Mapping and characterization of urban forest in Mexico City

Alejandra A. López-Caloca*a, b, Franz Mora*a and Boris Escalante-Ramírez*a, b

a Centro de Investigación en Geografía y Geomática “Ing. Jorge L. Tamayo”, A.C.
Blvd. Cataratas No.3, Jardines del Pedregal, México, D.F., 01900
b Universidad Nacional Autonoma de México, División de Posgrado de Ingeniería.
Apdo. Postal 70-256, México, D.F., 04510.

ABSTRACT

Mapping and characterization of forest and vegetation are particularly challenging in urban areas. High resolution imagery is needed for mapping and characterization purposes, due to the areal extent of urban forests, parks and recreational areas. Fusion techniques of panchromatic (1m resolution) and multiband (4m resolution) IKONOS data were used for mapping and characterization of land covering characteristics of urban green areas, allowing the identification of parks, tree areas and fields with a minimal mapping unit of 160 m². Techniques, that integrate the fine details of the input data into the fused image, are used. Experimental results for different image fusion methods (Laplacian, Gradient pyramids, Principal Component Analysis and Wavelet transform) are also demonstrated in order to improve spatial resolution.

Classification of urban areas, mapped with fused data, results in higher accuracies than when using a multiband approach with 4 m data alone. Furthermore, high spatial resolution data permitted to obtain new areal extents of green areas of the city, giving a better estimate of international indicators for a suitable green areas policy. Vegetation indexes derived from red and near infrared data IKONOS are used to evaluate vegetation conditions, which, along with their distribution, location and urban context, resulted in better indicators of green areas.

Keywords: urban green area mapping, image fusion, pyramidal schemes, wavelets, vegetation index.

1. INTRODUCTION

The interest in the mapping urban green areas has grown due to urban environmental study and urban planning. Policy-makers and planners have become aware of the environmental and economical benefits that trees and green areas provide in terms of air quality and energy conservation. Planners today need tools to map urban forests and to incorporate new green urban infrastructure. Green areas are a fundamental component in the structure and urbanized operation of the landscape. They not only provide recreational and relaxation spaces for inhabitants of cities, but also constitute a cooling mechanism within the urban landscape. Trees and urban parks provide shade to pedestrians and buildings. Traditionally, the extraction and mapping of urban green areas has been made through visual interpretation and field work.

The mapping and characterization of the efforts components vegetation cover in urban areas within Mexico City are made through the use of information of high resolution images (IKONOS). In this study we used IKONOS 1m and 4m multispectral resolution. By means of fusion techniques it is possible to improve the spatial resolution of the IKONOS images. Image fusion is used to combine two or more images to create a new image8. The essential problem when fusing images is keeping relevant patterns of all source images, that is to say important details in the images must be preserved in the resulting image. During the fusion process spurious elements that can interfere in subsequent analyses must not introduced.

The objective of image fusion is to generate synthetic images with higher resolution that attempt to preserve the radiometric characteristics of the original multispectral data. This is very important for vegetation analysis. It is desirable that any procedure to fuse high resolution panchromatic data with low resolution multispectral data preserves the original spectral characteristic as much as possible. The procedure should be optimal, in the sense that, only the additional spatial information available in the higher resolution data is blended into the multispectral bands.

* alopez@centrogeo.org.mx; (+52) 5135-2635; fax(+52) 5135 2638
In the present paper, the efficiency of different techniques to merge multiresolution and multispectral data has been tested. The panchromatic band and the multispectral bands of IKONOS have been merged into new images that combine the spectral characteristics of IKONOS data with the relatively high spatial resolution of the panchromatic band.

The fusion techniques studied here are the Laplacian pyramid, Gradient pyramid, Wavelet transform and the standard method of Principal Component Analysis. The first three techniques are multiresolution methods that consist of decomposing an image into sub-images, that representing the spatial contents of the original image at different levels of resolution. Consequently, each multispectral image and panchromatic image were decomposed in a pyramidal structure. In order combine the information carried by each image, some coefficients are selected from the source image pyramids. The rule used to select coefficients is based on match and salience measures. Finally, the composite image is recovered through an inverse pyramidal transformation.

From the resulting fusioned images, a vegetation index, derived from red and near infrared IKONOS data, is used to evaluate vegetation conditions for urban green areas, and used as a quantitative measure of photosynthetic activity to characterize vegetation condition.

The result and conclusions of this paper may be useful to improve the spatial resolution and to enhance urban areas imaging.

2. IMAGE DATA AND METHODOLOGY

The mapping and characterization of the vegetation cover in Mexico City include three main stages: image pre-processing, the study and application of fusion techniques for the identification of green areas; and image classification for the identification of green areas.

2.1 Pre-processing

The pre-processing stage includes: geometrical and radiometric corrections. The 1m and 4m IKONOS images have different pixel sizes, so the lower resolution image is resampled to match the higher resolution size by means of cubic interpolation. The next step involves a reference image to make the coregistration between the panchromatic and the multispectral images. Geometric correction for both types of images has to be done carefully and with sufficient ground control points in order to avoid mis-registration. Due to the terrain conditions of Mexico City, image orthorectification is the first important task in spatial analysis. The approach used here considers an elevation terrain model as part of the correction. With a digital elevation model (DEM), the correction of displacement by relief can be made. Orthorectification is made through corresponding points of the reference image with the DEM. When sufficient control points are collected, new dimensions for the image are calculated (a cubic resampling process is selected to preserve the spectral response in both images), and the final step in the generation process is to calculate digital values for every pixel in the output image.

2.2 Study and applications of fusion techniques

One way to take advantage of the panchromatic and multispectral images of IKONOS is to synthesize composed hybrid images by means of an image multiresolution fusion method. For this purpose experimental tests were implemented to select the decomposition depth, coefficients combination criteria were carried out. Different fusion methods were compared (Laplaciana pyramid, Gradient pyramid, Wavelet transform, and Analysis of Principal Component).

The concept of multiresolution analysis plays a central role in the fusion process. Figure 1 shows the scheme of multiresolution fusion. The idea of pyramidal methods was first presented in image processing by Burt and Adelson in 1983 in a scheme which they called Laplacian pyramid. The proposed pyramid method uses Gaussians functions taken from their relevance in the human vision system. A pyramid can be described like a filtered copy collection in which the image is reduced in each step. From this work different pyramid models have arised such as the Gradient pyramid.
The image pyramid is a data structure designed to support efficient scaled convolution through reduced image representation. It consists of sequential copies of the original image in which the resolution is decreased in regular steps. These reduced resolution levels of the pyramid are themselves obtained through a highly efficient iterative algorithm.

The first step to construct a Laplacian pyramid is to apply a low pass filter on the original image $G_0$ and then get a subsample by a factor of two in order to obtain the next pyramid level, $G_1$. In the same way $G_1$ is then filtered and subsampled to obtain $G_2$. Subsequent repetitions of this filter step generate the remaining pyramid levels. Then the $k^{th}$ level of the Gaussian pyramid is defined as: $G_k = \left[w \ast G_{k-1}\right]\downarrow$, where $w$ is the generating kernel, and the notation $\left[\ldots\right]\downarrow$ indicates that the image array is subsampled by a factor $n$. The Laplacian pyramid is defined as the difference between successive levels of a Gaussian pyramid: $l_k = G_k - 4w \ast \left[ G_{k+1}\right]\downarrow$, and $\left[\ldots\right]\uparrow$ indicating up sampling. An image is recovered from its Laplacian pyramid by reversing these steps.

For the construction of a Gradient pyramid, a 5 by 5 binomial kernel $w$ is used. The Gradient pyramid ($D_{kl}$) is obtained from $G_k$ through convolution with filters $d_i$, where $d_i$, for $l = 1\ldots4$ are a set of gradient-oriented filters.

$D_{kl} = d_i \ast \left[ G_k + \hat{w} \ast G_k \right]$.

To reconstruct an image from its Gradient pyramid, each gradient pyramid level $D_{kl}$ is converted to a corresponding second derivate pyramid (or oriented Laplacian) level $l_{kl}$:

$l_{kl} = -\frac{1}{8} d_i \ast D_{kl}$.

The oriented Laplacian pyramids are then summed to form a filter-subtract-decimate (FSD) Laplacian pyramid, $L_k$:

$L_k = \sum_{l=1}^{4} l_{kl}$.
Reconstruction is completed by converting the Laplacian to the FSD Laplacian, and then converting the FSD pyramid to the Gaussian pyramid.

The Laplacian pyramid ideas motivated S. Mallat in 1989 to demonstrate that a set of orthonormal wavelet provides an alternative to achieve this type of analysis. Thus Mallat\(^6\), in collaboration with Meyer, demonstrated the mathematical model of the wavelet multiresolution process. Recently, the wavelet transform has been shown as a powerful fusion method that preserves the spectral characteristics of images\(^7,\(^8,\(^9\). In this study we used the Haar Wavelet transform. The motivation for using the Haar Wavelet is the easy implementation. Four pyramidal levels were used in this study. This value has been found by using a trial and error technique, and they are related to the level of detail in the images being considered.

2.3 Rule of fusion pyramid

After the construction of the input pyramids, different methods of extracting information from the input image can be used for the selection of specific patterns. The method exploited throughout this paper is based on salience \( (S) \) and match measures \( (M) \). This method is based on the selection of the higher value coefficients.

In order to acquire the information carried by each pyramid, the salience of each pattern has been considered\(^2,\(^10\). The salience of a particular component pattern is high if that pattern plays an important role in the information the scene. A good measurement of pattern salience is pixel intensity. It can be defined as the local energy of the incoming pattern with neighborhood \( P \), as follows

\[
S(i, j, k, l) = \sum_{i,j} p(i, j) P(i + i', j + j', k, l)^2,
\]

where \( p \) is a window function around the considered pixel, \( P \) is the pyramid and \((i, j, k, l)\) are the row and column sample position, level and orientation indexes inside the pyramid structure, respectively. When the salience of the coefficient \((i, j)\) with level \( k \) and orientation \( l \) is computed, the local window allows to acquire the 3x3 pixel neighboring coefficients around the considered point. The value of \( P(i, j) \) is the weight of the sample, within the area of neighborhood \( P \). \( P(i + i', j + j') \) returns the pixel value inside the image at location \((i, j)\), with on offset \( i' \) and \( j' \). After this salience computation is applied to each one of the single pyramids, \( A \) and \( B \), a match measure is computed to combine the information carried by each pyramid; this match measure is defined on the basis of the salience measure, as the following

\[
M_{AB}(i, j, k, l) = \frac{2 \sum_{i,j} p(i, j) P_A(i + i', j + j', k, l) P_B(i + i', j + j', k, l)}{S_A(i, j, k, l) S_B(i, j, k, l)}
\]

Each level of the pyramid can now be fused by testing the match measures between the two images at the given level. The combined pyramid \( C(i, j, k, l) \) can be described by using the following expression:

\[
C(i, j, k, l) = w_A(i, j, k, l) P_A(i, j, k, l) + w_B(i, j, k, l) P_B(i, j, k, l),
\]

where \( w_A \) and \( w_B \) are the assigned weights of the source images. The way to implement the weighting function is to select some threshold \( \alpha \),

if \( M_{AB}(i, j, k, l) \leq \alpha \Rightarrow w_{\text{min}} = 0 \) and \( w_{\text{max}} = 1 \).
otherwise $M_{Ab}(i, j,k,l) > \alpha \Rightarrow w_{\min} = \frac{1}{2} - \frac{1}{2} \left( 1 - \frac{N_{Ab}}{1 - \alpha} \right)$ and $w_{\max} = 1 - w_{\min}$

if $S_A(i, j,k,l) \geq S_B(i, j,k,l) \Rightarrow w_A(i, j,k,l) = w_{\max}$ and $w_B(i, j,k,l) = w_{\min}$.

else if $S_A(i, j,k,l) < S_B(i, j,k,l) \Rightarrow w_A(i, j,k,l) = w_{\min}$ and $w_B(i, j,k,l) = w_{\max}$

where $\alpha$ is a threshold fixed by the operator. Match and saliency measures are computed within a 3 X 3 neighborhood ($p$) and with the transition $\alpha$ between selection and fusion modes equal to 0.85. The final step is to use the inverse transform to obtain the fused image.

Another studied method is based on Principal Component Analysis (PCA). This is a commonly used technique for analysis of remote sensing images. The principle of image fusion is to get the best resolution without altering the spectral contents of the image. The idea of the PCA method by substitution, is to replace the first principal component (PC1) by a higher resolution image. This PC1 accounts for the maximum variance, which results in an improved resolution of data in the fused image. Figure 2 shows the visual results of the different fusion methods.

Figure 2. Comparison of fusion methods. a) Laplacian pyramid, b) Gradient pyramid, c) Original multispectral 4m
As compared to the original color image figure 2c, the color of the fused images obtained by Laplacian pyramid, Gradient pyramid, Wavelet transform and Principal Component Analysis provide good results. The red areas in these images represent the vegetations areas. These four images possess about the same color and brightness as those in the original IKONOS (4m) image.

2.3 Classification of images for the identification of green areas

At present, a few quantitative evaluation methods of image fusion techniques have been carried out. Most of the existing evaluations are based on visual appreciation. Obviously, trees and meadows cannot be classified properly because they similar spectral properties. In this work we use a vegetation index derived from red and infrared bands and an ISODATA classification method.

The vegetation index was applied to the results of the images obtained by the previous fusion methods, that is to say, to the new synthetic images. The vegetation index was obtained by calculating the ratio of red and near infrared bands, \( \frac{\text{red} - \text{nir}}{\text{red} + \text{nir}} \), whose results are shown in figure 3. The resulting fused images do show the presence of vegetation, remarkably on the basis of the difference between visible and NIR reflectance.
In order to validate the vegetation enhancement data, we used again the PCA method. The first principal component was then replaced by the ratio \( \frac{\text{red} - \text{nir}}{\text{red} + \text{nir}} \) information, and the resulting images were then obtained by inverting the PCA transform. Afterwards, the images were classified by means of the ISODATA unsupervised classification approach to segmentation. Figure 4 shows the comparison of the classified results.

![Classification images](image)

Figure 4. Classification obtained from a) Laplacian pyramid, b) Gradient pyramid, c) Original multispectral 4m, d) Haar Wavelet transform and e) Principal Component Analysis. The dark gray color is the tree class and gray is field class.

The accuracies of the conventional classification are reported in Table 1. The accuracy assessment was carried out using the random method in an area of 1500 x 1500 pixels. Three hundred randomly selected points were assessed in the area. In order to validate the registered random points, they were compared with visual reference data, in the IKONOS bands red and NIR, and also with field work data.

On table 1 we note that for the tree class the best classification performance is achieved by PCA followed by Haar Wavelet transform, while for the field class the best performance is obtained with the original IKONOS image followed by the Haar wavelet transform and PCA. It is important to mention that in the case of PCVA and Haar Wavelet transform class boundaries are better defined than with the other methods.
Table 1. Accuracy of the trees and field areas classification using a Kappa Statistics

<table>
<thead>
<tr>
<th>Fusion method</th>
<th>Classification Results</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Class 1 Trees</td>
<td>Class 2 Field</td>
<td></td>
</tr>
<tr>
<td></td>
<td>User’s Accuracy (%)</td>
<td>Kappa Statistics</td>
<td>User’s Accuracy (%)</td>
</tr>
<tr>
<td>IKONOS 4m original</td>
<td>94.286 0.9356</td>
<td>98.312 0.9244</td>
<td></td>
</tr>
<tr>
<td>Laplacian Pyramid</td>
<td>83.333 0.8038</td>
<td>75.00 0.6998</td>
<td></td>
</tr>
<tr>
<td>Gradient Pyramid</td>
<td>92.105 0.9068</td>
<td>87.234 0.8449</td>
<td></td>
</tr>
<tr>
<td>Haar Wavelet Transform</td>
<td>94.595 0.9362</td>
<td>97.500 0.9482</td>
<td></td>
</tr>
<tr>
<td>Principal Component Analysis</td>
<td>99.583 0.9788</td>
<td>92.593 0.9189</td>
<td></td>
</tr>
</tbody>
</table>

Figure 5, demonstrates that the classification of the PCA result with the substitution of PC1 for the ratio of red and infrared band is sufficient for the extraction of trees and meadows in high resolution. The image of 5a presents the fused image of Mexico City (Baseball Stadium (Foro Sol), Sport Center (Madgalena Mixuca), Racing Circuit (Hermanos Rodríguez) and figure 5b the classified image, extracting the vegetation cover of forest and grass. It is possible to observe that the Baseball Stadium is not included within the vegetation cover because it features synthetic grass. In spite of having similar color, trees and field areas were discriminated from synthetic lawns and playgrounds thanks to their spectral response.

Figure 5a Images fused IKONOS Baseball Stadium (Foro Sol), Sport Center (Madgalena Mixuca), Racing Circuit (Hermanos Rodríguez), 5b Classification

The results obtained from the definition and location of the different types of green areas in Mexico City, generates a great amount of information. When converting the classified image to raster format, it generates many polygons, since the IKONOS images offer the possibility of identifying individual elements such as trees. Therefore, it was necessary to choose a minimal mapping unit of 160m², allowing the identification of parks, tree areas and fields.
3. CONCLUSIONS

The fusion methods evaluated, maintained the radiometric properties of the original multispectral image. Furthermore, with fusion, the spatial resolution of the images was improved.

A comparison of the synthetic images with the original ones revealed that the best method to merge panchromatic and multispectral IKONOS is Principal Component Analysis followed by Wavelet transform and Gradient pyramid.

It is necessary to continuing with the study of image fusion testing other types of fusion methods similar to the Wavelet transform such as the Hermite transform. It is worthwhile to consider. The Hermite transform provides a better image representation model for image fusion due to the operators based on Gaussian derivatives. Furthermore noise removal is applied to each image to fuse, allowing for image restoration.

This improved method of monitoring urban green areas allows for a more efficient extraction of the forest cover (trees) areas, better than manual interpretation. The effects of shadows and sun angle were not taken into account.

ACKNOWLEDGMENT

This work was sponsored by Centro de Investigación en Geografía y Geomática.

REFERENCES