The use of Discrete Cosine Transform for satellite images segmentation and comparison to statistical metrics

Ana Carolina Quintão Siravenha[1]
Evaldo Golçalves Pelaes[1]

[1] Federal University of Pará – UFPa
Augusto Correa, 1 – 66075-110 – Belém – Pa, Brazil
{siravenha,pelaes}@ufpa.br

Abstract. In the last years a large amount of methods have been proposed to segment images, and it is show that this task is still an open problem. Remote sensing is one field on Digital Image Processing (DIP) that is becoming increasingly widespread. In this paper we describe a novel approach based on Discrete Cosine Transformation (DCT) and compare it with some approaches based on statistical metrics. Both, the DCT and statistical approaches, make a texture analysis based on components of a gray scale satellite image from QuickBird QB2 and Word View WV02 sensors. The texture analysis are based on the similarity property within a group of pixels that allows simple sub-blocks manipulation. The DCT is applied for the image and the AC energy component is calculated individually for square blocks. The segmentation rules of the proposed method are defined based on those values of energy. The statistical metrics carries out a local analysis of pixel intensities where the statistical information of pixels intensities are calculated in image split into $N \times N$ blocks. The statistical metrics adopted for comparison are the mean, standard deviation, median and variance. The segmentation of statistical metrics is done according a maximum likelihood classifier and the approaches evaluation are done by an entropy analysis. The results demonstrate the effectiveness of the proposed DCT-based method when compared to the other statistical metrics.

Keywords: Discrete Cosine Transformation, image processing, textural analysis .

1. Introduction

For a few decades image segmentation has been extensively studied in order to provide better results and nowadays is the most important task in digital image processing. According Gonzalez e Woods (2008) segmentation is the routine that subdivides an image into its constituent regions or objects.

With advances in remote sensing, the images coming from artificial satellites has been largely used to environmental monitoring and other various proposals (JOHANSEN et al., 2007). Traditional methods usually are based on region or edge detection. Region based algorithm ordinarily uses region growing or texture information (SHIH; CHENG, 2005; CHEN; CHEN; CHIEN, 2009) whilst edge detection based algorithm can use techniques as Dynamic Programing (DP), as did by Mohamed e Gader (1996). The use of traditional methods is improved by *a priori* knowledge about the segmented object, that in most applications is provided by experts interventions, in a non-automatic form (MCINERNEY; TERZOPOULOS, 1996).

Gamanya, Maeyer e Dapper (2007) developed an automated classification design based on region-merging segmentation technique that incorporates some objects properties (spectral and textural) to segmentation task. The experiments with Landsat and Aster images resulted in mutually exclusive classes with clear and unambiguous class definitions.

Time-series can be applied to mapping the annual cycle of semi-natural habitats and agricultural land cover. Lucas et al. (2007) used a rule-based classifier after segment a Landsat sensor image with eCognition Expert, achieving good representation of habitats and agricultural land, with system accuracies exceeding 80%.

Johansen et al. (2007) used Terrestrial Ecosystem Mapping (TEM) in high spatial resolution satellite image (from QuickBird) in order to discriminate structural stages of vegetation in riparian and adjacent ecosystems. Based on semi-variogram for calculation of gray-level co-occurrence matrix, they showed that the classification accuracy of system was significantly increased with inclusion of image texture information.

The texture analysis plays important role in DIP and can be applied in image segmentation or classification tasks because these experiments are intrinsically related with correlation among pixels (GONZALEZ; WOODS, 2008; CHEN; NIXON; THOMAS, 1995). The group of pixels in a satellite image must be undertaken in order to specify the texture characteristics of a class.

Beyond DCT technique be largely used for compression purposes, it also can be used to represent texture features. It can achieve a computational efficiency even compared to Discrete Wavelet Transformation (DWT) based segment system (PUN; ZHU, 2010).

In this paper we present a segmentation algorithm based on DCT properties. Our approach is based on the fact that the AC coefficients are related with region activity, and due this, the DCT is used to identify homogeneous regions on satellite images. In parallel, we developed segmentation approaches based on *mean*, *standard deviation*, *variance* and *median*. These statistical metrics take into account the correlation within a group of pixels (HAIHUI et al., 2010).

In paper sequel, we give reviews of DCT in Subsection 2.1, Mean (Subsection 2.2), Standard deviation (Subsection 2.3), Variance (Subsection 2.4) and Median (Subsection 2.5). We describe the methodology involved on algorithm development in Section 3. The experimental results and discussions about them are exposed in Section 4. Finally, the conclusions about the proposal are presented in Section 5.

2. Textural analysis

In this section we describe the techniques used to segment satellite images. A review of DCT will be made in order to introduces the reader to concepts used on proposal design. Comparative methods are also presented in the follow subsections.

2.1. Discrete Cosine Transform

The DCT is a technique that converts a signal into elementary frequency components. The energy is concentrated in a minor amount of coefficients, which becomes a big draw for image and video compression systems (WATSON, 1994).

Using only real data with even symmetry, the DCT method is closely related to the Discrete Fourier Transform (DFT). The one-dimensional DCT (1D-DCT) $X[k]$ of a sequence of N numbers $x[n]$ is defined by

$$X[k] = \sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} x[n] cos \frac{\pi}{N} k(n + \frac{1}{2}),$$

where $0 \leq k \leq N - 1$. Each coefficient of list $X[k]$ is the inner product of input sequence $x[n]$ and a basis vector, that are orthogonal and normalized.

The 1D-DCT is useful to processing one-dimensional signals, such as general waveforms. To processing two-dimensional signals, such as images, a two-dimensional transform is desired. Thus, for an N_1xN_2 array, the two-dimensional DCT (2D-DCT) is given by

$$X[k_1,k_2] = \frac{2}{\sqrt{N_1N_2}} \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} x[n_1,n_2] cos\frac{\pi}{N_1}k_1(n_1+\frac{1}{2}) cos\frac{\pi}{N_2}k_2(n_2+\frac{1}{2}).$$

Because DCT is a separable operation, the 2D-DCT can be replaced by a 1D-DCT applied separately in rows and columns. In case of 2D-DCT, the basis functions are N_1xN_2 matrices, where each one is the outer product of two of one-dimensional basis vectors.

For processing an image, the 2D-DCT commonly is applied separately into nxn blocks, called blocked DCT. Actually, the image is not split into blocks, since 2D-DCT is separable, one can partition each row into lists of length n, applying the DCT to them. The resulting lists are rejoin and the whole image is transposed for repeat the process (WATSON, 1994).

In a 2D transformed matrix X the element $X[0,0]$ corresponds to DC coefficient, in other words, is the zero frequency coefficient and represent the mean value of pixels in a block. The other elements with non-zero coefficients are called AC coefficients.

According Parseval theorem, the mean energy of 2D-DCT of a sequence with N_1xN_2 dimension is given by

$$E = \frac{1}{N_1N_2} \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} X^2[n_1,n_2]$$

As mentioned before, the DCT is applied into blocks independently. Thus, the AC energy is related with block activity in spatial domain, i.e., is related with fast or slow changes in pixel sequence, or in the presence of edge(s). The AC energy can be calculated by

$$E_{AC} = \frac{1}{N_1N_2} \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} X^2[n_1,n_2] - E_{DC}$$

where $E_{DC} = \frac{X^2(0,0)}{N_1N_2}$.

2.2. Mean

In statistic mean is defined as the expected value of a random variable. For a data set, the mean is characterized as the sum of values divided by the number of values of this data set.

The 2D-mean is defined by

$$\mu = \frac{1}{n_1n_2} \sum_{i=0}^{n_1} \sum_{j=0}^{n_2} x_{i,j}$$

2.3. Standard deviation

Standard deviation is a measure of dispersion from the mean. It is defined by the square root of data or population variance. The closer the mean is the dispersion, lower is the standard deviation. This measure is described as follow,

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)^2}$$

2.4. Variance

Variance is a descriptor of distribution, describes how far values lie from the main, as showed bellow. The advantage of this approach is the computational simplicity because are based on moments.

$$Var(X) = \sum_{1}^{ws}\sum_{1}^{ws}\frac{(X-\mu)^2}{N^2-1},$$

where ws is the size of window defined on algorithm beginning, X is the image window analyzed and N represents the images block dimension.

2.5. Median

A median is described as the numeric value separating the higher half of a sample from the lower half.

Of a finite list of numbers can be found by arranging all the observations from lowest value to highest value and picking the middle one. If there is an even number of observations, then there is no single middle value; the median is then usually defined to be the mean of two middle values

$$Median = x_{i1} + \left(\frac{(N_M/2) - N_{i-1}}{f_i}\right)(x_{i2} - x_{i1}),$$

where x_{i1} and x_{i2} is the minimum and maximum values of data set, respectively.

3. Proposed segmentation algorithm

For convenience the input images were resized to 2^nx2^n dimension and were converted to gray scale domain. Because our proposal did not include any color domain we ensure the use for a large amount of images.

The proposal was implemented in Matlab[1] and were used images available in Digital Globe[2] sample catalog. The images from QuickBird and WorldView satellite (QB2 and WV02 sensors, respectively) are presented in Figure 1. The images chosen had total cloud cover percentage less than or equal to 10%. The acquisition dates vary from 2004 to 2010.

Four classes for images segmentation were identified: Water -WT-, Rain forest -RF-, Cloud -CL- and Urban sprawl-US. These classes were visually identified and the eventual existence of bare soil was treated as US.

The DCT algorithm is based on AC energy and mean analysis of a 2D transformed images. The algorithm makes the segmentation based on estimation of input values and after a combination of thresholds of AC energy and mean values. To complete this process, the segmented image is then filtered by a statistical filter that is equivalent to dilate operation. In practice, each pixel of the input image is replaced by the maximum value among all pixels in the neighborhood of this pixel (MATHWORKS, 2010).

To statistical metrics, the sample blocks selected during algorithm beginning provides an approximated values that distinguish the classes. Then, three boundaries were created to define the classes separation. Once the boundaries are dependent of sample regions and these values do

[1]http://www.mathworks.com/
[2]https://www.digitalglobe.com/

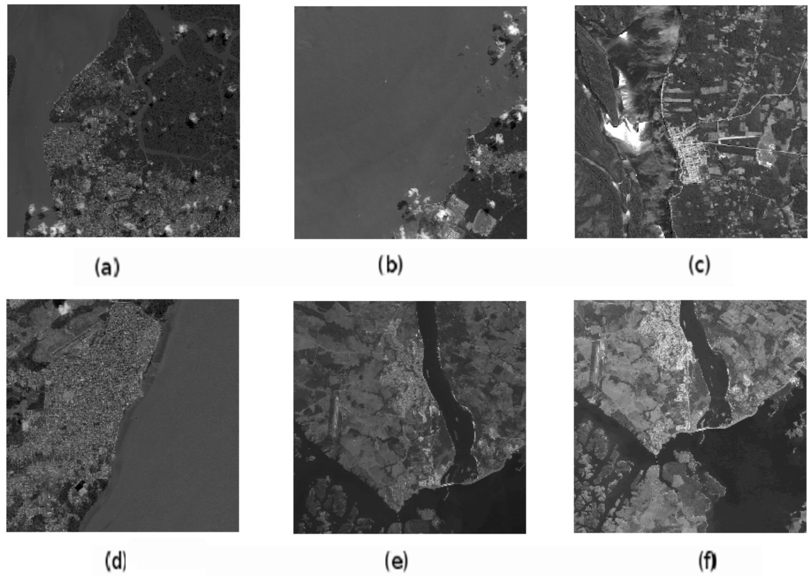

Figure 1: Satellite images from: (a) QB02 sensor from metropolitan region of Belém-Pa. (b) QB02 sensor from region of *Capim* bay in state of Pará. (c) QB02 sensor from a region of *Tocantins* river in state of Pará. (d) QB02 sensor from city of Macapá-Ap. (e) QB02 sensor from city of Tucuruí-Pa. (f) WV02 sensor from a region of artificial lake of Tucuruí-Pa.

not provide boundaries that serve to different images, the boundaries attribution was empirically defined according each image tested.

The approach evaluation was done by a entropy analysis. Entropy is a measure of randomness of an amount of information that can characterize the texture of an image (MATHWORKS, 2010). Low entropy images have little contrast and large number of pixels with similar or the same value. In contrast, high entropy image revels great contrast of pixels. The entropy of an image equals to zero represents a perfectly flat color image. The entropy (H) is defined by the relation:

$$H = -\sum_{i=1}^{n} p(x_i) \log_b p(x_i)$$

Since the smaller the entropy more homogeneous is the image, it is preferable that the entropy of a segmented image has lower values, making sure that falls within acceptable levels, to avoid flat color or nearly flat color images.

In addition to the comparisons among statistical metrics, in order to complete the validation of DCT algorithm, will be held a comparison of results by DCT-based method and the achievements of SPRING software (CAMARA et al., 1996).

4. Results and discussions

Since entropy can be used to represent an image texture feature and lowest values of entropy defines regions of greater homogeneity, it is expected that, when an image is segmented, the resulting entropy presents a decay proportional to manipulations applied into an image. In this context, Figure 2 presents the entropy resultant of approaches described previously.

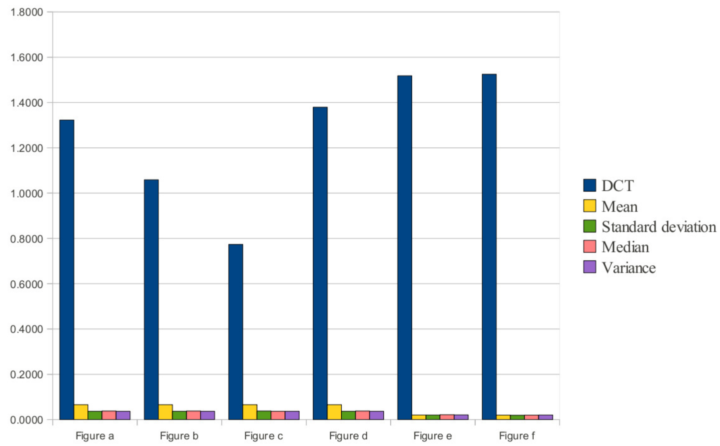

Figure 2: Comparison of entropy values among statistical metrics and DCT-based method.

Figure 2 shows that the images segmented by statistical metrics suffered an over-segmentation. All values of statistical metrics entropy are located in a region bellow of 0.2, while the DCT method provides entropy values located, generally, higher than 1.00. The exception is found in Figure 1 (c), whose entropy is 0.7735. This is caused by regions of confusion showed in this figure. These confusions occurs, mainly, because the informations about the RF and WT classes have similar features of AC energy and mean, resulting in a poor separation between these classes. This value of entropy does not shows that occurs an over-segmentation, but just reflects an image particularity.

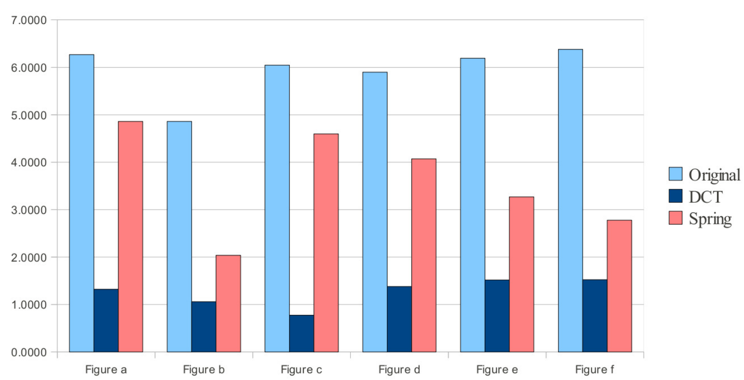

Figure 3: Entropy comparison among achievements of DCT-based method and SPRING software.

The comparison of method based on DCT with the segmentation provided by SPRING software is shown in Figure 3. To SPRING segmentation, were modified the similarity parameters and minimal size of regions in order to produce images with match conditions with DCT results. Thus, was observed that SPRING produces under-segmentation results when some restrictions are imposed to it. It is important stress that the SPRING is a state-of-the-art remote sensing image processing system, and this poor results can be caused by the restrictions imposed to it. However, the proposed method based on DCT shows that with little amount of informations it can produce acceptable results.

By visually inspecting the results, we have concluded that the DCT method gives really good results and its accuracy can be noted in Figure 4.

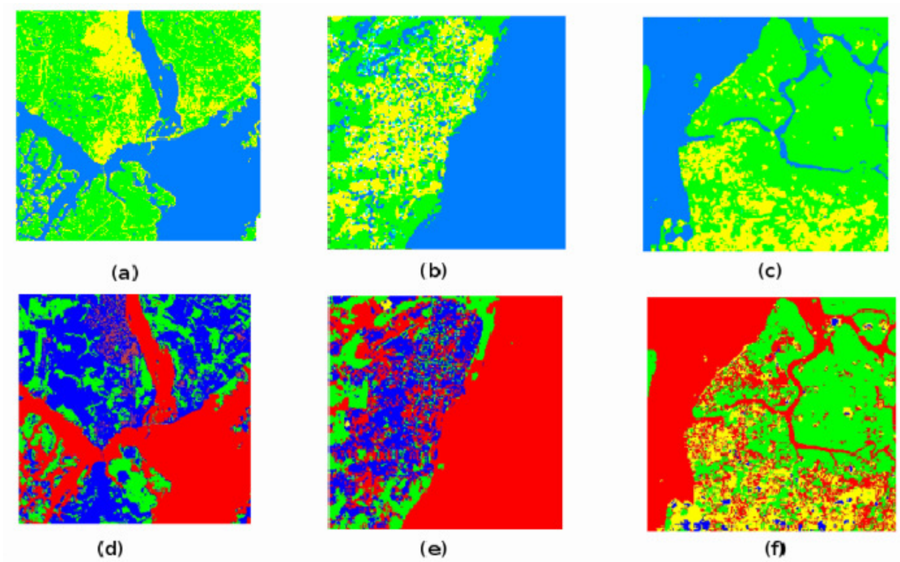

Figure 4: Achievements of: (a) to (c) DCT-based method. (d) to (f) SPRING software.

5. Conclusion

The proposed method using DCT showed to be an efficient segmentation approach. The results demonstrates the efficiency of proposed method in comparison with statistical metrics and the SPRING software achievements. Furthermore, it does not require a lot of information about the image that will be submitted. In terms of execution time, the approaches do not exceed a little dozens of seconds, which an average time of 13 seconds to DCT method. The tested images do not exceed the 2^9x2^9 dimension, but it is expected that the algorithm does not become expensive in term of execution time, since the most of satellite images are commonly bigger than that. In terms of computational costs, the use of DCT is not expensive, so we believe that the DCT method represent a novel approach to segmentation task, covering a lot of images with a wider range of randomness.

The future works are focused on DCT method enhancement. It is expected that the thresholds being automatically selected according to the classes features in an image. In addition, other thresholds may be introduced to improve the method accuracy, beyond the development of an unsupervised classifier to complete the image analysis.

6. Acknowledgements

This work was supported by Fundação de Amparo à Pesquisa do Estado do Pará, grant no. 021/2008, *Desenvolvimento de um sistema de interpretação de imagens de satélite baseado em modelos híbridos*.

References

CAMARA, G. et al. **SPRING: Integrating Remote Sensing And Gis By Object-Oriented Data Modelling**. In: *SPRING Programming Manual*. [S.l.: s.n.], 1996. São José dos Campos, INPE.

CHEN, Y.-L.; CHEN, T.-W.; CHIEN, S.-Y. Fast texture feature extraction method based on segmentation for image retrieval. In: ***IEEE 13th International Symposium on Consumer Electronics***. [S.l.: s.n.], 2009. p. 941–942.

CHEN, Y. Q.; NIXON, M. S.; THOMAS, D. **Texture classification using statistical geometrical features**. In: *Pattern Recognition*. [S.l.: s.n.], 1995. p. 537–552.

GAMANYA, R.; MAEYER, P. D.; DAPPER, M. D. An automated satellite image classification design using object-oriented segmentation algorithms: A move towards standardization. In: ***Expert Systems with Applications***. [S.l.: s.n.], 2007. v. 32, p. 616–624.

GONZALEZ, R. C.; WOODS, R. E. ***Digital Image Processing***. [S.l.]: Addison-Wesley Publishing Company, 2008.

HAIHUI, W. et al. Images segmentation method on comparison of feature extraction techniques. In: ***2nd International Workshop on Telecomunication***. [S.l.: s.n.], 2010. p. 1–4.

JOHANSEN, K. et al. Application of high spatial resolution satellite imagery for riparian and forest ecosystem classification. In: ***Remote Sensing of Environment***. [S.l.: s.n.], 2007. v. 110, p. 29–44.

LUCAS, R. et al. Rule-based classification of multi-temporal satellite imagery for habitat and agricultural land cover mapping. In: ***ISPRS Journal of Photogrammetry and Remote Sensing***. [S.l.: s.n.], 2007. v. 62, p. 165–185.

MATHWORKS. ***Image Processing Toolbox***. 2010. Disponível em: <http://www.mathworks.com/help/toolbox/images/ref/ordfilt2.html>. Acesso em: 17 de Novembro de 2010.

MCINERNEY, T.; TERZOPOULOS, D. Deformable modelsin medical image analysis: A survey. In: ***Medical Image Analysis***. [S.l.: s.n.], 1996. v. 1, p. 91–108.

MOHAMED, M.; GADER, P. **Handwritten word recognition using segmentation-free hidden Markov modeling and segmentation-based dynamic programming techniques**. In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on*. [S.l.: s.n.], 1996. v. 18, p. 548–554.

PUN, C.-M.; ZHU, H.-M. Image segmentation using discrete cosine texture feature. In: ***International Journal of Computers***. [S.l.: s.n.], 2010. v. 4(1), p. 19–26.

SHIH, F. Y.; CHENG, S. Automatic seeded region growing for color image segmentation. In: ***Image and Vision Computing***. [S.l.: s.n.], 2005. v. 23, p. 877–886.

WATSON, A. B. **Image compressing using the discrete cosine transform**. In: *Mathematica Journal*. [S.l.: s.n.], 1994. v. 1, p. 81–88.