, 5 bins for the other three properties, a 5x5 pixels neighborhood and 4 distance values (1, 3, 5, and 7). L1 distance function was used.

Border/Interior Pixel Classification (BIC) [14] classifies image pixels as *border* or *interior* pixels. If a pixel has the same color in the quantized space as their four neighbors (above, below, right, and left) it is classified as *interior*. Otherwise, it is classified as *border*. After the classification, two color histograms are computed: one for the interior pixels and other for the border pixels. The two histograms are concatenated and stored into the feature vector. The implemented version quantized the RGB color space in 64 bins. The *dLog* distance function was used.

Color-Based Clustering (CBC) [15] is a method for feature extraction based on image segmentation . The method decomposes the image into disjoint connected components. Each region has a minimum size and a maximum color difference. A region is defined by its average color in the CIE Lab color space, by its horizontal and vertical center, and by its size in relation to the image size. The distance function is a combination of L2 distance and *Integrated Region Matching (IRM)* functions.

Color Bitmap descriptor [16] analyzes image color properties globally and locally. Its extraction algorithm computes the mean and the standard deviation of each of the R, G, and B channels independently. After that, the image is split into m blocks and the mean of each block is computed for each channel. If the block mean is greater than the image mean, the correspondent feature vector position receives 1; otherwise, it receives 0. The implemented version used 100 blocks. The distance was computed in two steps: L2 function for the mean and standard deviation values; and Hamming distance for the binary values.

Color Structure (CSD) [17] is one of the color descriptors used in the MPEG-7 standard. The CSD extraction algorithm uses the HMMD (hue, max, min, diff) color space and scans the image with a 8x8 pixels structuring element. A histogram $h(m)$ is incremented if the color m is inside the structuring element, where m varies from 0 to $M-1$ and M is the color space quantization. The implemented version quantized the space in 184 bins as suggested in [17] and used the L1 distance function. According to previous comparisons among the MPEG-7 color descriptors [20, 21], the CSD achieved the best performance.

Color Wavelet HSV (CW-HSV) [18] considers image color properties in wavelet domain. Its extraction algorithm uses the HSV color space quantized into 64 bins and computes a global color histogram for the image. After that, the Haar wavelet coefficients are hierarchically computed. This is done recursively by dividing the histogram in the middle: if the sum of the values from the first half are greater than the sum of the values from the second half, the correspondent feature vector position receives 1; otherwise, 0. The process is repeated until the last possible level of division, what leads to 63 bits in the feature vector.

Chromaticity Moments (CM) descriptor [19] characterizes the image by chromaticity values. Its extraction algorithm first converts the image to the CIE XYZ color space. The chromaticity values (x,y) are computed as $x = \frac{X}{X+Y+Z}$ and $y = \frac{Y}{X+Y+Z}$. After that, two features are computed: the *trace*, that indicates the presence or not of each (x,y) value, and the histogram of chromaticities. The trace and the histogram are used to define the chromaticity moments. In the implemented version 6 moments were used, leading to 12 values in the feature vector. The distance function cumulates the modular differences between the corresponding moments.

3 TEXTURE DESCRIPTORS

The texture descriptors evaluated in this work are: LBP [22], HTD [23], SID [24], CCOM [25], Unser [26], QCCH [27], and LAS [28].

Local Binary Pattern (LBP) [22] is a simple texture descriptor that is invariant to rotation and variations in the gray scale values. Its extraction algorithm defines a window with radio R and a quantity of neighbors P and scans the image counting the quantity of positive and negative variations between the gray values of the neighbor pixels and the central pixel of the window. For gray scale invariance, only the signal of the variation is considered, being 1 for positive and 0 for negative variation. After that, the quantity of 0/1 and 1/0 transitions are computed, what guarantees the rotation invariance. If the quantity of transitions is less than 2, the LBP value for that window position is equal to the quantity of 1 signals in the neighborhood. Otherwise, the LBP value is $P+1$. After all the image is scanned, a histogram of LBP values is computed. In our experiments $R = 1$ and $P = 8$ values were used. The distance function used was the L1 distance.

Homogeneous Texture Descriptor (HTD) [23] is one of the texture descriptors from the MPEG-7 standard. Its extraction algorithm applies a set of filters sensitive to different scales and orientations. The output of each filter is an image from which the average and standard deviation values are computed. The most common filters used are Gabor filters. In the implemented version, Gabor filters sensitive to 4 scales and 6 orientations were used, leading to a feature vector with 48 values. The distance function computes the difference between each correspondent average and standard deviation values.

Invariant Steerable Pyramid Decomposition (SID) has extraction algorithm similar to the HTD descriptor. The image is processed by a set of filters sensitive to different scales and orientations. In SID, the image is first decomposed into two sub-bands by a high-pass filter and a low-pass filter. Next, the low-pass sub-band is decomposed recursively into K sub-bands by band-pass filters and into one sub-band by a low-pass filter. Each recursive step captures different directional information about each scale. The mean and standard deviation of each sub-band are used as feature vector values. The invariance to scale and orientation is obtained by applying circular shifts in the feature vector. The implemented version uses 2 scales and 4 orientations, leading to a feature vector with 16 values. The distance function computes the difference between each corresponding mean and standard deviation values.

Color Co-Occurrence Matrix (CCOM) [25] is a variation of Gray Level Co-Occurrence Matrix (one of the most common approaches for texture analysis and classification). CCOM extracts the feature vector by first quantizing the color space and then scanning the image to compute the co-occurrence matrix $W(c_p, c_q, d)$. For each pair of image pixels p, q with distance d between themselves, $W(c_p, c_q, d)$ is incremented by one, where c_p is the color of pixel p in the quantized space, c_q is the color of pixel q in the quantized space, and d is the distance between them. The feature vector stores the positive values of the matrix that are below a superior threshold, leading to a variable size feature vector. The implemented version quantized the RGB color space into 216 bins and used d equal to 1. The distance function computes the differences between the corresponding W values.

Unser descriptor [26] was proposed aiming to reduce the complexity of co-occurrence matrices while keeping good effectiveness. Its extraction algorithm computes a histogram of sums H_{sum} and a histogram of differences H_{dif}. The image is scanned and, for each angle a and distance d defined, the histogram of sums is incremented considering the sum and the histogram of differences is incremented considering the difference between the values of two neighboring pixels. As with gray level co-occurrence matrices, measures like energy, contrast, and entropy can be extracted from the histograms. In our experiments 256 gray levels and 4 angles were used (0°, 45°, 90° e 135°), as well as distance d equal to 1.5 and eight different measures were extracted from the histograms. The feature vector was composed by 32 values. The L1 distance function was used.

Quantized Compound Change Histogram (QCCH) descriptor [27] characterizes texture information using the relation between pixels and their neighbors. It generates a representation invariant to rotation and translation. Its extraction algorithm scans the image with a squared window. The average gray value of the window is computed for each position in the image. Then, four variation rates are computed considering the average gray values and the horizontal, vertical, diagonal and anti-diagonal directions. The average of these four variations is computed for each window position and they are quantized into 40 bins. A histogram of these values is computed. L1 distance function was used.

Local Activity Spectrum (LAS) descriptor [28] captures textures spatial activity in four different directions separately: horizontal, vertical, diagonal, and anti-diagonal. The four activity measures are computed for a pixel (i, j) by considering the values of neighboring in the four directions. The values obtained are used to compute a histogram that is called *local activity spectrum*. Each component g_i is quantized independently. In our experiments each component was non-uniformly quantized into 4 bins, leading to a histogram with 256 bins. The distance is computed by L1 function.

4 EXPERIMENTS

This section presents the databases used in the experiments and the measures used to evaluate the descriptors.

4.1 Image databases

Two image databases were created to evaluate image descriptors based on distinct RSIs. One of them can be classified as "easy recognition" (pasture image) while the other as "hard recognition" (coffee image). Information about the used RSIs is showed in Table 1.

In the experiments, one image is represented by a tile from the original RSI. The size of the tile was fixed according to the common extension value of a *region of interest*. Coffee crops are normally in small parcels on the same farm. We defined that 75×75 *meters* is a good value to the size of the partition. For pasture parcels, that are larger, the chosen value was 400×400 *meters*. The dimension of partitions are fixed in the experiments. We used 30×30 *pixels* to partition the coffee image and 20×20 *pixels* for the pasture image. The number of partitions for the pasture and coffee images was 5980 and 6400, respectively.

A "mask" containing all regions of interest from the RSIs used in the experiments was used to

Table 1: Remote Sensing Images used in the experiments.

	Image1	**Image2**
Region of interest	pasture	coffee
Terrain	plain	mountainous
Satelite	CBERS	SPOT
Spatial resolution	20 meters	2,5 meters
Bands composition	R-IR-G (342)	IR-NIR-R (342)
Acquisition date	08–20–2005	08–29–2005
Location	Laranja Azeda Basin, MS	Monte Santo County, MG
Dimensions (px)	1310×1842	2400×2400

know the class of each tile. A "mask" is a binary image where value 1 represents pixels of regions of interest. The "masks" used in our experiments were classified manually by agricultural specialists.

4.2 Evaluation measures

The main objective of the experiments was to evaluate and compare the descriptors considering effectiveness issues. For this purpose, we configured two experiments: retrieval effectiveness evaluation and overall accuracy classification.

To evaluate retrieval effectiveness, Precision $\times$ Recall curves were used. *Precision* quantifies the percentage of relevant images present in the retrieved results. *Recall* is a measure that represents the percentage of the relevant images that are retrieved. A Precision $\times$ Recall curve indicates the variation in Precision values as the rate of relevant images from the database (Recall) changes. Intuitively, the higher the curve, the better the effectiveness. The Precision and Recall curves were computed based on the average values obtained for each query image in each database. We used 340 and 100 queries in the Pasture and Coffee image sets respectively for all the color and texture descriptors presented in Section 2.

To compute the overall accuracy of each descriptor we implemented a variation of K-Nearest Neighboor (KNN) classifier. First of all, a set of tiles from the database was randomly selected to be used as training set. The set, corresponding to 10% of the database size, is composed by relevants and non-relevants samples in the same proportion found in the full database. To classify one image (tile), each descriptor evaluated was used to compute the distance between the given tile and all the training set tiles. Based on the descriptor distances, the training set is ranked and the first K tiles are weighted inversely proportional to their position in the rank. Finally, the sum of the tiles' weights for each class (relevant or non-relevant) is computed. The greater sum indicates the class of the input tile. To test the classification effectiveness of the descriptors 100 tiles were used for each RSI.

4.3 Results

Figures 1, 2, 3, and 4 show the Precision $\times$ Recall curves for color and texture descriptors in the databases used.

From Figure 1 we can see that good descriptors considering retrieval effectiveness are JAC, Color Bitmap, and ACC.

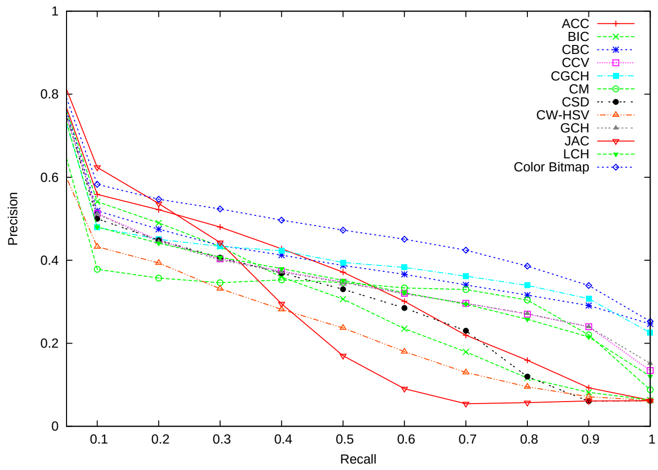

Figure 1: Precision × Recall curves for color descriptors in the Pasture Image Set.

From Figure 2, it is possible to see that JAC presents the highest Precision values even for small values of Recall and for Recall equal to 1.

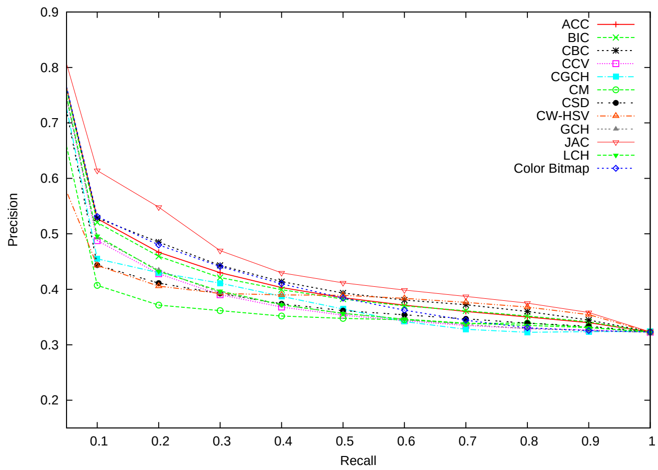

Figure 2: Precision × Recall curves for color descriptors in the Coffee Image Set.

Analyzing Figure 3 it is possible to notice that SID has the highest Precision values for all values of Recall.

Considering curves for the Coffee database in Figure 4, it is possible to see that the descriptors present very similar Precision values and these values are near 32% when Recall reaches 10%.

After analyzing the curves for color and texture descriptors it is possible to say that color descriptors are slightly better than texture descriptors for the databases used. For example, in the Pasture database, for Recall equal to 10%, the highest Precision value for color descriptors is around

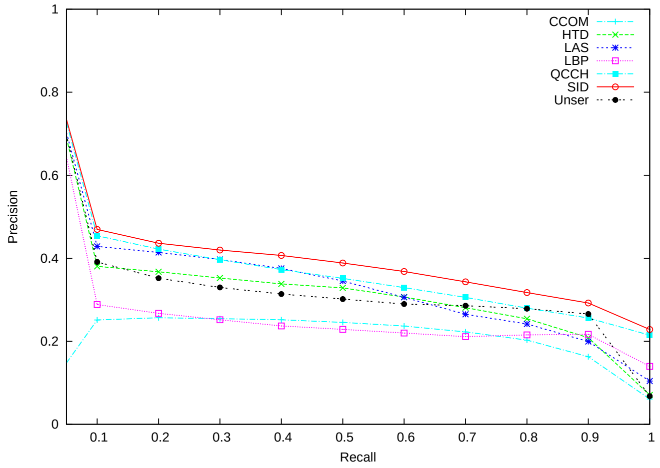

Figure 3: Precision × Recall curves for texture descriptors in the Pasture Image Set.

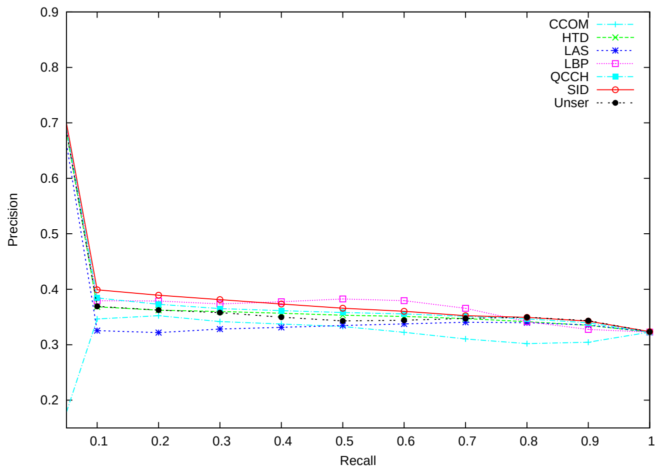

Figure 4: Precision × Recall curves for texture descriptors in the Coffee Image Set.

62% (JAC) and for texture descriptors is near 47%. For Recall equal to 1, color descriptors achieve Precision of 25% (Color Bitmap) and texture descriptors achieve almost 23% . For the Coffee database, it is possible to notice that, for Recall equal to 10%, the highest curve of a color descriptor reaches 61% (JAC) while the highest curve of a texture descriptor reaches almost 40% (SID). For Recall equal to 1, there is almost no difference in the Precision values.

According to the results for the coffee database presented in Figure 5, its is observed that some descriptors achieved high overall accuracy values. The color descriptors BIC, ACC, CBC, Color Bitmap, and JAC were the best ones reaching more than 60% of overall accuracy for any k. JAC produced the highest accuracy values, being the only one with values over 70% (72% for k=1, 79% for k=3, and 73% for k=7 and k=10). In relation to the texture descriptors, QCCH, SID, and

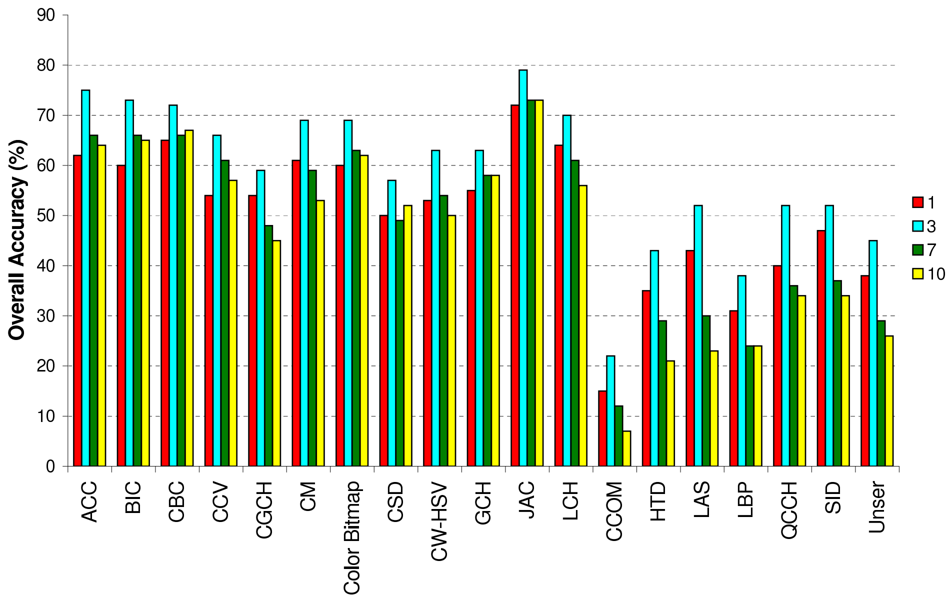

Figure 5: Overall accuracy classification of each descriptor for Coffee using KNN 1, 3, 7 and 10.

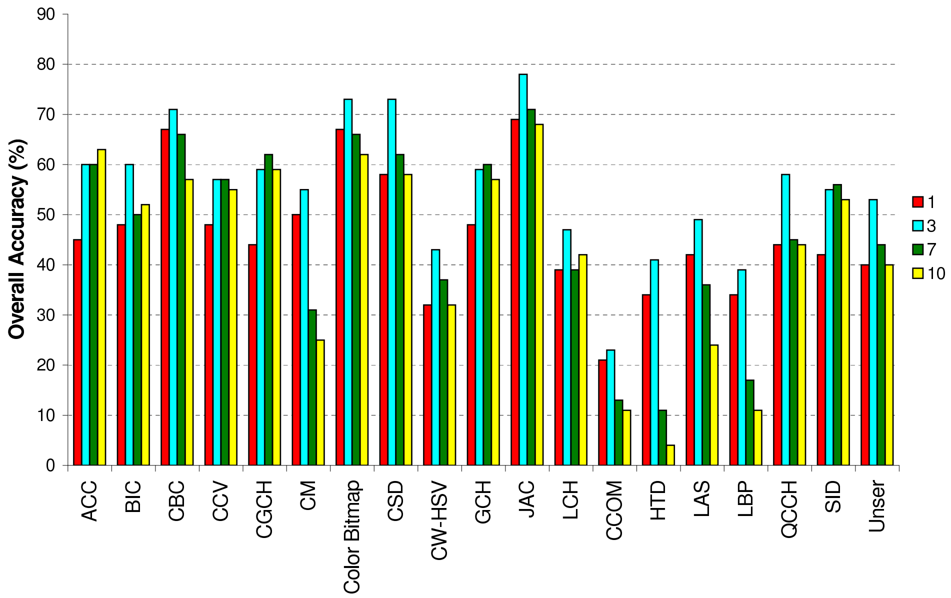

Figure 6: Overall accuracy classification of each descriptor for pasture KNN 1, 3, 7 and 10.

LAS yielded the highest accuracy values, 52% for k=3. For k values different than 3, the texture descriptors presented accuracy below 48%. The CCOM descriptor did not reach 25% of accuracy in any of the experiments in the coffee database.

According to the results for the pasture database presented in Figure 6 we can see that some descriptors yielded good accuracy values. The color descriptors JAC, Color Bitmap, and CBC reached near or more than 60% of overall accuracy. JAC descriptor was again the descriptor with highest accuracy value, reaching 78% for k=3 and being over 65% for all k values. The texture descriptors yielded lower accuracy values in relation to the majority of color descriptors. QCCH, SID and Unser were the only texture descriptors to reach accuracy above 50%. For k=3, QCCH reached 58% of accuracy, SID 55% and Unser 53%. CCOM descriptor reached the lowest accuracy values, being below 25% for all k values.

Considering the accuracy values in both image databases, we can point JAC as the best color descriptor. However, JAC generates big feature vectors and so, it is slower to compare them. If storage and time requirements are not critical, JAC is the best choice. Other descriptors with near effectiveness are CBC and Color Bitmap. CBC has complex extraction and distance function. Color Bitmap is the best choice among the color descriptors, which balance simple algorithms and good effectiveness. Amongst the texture descriptors, QCCH and SID reached the highest accuracy values, being SID more computationally complex than QCCH for features extraction.

References

[1] R. Showengerdt, *Techniques for Image Processing and Classification in Remote Sensing*. New York: Academic Press, 1983.

[2] D.-K. Mo, H. Lin, J. Li, H. Sun, and Y.-J. Xiong, “Design and implementation of a high spatial resolution remote sensing image intelligent interpretation system,” *Data Science Journal*, vol. 6, pp. S445–S452, 2007.

[3] I. Yildirim, O. K. Ersoy, and B. Yazgan, “Improvement of classification accuracy in remote sensing using morphological filter,” *Advances in Space Research*, 2005.

[4] D. hyung Kim, S. gyu Jeong, and C. hwa Park, “Comparison of three land cover classification algorithms - isodata, sma, and som - for the monitoring of north korea with modis multi-temporal data,” *Korean Journal of Remote Sensing*, 2007.

[5] C. Tusk, K. Koperski, S. Aksoy, and G. Marchisio, “Automated feature selection through relevance feedback,” in *Geoscience and Remote Sensing Symposium, 2003. IGARSS ’03. Proceedings. 2003 IEEE International*, vol. 6, July 2003, pp. 3691–3693 vol.6.

[6] A. Samal, S. Bhatia, P. Vadlamani, and D. Marx, “Searching satellite imagery with integrated measures,” *Pattern Recogn.*, vol. 42, no. 11, pp. 2502–2513, 2009.

[7] J. A. dos Santos, R. A. Lamparelli, and R. da Silva Torres, “Using relevance feedback for classifying remote sensing images,” in *Proceedings of Brazilian Remote Sensing Symposium*, Natal, RN, Brazil, April 2009, pp. 7909–7916.

[8] O. A. B. Penatti and R. d. S. Torres, “Color descriptors for web image retrieval: A comparative study,” *XXI Brazilian Symposium on Computer Graphics and Image Processing*, pp. 163–170, October 2008.

[9] M. J. Swain and D. H. Ballard, "Color indexing," *International Journal of Computer Vision*, vol. 7, no. 1, pp. 11–32, 1991.

[10] M. A. Stricker and M. Orengo, "Similarity of color images," in *Proc. SPIE Storage and Retrieval for Image and Video Databases III*, W. Niblack and R. C. Jain, Eds., vol. 2420, Mar. 1995, pp. 381–392.

[11] G. Pass, R. Zabih, and J. Miller, "Comparing images using color coherence vectors," in *Proceedings of the fourth ACM international conference on Multimedia*, New York, NY, USA, 1996, pp. 65–73.

[12] J. Huang, S. R. Kumar, M. Mitra, W. Zhu, and R. Zabih, "Image indexing using color correlograms," in *Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition*, Washington, DC, USA, 1997, p. 762.

[13] A. Williams and P. Yoon, "Content-based image retrieval using joint correlograms," *Multimedia Tools and Applications*, vol. 34, no. 2, pp. 239–248, 2007.

[14] R. de O. Stehling, M. A. Nascimento, and A. X. Falcão, "A compact and efficient image retrieval approach based on border/interior pixel classification," in *Proceedings of the eleventh international conference on Information and knowledge management*, New York, NY, USA, 2002, pp. 102–109.

[15] R. de O. Stehling, M. A. Nascimento, and A. X. Falcao, "An adaptive and efficient clustering-based approach for content-based image retrieval in image databases," in *Proceedings of the International Database Engineering & Applications Symposium*, Washington, DC, USA, 2001, pp. 356–365.

[16] T. Lu and C. Chang, "Color image retrieval technique based on color features and image bitmap," *Information Processing and Management*, vol. 43, no. 2, pp. 461–472, 2007.

[17] B. S. Manjunath, J.-R. Ohm, V. V. Vasudevan, and A. Yamada, "Color and texture descriptors," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 6, pp. 703–715, June 2001.

[18] A. Utenpattanant, O. Chitsobhuk, and A. Khawne, "Color descriptor for image retrieval in wavelet domain," *Eighth International Conference on Advanced Communication Technology*, vol. 1, pp. 818–821, 20-22 February 2006.

[19] G. Paschos, I. Radev, and N. Prabakar, "Image content-based retrieval using chromaticity moments," *IEEE Transactions on Knowledge and Data Engineering*, vol. 15, no. 5, pp. 1069–1072, 2003.

[20] J. Annesley, J. Orwell, and J.-P. Renno, "Evaluation of mpeg7 color descriptors for visual surveillance retrieval," *2nd Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance*, pp. 105–112, 15-16 October 2005.

[21] T. Ojala, M. Aittola, and E. Matinmikko, "Empirical evaluation of mpeg-7 xm color descriptors in content-based retrieval of semantic image categories," *International Conference on Pattern Recognition*, vol. 2, p. 21021, 2002.

[22] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, 2002.

[23] P. Wu, B. S. Manjunath, S. Newsam, and H. D. Shin, "A texture descriptor for browsing and similarity retrieval," *Signal Processing: Image Communication*, vol. 16, no. 1-2, pp. 33 – 43, 2000.

[24] J. A. M. Zegarra, N. J. Leite, and R. d. S. Torres, "Rotation-invariant and scale-invariant steerable pyramid decomposition for texture image retrieval," in *XX Brazilian Symposium on Computer Graphics and Image Processing*. Washington, DC, USA: IEEE Computer Society, 2007, pp. 121–128.

[25] V. Kovalev and S. Volmer, "Color co-occurence descriptors for querying-by-example," *Multi-Media Modeling*, vol. 0, pp. 32–38, October 1998.

[26] M. Unser, "Sum and difference histograms for texture classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 1, pp. 118–125, 1986.

[27] C. Huang and Q. Liu, "An orientation independent texture descriptor for image retrieval," *International Conference on Communications, Circuits and Systems*, pp. 772–776, July 2007.

[28] B. Tao and B. W. Dickinson, "Texture recognition and image retrieval using gradient indexing," *Journal of Visual Communication and Image Representation*, vol. 11, no. 3, pp. 327 – 342, 2000.