2021 18th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE). Mexico City, Mexico. November 10-12, 2021

Quantitative Comparisons of Edge Based and Region Based Feature Detection in Digital Aerial Imagery Analysis

Zhengmao Ye
College of Science and Engineering, Southern University, Baton Rouge, USA
zhengmao_ye@subr.edu

Abstract— **Digital image segmentation focuses on separation of one or more features in an image based on certain similarity or discontinuity criterion. It is inevitable for the aerial imagery to be contaminated by artifacts and noises, which causes extra blurry boundary, intensity inhomogeneity, and high complexity. Both edge based and region based feature detection are broadly applied to segmentation of digital aerial imagery. Numerous comparative results between the two have been obtained based on qualitative analysis, while conclusions are lack of convincing data-driven evidences instead. In this article, the systematic quantitative analysis approach is proposed to compare two dominating feature detection techniques using the RGB model and HSI model, from both the frequency domain and spatial domain perspectives. Three popular digital image segmentation schemes are employed for comparisons, including Canny edge detection, Gabor edge detection, and kernel fuzzy C-Means clustering. A useful set of quantitative information metrics has been introduced, such as the discrete entropy, mutual information, homogeneity, contrast, and dynamic range, in order to enhance the decision making accuracy on feature detection and ensure information integrity.**

Keywords— Digital Aerial Imagery, Canny edge detection, Gabor edge detection, Kernel Fuzzy C-Means clustering, Information Metrics

I. INTRODUCTION

Edge detection and region detection are two typical digital image processing methods to extract useful features such as changes and discontinuities in the intensity and texture. Edge detection could be search-based or zero-crossing based to locate and link edge pixels and then form potential contours. Region detection could be more robust since regions usually cover more pixels than edges with more information to describe. Edge detection is sensitive to noises however edges act as important image features to separate multiple regions. There are various feature detection approaches. The classical Canny edge detection, powerful Gabor edge detection, and popular kernel fuzzy C-Means clustering are among the most successful feature detection approaches, which are selected for the comparison purpose in the context [1-2].

Canny edge detection locates edges by a Gaussian filter to reach the local minima from the intensity gradient. Double thresholds are selected to find strong and weak edges. Path tracing is performed to detect true weak edges. To remove the unnecessary points on image edges, edge thinning can be used while the extended adaptive edge tracing is able to further enhance the accuracy. Artificial Intelligence has been widely applied to edge detection as well. For instance, Ant Colony Optimization (ACO) will help to produce more true edges and avoid false edges at the same time [3-5]. A Gabor filter acts as an energetic edge detection technique. Gabor wavelets can be used for object detection. The 2D Gabor function does not belong to orthonormal wavelet transform, but under certain conditions, it can still perform full representation with many real world applications. Integration of Gabor wavelet transform and adaptive contrast stretching instead can implement edge detection and contour tracing at the same time, where the Gabor filter and discrete wavelet transform are applied for contour tracing of true color RGB images. The wavelet functions are also capable of generating high resolution in both spatial and frequency domains [6-8]. The extensions of Gabor filters are optimal Gabor filter and Gabor pyramid. To identify and extract the surface topography features, Gabor filters and wavelets as well as pyramid decompositions are examined and applied to the selected set of texture characterization cases together with applications of topography partitioning [9]. The Gabor filter can also be easily applied to biometric feature identification systems. The ear features are extracted via the introduction of the Gabor operator. Both local features and global features are optimized using Genetic Algorithms [10].

The typical region-based segmentation is fuzzy C-Means clustering. The fuzzy C-partition has a variety of applications to interpret data. The convergence theorem can also involve in designing a biased fuzzy C-Means algorithm with a focal point, which is less sensitive to the initialization and faster in the convergence [11]**.** C-Means clustering partitions the feature space into multiple regions in terms of the position, intensity, color and texture. However contour information is missing sometimes. Fuzzy C-Means clustering instead aids in finding vague boundaries so that optimization regarding both contour information and region information can be implemented [12]. Fast kernel based fuzzy C-Means clustering has also been presented for robust segmentation with the high accuracy, via integration of watershed transform and level set schemes. In this case, edge based level set method and region based fuzzy C-Means clustering can be combined. To achieve fast segmentation, watershed transform is applied to locate the initial contour of the fast level set method. Initial cluster centers of fuzzy C-Means clustering are selected in closed contours to avoid misclassification. It has reduced computation time on segmentation of densely distributed aerial images [13].

978-1-6654-0029-9/21/$31.00 ©2021 IEEE

Quantitative comparisons between edge detection and region detection will be made, where the information metrics being chosen will cover both spatial domain and frequency domain features. In particular, metrics of the discrete entropy, discrete energy, mutual information, homogeneity, contrast, and dissimilarity will be selected in the RGB color model together with the dynamic range in the HSI color model.

II. RGB AND HSI COLOR SPACES

Nonlinear visual perception occurs using the Trichromatic Theory of additive color vision. The actual color is the mixture of RGB (Red, Green, Blue) components in the 3D Cartesian system. The RGB model is defined inside a cube with Red, Green and Blue as 3 primary color axes in the range of 0 to 255 (8 bits). The complete RGB color representation consists of 6 hexadecimal bits. Black and white are located at two corners of the main diagonal. A composite color is the vector being mapped into the cube. Projection of the RGB component to its main diagonal gives rises to the gray level. The dimensions of the RGB true color and gray level images are (M×N×3) and (M×N), respectively. The HSI (Hue, Saturation, Intensity) color model is similar to the human visual sense of colors. Each color has three components in the HSI model: hue angle, color saturation, and intensity. Intensity and color information will be independent instead. The hue component uses an angle to describe the color itself in the form of an angle between [0, 2π], such as 0 for red and π/3 for yellow. The saturation component has a range between 0 and 1. Intensity is between 0 and 1 as well, which limits the saturation range. The RGB and HSI color spaces are interchangeable using simple conversion formulas. Both spaces are popular in the machine vision.

III. ADAPTIVE CANNY EDGE DETECTION

The fundamental Canny edge algorithm detects edges of a digital image based on zero-crossings of the directional 2nd order derivative. Canny edge detection starts with Gaussian convolution against noises with an optimal smoothing filter shown in (1). The digital image is smoothed via Gaussian convolution in (2).

$$G(x, y, \sigma)= \frac{1}{2\pi\sigma^2}\exp(-\frac{x^2+y^2}{2\pi\sigma^2}) \quad (1)$$

$$H(x, y) = I(x, y) * G(x, y) \quad (2)$$

where I shows the intensity of the source image; H shows the intensity of the processed image after smoothing; G represents the Gaussian smoothing filter; * indicates convolution.

The edge strength and edge direction are expressed as the gradient magnitude and gradient direction respectively after the 2D spatial gradient is computed. Then the edge is located at the local optima of 1st order derivatives of the smoothed image H(x, y) in the direction of the gradient magnitude. It acts as the zero-crossing point of its 2nd order derivative. The edge direction will be rounded to one out of eight evenly distributed angles [0, 2π] to manifest vertical, horizontal and diagonal directions.

The next step is to apply non-maximal suppression. The edge direction is used to trace along the edge and also to suppress the non-edge pixels. Thus local maximal gradient strength is reached. The edge gradient of each pixel is compared with gradients of all 8 neighbors in the gradient direction. Finally in order to eliminate possible broken edges, adaptive edge tracking can be further introduced using the Chain Code criterion on each visiting node. A digital curve is formulated as an integer sequence in terms of the relative position between the actual edge node and all neighbors in the 2D spatial domain. The 8-connectivity Chain Code can be applied to illustrate the pixel thin line trajectory. Thresholding with hysteresis (high or low threshold) is implemented so that proper edge tracing and clear marking are achieved without broken edges.

IV. 2D GABOR WAVELETS FOR EDGE DETECTION

The practical 2D Gabor filter h(x,y) is designed as a typical Gaussian kernel function modulated by the complex sinusoid. It can provide diverse orientations and different frequencies to capture edges effectively without fragment connecting via the Chain Code needed, which is identical to human vision. In the spatial domain, the 2D Gabor function is expressed as a product of the Gaussian shaped envelope g(x,y) and complex sinusoid carrier c(x,y) as (3). In (4-5), σ_x and σ_y act as the standard deviations in x and y directions, u_0 and v_0 act as the central frequencies.

$$h(x,y)=c(x,y)g(x,y) \quad (3)$$

$$c(x, y)=e^{-j2\pi(u_0x+v_0y)} \quad (4)$$

$$g(x, y)=\frac{1}{2\pi\sigma_x\sigma_y}e^{-(\frac{x^2}{2\sigma_x^2}+\frac{y^2}{2\sigma_y^2})} \quad (5)$$

Thus the 2D Gabor filter is simply formulated as (6). The high frequency edges are relevant to high output values when the Gabor filter is applied to a digital image.

$$h(x,y)=c(x,y)g(x,y)=\frac{1}{2\pi\sigma_x\sigma_y}e^{-(\frac{x^2}{2\sigma_x^2}+\frac{y^2}{2\sigma_y^2})}e^{-j2\pi(u_0x+v_0y)} \quad (6)$$

In the frequency domain instead, frequency shifts occur in the u axis and v axis of the frequency response.

$$H(u, v) = G(u-u_0, v-v_0) \quad (7)$$

Via the convolution theorem, H(u, v) is shown as (8):

$$H(u, v)=\frac{e^{-[\frac{(u-u_0)^2}{2\sigma_u^2}+\frac{(v-v_0)^2}{2\sigma_v^2}]}}{2\pi\sigma_x\sigma_y}=2\pi\sigma_u\sigma_v e^{-2\pi^2[(u-u_0)^2\sigma_u^2+(v-v_0)^2\sigma_v^2]} \quad (8)$$

Still σ_u and σ_v represent standard deviations in the u and v axes, where we have (9).

$$\sigma_x = 2\pi\sigma_u;\ \sigma_y = 2\pi\sigma_v \quad (9)$$

Wavelet transform is widely applied to the time varying and transient spectral analysis. 2D Gabor wavelets can be produced by performing dilation and rotation operations on the Gabor function, using the scaling and phase factor.

$$g_{mn}(x, y) = g(x_0, y_0)/ \alpha^m \quad (10)$$

where the scaling factor $s = \alpha^{(-m)}$ (α=2 herein; m = 0, 1, ..., M−1) and phase factor $\theta_n = n\pi /N$ (n = 0, 1, ..., N−1). M and N indicate the total number of scales and orientations. x_0 and y_0 are formulated in (11-12).

$$x_0= x\cos\theta_n+ y\sin\theta_n \quad (11)$$

$$y_0= -x\sin\theta_n+ y\cos\theta_n \quad (12)$$

At each level of discrete wavelet transform, when α is equal to 2, it generates a 2×2 quadrant block with 4 subbands.

Both the optimal Gabor filter and Gabor wavelet pyramid are further expansions of the 2D Gabor filters. The former is based on different parametric optimization schemes. The latter produces low and high resolution by either downsampling and upsampling, respectively. Since the comparison between the edge detection and region detection is the major goal of this article, the fundamental Gabor filter is applied instead.

V. Kernel Based Fuzzy C-Means Clustering

C-Means clustering is the typical approach for region based feature segmentation. Fuzzy C-Means clustering is more practical for data analysis because the cluster boundary itself has the degree of uncertainty. Its objective is to partition a finite dataset into C fuzzy clusters in order to minimize the cost function. Gaussian fuzzy membership function is chosen. The pixels near the cluster boundary have smallest degrees of belonging and the pixels around cluster centers have largest degrees of belonging. To improve the accuracy and robustness, kernel fuzzy C-Means clustering is also introduced instead. The kernel function makes nonlinear to linear transformation by implicitly mapping source images into high dimensional kernel space. The kernel distance and Gaussian kernel functions are expressed as (13-14).

$$\|\Phi(x_i) - \Phi(x_j)\|^2 = G(x_i, x_i) + G(x_j, x_j) - 2G(x_i, x_j) \quad (13)$$

$$G(x_i, x_j) = \exp\left(\frac{-\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (14)$$

After cluster center initializations, the fuzzy membership function (15) is defined as the ratio of single cluster kernel distance measure to the sum of every kernel distances from the pixel to all C clusters selected. A fuzziness index m ($1 \le m < \infty$) is used for kernel fuzzy C-Means clustering.

$$\mu_{ij} = \frac{\|\Phi(x_j) - \Phi(c_i)\|^{-\left(\frac{2}{m-1}\right)}}{\sum_{i=1}^{C} \|\Phi(x_j) - \Phi(c_i)\|^{-\left(\frac{2}{m-1}\right)}} \quad (15)$$

The centroids can then be updated as (16).

$$c_i = \frac{\sum_{j=1}^{N} \mu_{ij}^m G(c_i, x_j) x_j}{\sum_{j=1}^{N} \mu_{ij}^m G(c_i, x_j)} \quad (16)$$

The cost function is corresponding to spatial information in the region based segmentation scheme. The position, intensity, texture, and color feature in the spatial space are all taken into account to formulate the fuzzy cost function as (17). The vague boundaries can be determined accordingly.

$$J^{\Phi}(X, C) = \sum_{i=1}^{C} \sum_{j=1}^{N} \mu_{ij}^m \|\Phi(x_j) - \Phi(C_i)\|^2 \quad (17)$$

where $X=\{x_j\}$ is a data matrix whose x_j is the j-th element of the fuzzy matrix. $C=\{c_i\}$ shows a vector of cluster centers whose c_i is the centroid of its i-th cluster. $\|.\|^2$ represents the L_2 norm. At each iteration, both fuzzy memberships and cluster centers will be updated until the cost function is minimized to achieve optimal image segmentation.

VI. Numerical Simulations

Numerical simulations will be conducted with respect to 3 three typical digital aerial images with diverse information complexity (i.e., bird's eye views of Rio de Janeiro, Machu Picchu, Cape Town). Both edge based detection and region based detection are implemented on these aerial images. The classical Canny edge detection and more practical Gabor edge detection are used to generate edge based detection outcomes, while kernel fuzzy C-Means clustering with different cluster numbers is used to generate region based detection outcomes. All three tristimulus (red, green and blue) components have the intensity distributions in a range from 0 to 255. Qualitative comparisons could be simply observed from those outcomes listed in Fig. 1, where three source images are followed by corresponding outcomes of Canny edge detection, Gabor edge detection, and 3 cases of kernel fuzzy C-Means clustering (3 clusters, 5 clusters, 10 clusters), respectively.

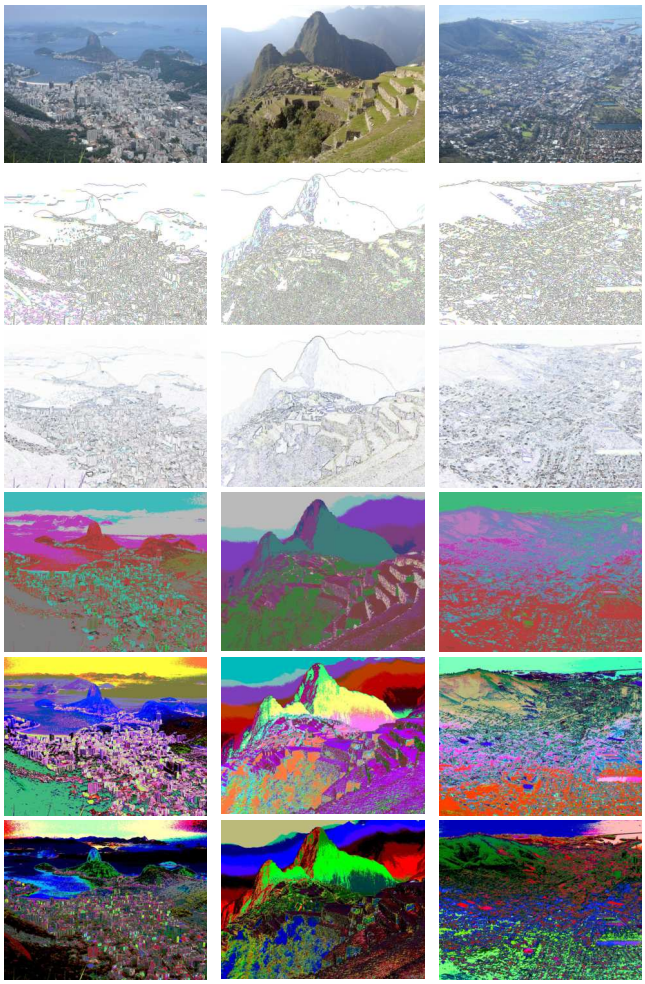

Fig. 1. Feature Detection of Digital Aerial Images with Rows in Order 1) Three Source Images; 2) Canny Edge Detection; 3) Gabor Edge Detection; 4) Fuzzy C-Means Clustering (3-Cluster); 5) Fuzzy C-Means Clustering (5-Cluster); 6) Fuzzy C-Means Clustering (10-Cluster)

Rather than discussing the merits and drawbacks from the qualitative analysis being widely conducted in the literature, objective evaluations via thoroughly quantitative analysis is the focus here, which has never been systematically taken into

consideration yet. In particular, frequency domain analysis and spatial domain analysis using the RGB model and dynamic range analysis using the HSI model will all be implemented. A comparison on computation time is relatively simple however not the focus as well. Computational complexity depends on the digital image resolution. For certain aerial image with the fixed resolution, it can be easily demonstrated that edge detection is more time efficient than region detection in general. From a typical case, computation time (in secs) of 5 schemes is listed as below: Canny edge detection (3.019844), Gabor edge detection (18.329676), kernel fuzzy C-Means clustering (3-Cluster: 107.366739), kernel fuzzy C-Means clustering (5-Cluster: 185.841550), kernel fuzzy C-Means clustering (10-Cluster: 393.031783), respectively.

VII. Quantitative Analysis in Frequency Domain

In frequency domain analysis, the set of information metrics being selected includes the discrete entropy, discrete energy and mutual information in this comparative study. To manifest the color balance in three individual channels (Red, Green, Blue) from the RGB model, all 3 channels will be subject to independent quantitative analysis simultaneously. A similar approach to grayscale image analysis is performed, where the occurrence frequency of pixel counts is computed for each of the 256 intensity levels in RGB channels.

A. Entropy

Discrete entropy indicates average information conveyed from the source image or processed image. It is a measure of uncertainty or randomness as shown in (18). It is formulated as the sum of product terms of the outcome probability and the logarithm of its inverse, covering every outcomes of the event $\{x_1, x_2, \ldots, x_k\}$. In (18), p(i) shows the probability distribution function of histogram, while k indicates the actual count of intensity levels.

$$H(x)=\sum_{i=1}^{k} p(i)\log_2 \frac{1}{p(i)} = -\sum_{i=1}^{k} p(i)\log_2 p(i) \tag{18}$$

B. Energy

Discrete energy of the digital image follows the energy definition for those discrete time sequences with the finite lengths. It is expressed as (19) which also describes the randomness. p(j) still represents a probability distribution function in terms of the histogram for each RGB channel. The energy has its maxima of one when the whole image has the constant intensity. The higher energy is in fact corresponding to lower information content.

$$E(x)=\sum_{j=1}^{k} p(j)^2 \tag{19}$$

C. Mutual Information

Mutual information acts as a symmetric function which shows statistical dependence between variables. It indicates the information amount extracted on one random variable through observing another random variable. It is defined as (20), where mutual information is expressed as I(X; Y). H(X) and H(X|Y) instead represent the discrete entropy and conditional entropy. When two digital images are totally independent, I(X;Y) reaches zero. Large mutual information depicts a big reduction in uncertainty while small mutual information depicts a slight reduction.

$$I(X;Y)=H(X)-H(X|Y)=\sum_{X,Y} p_{XY}(X, Y)\log_2 \frac{p_{XY}(X, Y)}{p_X(X)p_Y(Y)} \tag{20}$$

In Tables 1-3, information metric outcomes of quantitative analysis in the frequency domain are listed. Same conclusions can be made across all 3 cases of diverse source images. The source images have the largest discrete entropy, followed by kernel fuzzy C-Means clustering (10-Cluster), kernel fuzzy C-Means clustering (5-Cluster), Gabor edge detection, and kernel fuzzy C-Means clustering (3-Cluster), while Canny edge detection produces the smallest discrete entropy, which contains least amount of information. The source image does contain maximal amount of information. However Gabor edge detection can even show more information than region based kernel fuzzy C-Means clustering with a small cluster number.

However, source images have the smallest discrete energy, followed by kernel fuzzy C-Means clustering (10-Cluster), kernel fuzzy C-Means clustering (5-Cluster), Gabor edge detection, and kernel fuzzy C-Means clustering (3-Cluster), while Canny edge detection produces the largest value of discrete energy, corresponding to least amount of information. The source image is corresponding to the smallest energy but maximal amount of information. Gabor edge detection instead produces smaller energy and shows more information than the region based kernel fuzzy C-Means clustering with a small cluster number.

For mutual information between the source images and corresponding detected images, Canny edge detection has produced the smallest mutual information, followed by Gabor edge detection, kernel fuzzy C-Means clustering (3-Cluster), kernel fuzzy C-Means clustering (5-Cluster), and kernel fuzzy C-Means clustering (10-Cluster). It manifests the fact that edge detection outcomes generally show less dependency on source images than region detection outcomes.

TABLE 1 QUANTITATIVE METRICS IN FREQUENCY DOMAIN – CASE 1 (RED, GREEN, BLUE) CHANNELS

RIO	Source	Canny	Gabor	3-Cluster	5-Cluster	10Cluster
Entropy	7.286801	3.762427	5.973869	5.922676	6.957284	7.280495
Energy	0.007626	0.314093	0.029768	0.038884	0.026043	0.016582
Mutual		0.115382	0.427524	1.113105	1.211792	1.674995
Entropy	7.240504	3.829991	5.905175	5.903573	6.687829	7.169392
Energy	0.007707	0.280669	0.032415	0.048471	0.032844	0.016919
Mutual		0.143829	0.434015	0.816597	0.957272	1.721606
Entropy	7.435901	3.734364	5.950179	5.904597	7.257223	7.373159
Energy	0.006708	0.333992	0.035734	0.058523	0.015403	0.014774
Mutual		0.203885	0.457256	1.090746	1.183396	1.59985

TABLE 2 QUANTITATIVE METRICS IN FREQUENCY DOMAIN – CASE 2 (RED, GREEN, BLUE) CHANNELS

Machu	Source	Canny	Gabor	3-Cluster	5-Cluster	10Cluster
Entropy	7.478202	3.999199	5.896837	5.611307	6.763562	7.105285
Energy	0.007825	0.301083	0.030697	0.055195	0.028251	0.015120
Mutual		0.107376	0.407902	1.045900	1.190451	1.541148
Entropy	7.498958	3.931777	5.805637	5.319919	6.294713	7.106725

Energy	0.007921	0.290977	0.034240	0.059878	0.038507	0.015760
Mutual		0.147006	0.502755	1.094993	1.367240	1.778151
Entropy	7.698102	4.051716	5.981634	5.952255	6.489385	7.272714
Energy	0.007271	0.318659	0.037402	0.037804	0.023290	0.017839
Mutual		0.246308	0.649304	1.628131	1.424969	1.654642

TABLE 3 QUANTITATIVE METRICS IN FREQUENCY DOMAIN – CASE 3 (RED, GREEN, BLUE) CHANNELS

Cape	Source	Canny	Gabor	3-Cluster	5-Cluster	10Cluster
Entropy	7.779091	4.566382	6.589532	6.544005	7.068440	7.709158
Energy	0.007426	0.200812	0.016571	0.017444	0.016274	0.007880
Mutual		0.103498	0.287686	1.017801	1.038518	1.433205
Entropy	7.358212	4.517357	6.512067	6.388109	7.303062	7.332833
Energy	0.007058	0.184151	0.014380	0.022266	0.010959	0.010526
Mutual		0.125940	0.326198	1.085338	1.094416	1.253310
Entropy	7.713123	4.584348	6.726122	6.631067	7.416098	7.642120
Energy	0.006418	0.229275	0.018985	0.019650	0.010023	0.008633
Mutual		0.118322	0.247020	0.679529	0.814145	1.187898

VIII. QUANTITATIVE ANALYSIS IN SPATIAL DOMAIN

In spatial domain analysis, the set of information metrics being selected covers homogeneity, contrast and dissimilarity in this comparative study. Now the co-occurrence matrix is introduced in the spatial domain analysis. It is a matrix which is relevant to the distribution of grayscale values or RGB color components in rows and columns of a digital image.

D. Homogeneity

Homogeneity serves as the direct measure of the local variations of the true color or gray level images. Small homogeneity values result from high structural variations and large homogeneity values result from low structural variations. The homogeneity S is expressed as (21).

$$S=\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\frac{1}{1+(i-j)^2}g(i,j) \tag{21}$$

where M and N stand for total numbers of pixels in the row and column of the digital image; i and j are indexes at the x and y axes of the co-occurrence matrix; g(i, j) shows the pixel intensity in the co-occurrence matrix at the coordinates i and j.

E. Contrast

Contrast has been used to compute variations of intensity distribution in each color channel. It illustrates the intensity differences between the features and background within the same range being shown. It is formulated as (22), where g_{AVG} shows the average intensity.

$$C=\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\frac{[g(i,j)-g_{AVG}]^2}{M*N} \tag{22}$$

F. Dissimilarity

Dissimilarity depends on both the perceptual intensity and local distance representations between 2 specific pixels in the co-occurrence matrix. It has been defined as (23), where g(i, j) indicates the pixel intensity in the co-occurrence matrix at coordinates i and j.

$$D=\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}g(i,j)\,|i-j| \tag{23}$$

In Tables 4-6, information metric outcomes of quantitative analysis in the spatial domain are listed. Same conclusions can still be made across all 3 cases of source images. Canny edge detection gives rise to lowest homogeneity, followed by kernel fuzzy C-Means clustering (10-Cluster), kernel fuzzy C-Means clustering (5-Cluster), source images and kernel fuzzy C-Means clustering (3-Cluster), while Gabor edge detection gives rise to highest homogeneity.

The source image and kernel fuzzy C-Means clustering have the homogeneity values between Canny edge detection and Gabor edge detection with highest and lowest structural variations, respectively. Kernel fuzzy C-Means clustering can produce different structural variations according to the cluster number, smaller than that of source images at the low cluster number cases, but larger than that of source images at large cluster number cases.

TABLE 4 QUANTITATIVE METRICS IN SPATIAL DOMAIN – CASE 1 (RED, GREEN, BLUE) CHANNELS

RIO	Source	Canny	Gabor	3-Cluster	5-Cluster	10Cluster
Homo	0.874548	0.753787	0.918203	0.882118	0.779212	0.768068
Contrast	0.358165	5.052915	0.172899	0.329910	0.931690	0.935484
DisSim	0.268563	0.934018	0.165145	0.268120	0.519881	0.539150
Homo	0.875124	0.759389	0.918299	0.885792	0.780552	0.758410
Contrast	0.355518	5.332076	0.172465	0.351541	1.026565	1.099633
DisSim	0.267173	0.951290	0.164912	0.248937	0.532930	0.581860
Homo	0.876816	0.769787	0.922346	0.885524	0.777118	0.785936
Contrast	0.349934	4.787596	0.163897	0.305529	0.822946	0.858759
DisSim	0.263430	0.917648	0.156740	0.241712	0.537218	0.539915

TABLE 5 QUANTITATIVE METRICS IN SPATIAL DOMAIN – CASE 2 (RED, GREEN, BLUE) CHANNELS

Machu	Source	Canny	Gabor	3-Cluster	5-Cluster	10Cluster
Homo	0.867561	0.766997	0.924562	0.916111	0.814865	0.775777
Contrast	0.373723	5.571923	0.158377	0.187512	0.575900	0.953935
DisSim	0.282721	1.069370	0.152126	0.171066	0.404298	0.530461
Homo	0.867944	0.776352	0.924762	0.923559	0.837971	0.806795
Contrast	0.372727	5.826186	0.157860	0.165699	1.193523	1.213557
DisSim	0.281918	1.075224	0.151707	0.168352	0.561684	0.573385
Homo	0.869182	0.756898	0.925964	0.920613	0.835005	0.833363
Contrast	0.367289	5.060641	0.154729	0.156940	0.809403	0.893676
DisSim	0.278974	1.041228	0.149181	0.149468	0.51197	0.559598

TABLE 6 QUANTITATIVE METRICS IN SPATIAL DOMAIN – CASE 3 (RED, GREEN, BLUE) CHANNELS

Cape	Source	Canny	Gabor	3-Cluster	5-Cluster	10Cluster
Homo	0.834687	0.708989	0.887475	0.868458	0.718358	0.709405
Contrast	0.507937	6.601098	0.245139	0.291554	1.200992	1.282562
DisSim	0.359368	1.231889	0.228398	0.267829	0.667259	0.689395
Homo	0.834825	0.708679	0.884891	0.879870	0.741378	0.718552
Contrast	0.503582	6.884240	0.250270	0.272898	1.074686	1.214865
DisSim	0.358468	1.240163	0.233560	0.245699	0.606618	0.666949
Homo	0.836484	0.707741	0.885288	0.884205	0.718849	0.716366
Contrast	0.491762	5.936572	0.248738	0.295027	1.161911	1.182490
DisSim	0.353829	1.192079	0.232643	0.242162	0.664529	0.687425

On the other hand, Canny edge detection leads to highest contrast and dissimilarity, followed by kernel fuzzy C-Means clustering (10-Cluster), kernel fuzzy C-Means clustering (5-Cluster), source images and kernel fuzzy C-Means clustering (3-Cluster), while Gabor edge detection leads to the lowest

contrast and dissimilarity, corresponding to smallest intensity difference and spatial difference. Kernel fuzzy C-Means clustering will produce smaller intensity difference and spatial difference than that of source images with the low cluster number, but larger intensity difference and spatial difference than that of source images with the large cluster number.

IX. Quantitative Analysis of HSI Model

The RGB model can be converted to the HSI model directly where the intensity I of a HSI model can be simply obtained by taking the intensity average out of three components R, G, B in the RGB model via I=(R+B+G)/3.

G. Dynamic Range

Dynamic range is equal to a ratio of the maximal intensity level measurable to the minimal detectable intensity level in digital imagery. By convention, the maximal intensity is relevant to saturation and the minimal one can be reached at the noise level. Dynamic range has been formulated as a logarithm function in (24), where I_{sat} shows the saturation intensity and I_{min} shows the minimum detectable intensity.

$$DR = \log_2 (I_{sat}/I_{min}) \quad (24)$$

Since conclusions for dynamic range comparisons among all 3 source images and corresponding feature detection outcomes are identical to each other, outcomes from one case are listed in context below without loss of generality. In general, kernel fuzzy C-Means clustering (3-Cluster: 3.973) is accompanied by the narrowest dynamic range, followed by Canny Edge detection (6.579), Gabor Edge detection (6.771), kernel fuzzy C-Means clustering (5-Cluster: 9.543), and kernel fuzzy C-Means clustering (10-Cluster: 9.570), while source images have the broadest dynamic range (9.579). Thus, edge based detection has the measurable intensity level within the region based detection with different cluster numbers. The source images contain the highest measurable intensity level.

In summary, neither edge based detection nor region based detection takes a dominating role in feature detection, from the quantitative analysis in multiple aspects. When the detection quality and computational complexity are both considered, the Gabor edge detection and kernel fuzzy C-Means clustering with an appropriate cluster number (e.g. No=5) can produce better overall feature extraction outcomes than others.

Conclusions

Feature detection in digital aerial imagery analysis has been presented in this study, in terms of typical edge based and region based schemes. Instead of qualitative visual evaluations, quantitative comparisons are proposed and also systematically conducted between outcomes from multiple edge detection and region detection approaches. The information metrics are selected in both spatial domain and frequency domain with respect to the RGB model or HSI model, including discrete entropy, discrete energy, mutual information, homogeneity, contrast, and dissimilarity as well as dynamic range. Numerical simulations have provided a plenty of novel outcomes. In frequency domain analysis, although edge detection outcomes show less dependency on source images than region detection outcomes generally, it is possible that the edge detection scheme (e.g. Gabor Filter) will produce outcomes with more information content than some region detection schemes. In spatial domain analysis, both edge detection and region detection could generate either higher or lower structural variations than source images. However, the role of edge detection is more significant where Gabor edge detection could generate lowest structural variations and contrast, but Canny edge detection could generate highest structural variations and contrast. In dynamic range analysis, the source image has the highest measurable intensity level, while edge based detection has the less measurable intensity level than the region based detection with a high cluster number. As a tradeoff, it is concluded from the context that Gabor edge detection and kernel fuzzy C-Means clustering with 5 clusters provide the relatively better overall performance in feature detection than others. These discoveries could be further extended to complex multispectral and hyper-spectral imagery analysis with high spectral resolution, where similar conclusions can be made.

References

[1] R. Gonzalez and R. Woods, “Digital Image Processing,” 2nd Edition, Prentice-Hall, 2002

[2] R. Schilling, S. Harris, “Fundamental of Digital Signal Processing using Matlab”, Cengage Learning, 2005

[3] Z. Ye, H. Cao, S. Iyengar and H. Mohamadian, "Medical and Biometric System Identification for Pattern Recognition and Data Fusion with Quantitative Measuring", Systems Engineering Approach to Medical Automation, Chapter Six, Artech House Publishers, 2008, pp. 91-112, Oct 2008

[4] X. He, W. Jia and Q. Wu, “An Approach of Canny Edge Detection with Virtual Hexagonal Image Structure”, 2008 10th International Conference on Control, Automation, Robotics and Vision, Hanoi, Vietnam, December 17–20, 2008

[5] Z. Ye, H. Mohamadian and Y. Ye, "Quantitative Analysis of Feature Detection Using Adaptive Canny Edge Detector and Enhanced Ant Colony Optimization", International Journal of Modeling and Optimization, pp.384-390, V2, No. 4, 2012, International Association of Computer Science and Information Technology Press

[6] T. Lee, “Image Representation Using 2D Gabor Wavelets”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 18, No. 10, 1996

[7] J. Arrospide, L. Salgado, “Log-Gabor Filters for Image-Based Vehicle Verification”, IEEE Transactions on Image Processing, Vol. 22, No. 6, pp. 2286-95, June 2013

[8] Z. Ye, H. Mohamadian, H. Yin and Y. Ye, "Practical Contour Tracing Via Integration of Adaptive Contrast Stretching and Gabor Wavelet Transform", Proceedings of 2014 Second International Conference on Advances in Computing, Electronics and Communication, pp. 93-97, Oct. 25-26, 2014, Zurich, Switzerland

[9] N. Senin, R. Leach, S. Pini, L. Blunt, “Texture-Based Segmentation with Gabor Filters, Wavelet and Pyramid Decompositions for Extracting Individual Surface Features from Areal Surface Topography Maps”, Measurement Science and Technology, v 26, n 9, (14 pp.), Sept, 2015

[10] S. Sajadi, and A. Fathi, "Genetic Algorithm Based Local and Global Spectral Features Extraction for Ear Recognition", Expert Systems with Applications, v 159, November 30, 2020

[11] P. Fazendeiro, J. Oliveira, "Observer-Biased Fuzzy Clustering", IEEE Transactions on Fuzzy Systems, v23, n1, p 85-97, Feb 2015

[12] Z. Ye and H. Mohamadian, "Enhancing Decision Support for Pattern Classification via Fuzzy Entropy Based Fuzzy C-Means Clustering", Proceedings of the 2013 52nd IEEE Conference on Decision and Control, pp. 7432-7436, December 10-13, 2013, Florence, Italy

[13] Z. Ye, H. Yin, Y. Ye, "Aerial Scene Classification and Information Retrieval via Fast Kernel Based Fuzzy C-Means Clustering", CH. 11, pp. 98-111, Information Management and Big Data: Communications in Computer and Information Science Book Series (Vol. 898), Springer International Publishing, Switzerland, January, 2019