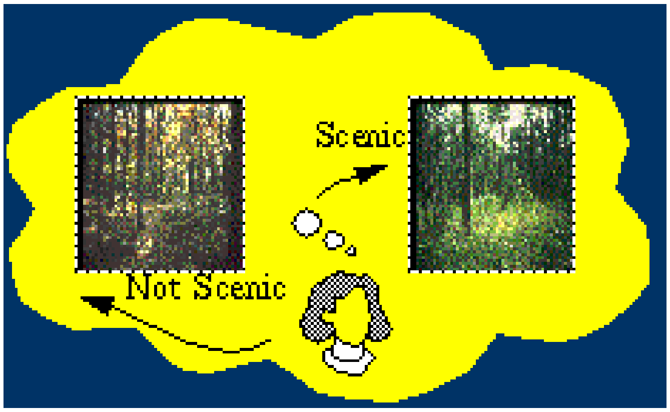
Scenic Beauty Estimation of Forestry Images

prepared for:

Southern Forest Experiment Station
United States Forest Service
201 Lincoln Green
Starkville, MS 39759-0906

December 15, 1998



submitted by,

N. Kalidindi, V. Ramani, and J. Picone
Institute for Signal and Information Processing
Department of Electrical and Computer Engineering
Mississippi State University
Box 9571
413 Simrall, Hardy Rd.
Mississippi State, MS 39762
Tel: (601) 325-3149
Fax: (601) 325-3149
email: {kalidindi, ramani, picone}@isip.msstate.edu

EXECUTIVE SUMMARY

This project involved developing an automatic algorithm to estimate the scenic beauty rating of the forestry images. Scenic beauty estimation (SBE) of forestry images is important to the United States Forest Service (USFS) for management of timber production and forest growth. Federal and regional agencies have also stressed the need for aesthetic evaluation to assess the impact of incentive programs in order to distribute public funds for forest management, and to plan timber harvest schedules on national forest land. The goal of this project is to use signal processing techniques to extract relevant features from these images, and to use this information to predict SBE.

A well-developed database, large enough to provide sufficient training and testing, was necessary for the evaluation of algorithmic performance. For these reasons, an extensive database was developed in conjunction with United States Forest Service. The images included in this database were drawn from a study spanning four years dealing with the Ouachita National Forest in Arkansas, U.S. Photographs taken under controlled conditions were digitized using an extremely high quality scanning process and converted into computer readable data. The database consists of 700 images, with images taken over two different sessions from 1990-91 and 1994-95 and sampled over all the seasons of the year at a number of different angles. Subjective beauty ratings are available for each of the images in the database. In this phase of the project, we have added an additional 1145 images, bringing the total to over 1900 images.

The scenic beauty (SB) of an image is highly dependent on the complexity of an image. Complexity is defined as the degree of variation derived from the visual qualities of an image. The features that were used as measures of the image complexity were the color, the density of the trees, the sharpness, the standard deviation of the intensity of the pixels, the compression ratio, the entropy of the pixel intensities and the fractal dimension of the image. We used a classification approach and regression analysis to statistically normalize the features extracted from the image and combine these features into a variety of feature sets. Different combinations of these features were used, including individual colors (red, green, and blue); colors combined with the output of a line detection algorithm (short lines and long lines); and colors combined with information theoretic measures such as entropy and fractal dimension.

A system using color, long lines, and entropy yielded our lowest overall classification error rate — 38.47%. This combination of features also had high correlation with the reference SBEs — 0.59. Most of the errors were observed to be in the regions of overlap between images rated as having medium to low scenic beauty and medium to high scenic beauty. This suggests that the SBE value as determined by human perception is somewhat imprecise and needs a more fundamental understanding and calibration. This analysis system suggests that the color and density of the trees plays a major role in estimating the scenic beauty of the forestry images. Future research is planned to augment the analysis with an ability to detect and quantify key objects within the image, such as trees, sky, and bushes using object recognition techniques. We believe that developing an understanding of the composition of the scene will be crucial to improving our ability to classify and analyze images.

TABLE OF CONTENTS

1. ABSTRACT

The aesthetic quality of forests in the U.S is actively managed by the United States Department of Agriculture and Forest Service (USFS). The rising public concern for preserving the beauty of the natural environment and the need to preserve the aesthetic quality of forests were responsible for the enactment of legislature to preserve the beauty of the forests. Traditional methods used to determine the scenic quality were very tedious and involved a large group of people manually rating each of the images. Our goal was to develop an algorithm which could automatically classify the images as having low, medium or high scenic value.

The primary factors which were known to relate to the scenic beauty were color content and image complexity. Some of the features which determined the complexity of the image were the density of the trees in an image, the entropy of the image, the sharpness of the image, the compression ratio, the standard deviation of the pixel intensity in the image and the fractal dimension of the image. The features extracted were compared to model files using a standard pattern matching paradigm. We also developed an extensive database in conjunction with the United States Forest Services to support the algorithm development. The database consisted of 637 unique images, each image having various subjective ratings such as the scenic beauty. The database extensively sampled several dimensions of the problem including year, season, time of day, angle, and treatment.

2. INTRODUCTION

Natural forests and wildlands are important sources of scenic beauty. Preserving their aesthetics has been a primary concern to the authorities managing these public forests. This required researchers to identify the features effecting scenic beauty and relating these features to public perceptions. The challenge was to extract these features from the image and combine them using statistical normalization techniques and relate them to the scenic beauty.

2.1. Historical Background

The United States Forest Service (USFS) was required to manage the forest land. Due to the increasing public concern for preserving the aesthetic quality of forests, the USFS was required to identify the scenic quality of the forest areas so that they could plan to manage the forests and simultaneously maintain the scenic beauty of the area. Legislation also encouraged maintaining the forests for recreational use. The Multiple-Use Sustained-Yield Act of 1960 required that national forests be managed for the full range of forest products as well as outdoor recreation. The National Environmental Policy Act of 1969 required that federal agencies identify and develop methods and procedures that require appropriate consideration to be given to the aesthetics in decision making. This gave rise to a lot of think tank for automatically determining the scenic beauty content of a given image and this project has come a long way from the time traditional methods were used for scenic beauty estimation to the present where the scenic beauty estimation can be done automatically.

2.2. Problem Perspective

Determining the scenic beauty of forestry images required identification of the features which contributed to the scenic beauty of the image.Signal processing techniques were used to relate features extracted from the images to people's perceptions. Human perceptions were made available through a survey which required participants to rate the scenic beauty of the images. To analyze people's preferences, let us look at the two images in Figure 1. One of them was rated as having low scenic beauty and the other as having high scenic beauty by the public.

SBE = -122.31

SBE = 84.63

Figure 1. a low scenic and a high scenic image

The feature which primarily differentiates the low scenic from high scenic image is the color content. The high scenic beauty image is darker, having lower mean values of the colors compared to the low scenic beauty image. The low scenic beauty image has more light penetrating through the forest area. The low scenic beauty image also has random distribution of the trees and a number of short bushes. This relates to the ratings in that people typically prefer visually penetrative images – images which have more long trees and fewer short bushes.

The density of the trees and bushes in the image was estimated by computing the vertical lines in the image and the length of each of the vertical lines. Randomness in the image was also an important feature in a person's perception of scenic beauty. Thus, we selected features relating to the color, visual penetration and randomness of an image for use in modeling the scenic beauty of an image.

The traditional methods used for determining scenic quality involved assessment from forest managers and evaluation by the general public. Two main approaches were used in these studies for determining scenic beauty: 1.Descriptive Inventories 2. Public Preference Models. Each of these approaches has been presented here. Both these techniques describe the features contributing to the estimation of scenic beauty but neither of these attempts to extract the features.

2.3. Descriptive Inventories

Descriptive inventories is the largest used method for assessing scenic resources. They comprise of the professionals directly involved in managing forestry resources. Descriptive inventories mainly involve describing the landscape attributes which contribute to scenic beauty. There are typically two methods for doing this: non-quantitative and quantitative methods.

Non-quantitative measures are given by professional landscape designers who describe scenic beauty verbally or graphically in terms of design components such as color, contrast, dominance and depth of the field. The drawbacks are that there are no numerical weights assigned to each of these features, and the final decision about the scenic beauty of an image is left to the professional. These measures provide a detailed description of the scenic areas but fail to give a measure of the general public perceptions. The untrained observer may disagree with the professional perceptions, but it is the general public that ultimately evaluates the scenic beauty of the land. For instance, the public rated diseased trees as more picturesque [1] and thus preferred these areas over areas with healthy trees, while the professionals do not prefer to have diseased tress in the area. This method can offer a simple and inexpensive analysis of estimating the scenic beauty but it depends largely on the attitudes, perceptibility, and experience of the evaluators.

Quantitative methods are an improvement over the non-quantitative analyses since each of the factors contributing to scenic beauty is represented by a numerical weight allowing objective comparisons between results. The advantage of quantitative methods is that the relative contribution of various landscape factors to scenic beauty can be indicated by weights. The features may have both positive and negative impact on the scenic beauty, so without relative weighting of those features, it is difficult to determine which feature add to the scenic beauty and which features reduce the scenic beauty. This method is expensive, time consuming and complex, as it is difficult to weigh all the features on the same scale. Descriptive inventories played an important role [1] in introducing aesthetic criteria but methods which represent public preferences and which may be more directly applicable to scenic planning are desirable.

2.4. Public Preference Models

Increasing concern among the public to preserve aesthetic resources resulted in the development of scenic assessment models based on input from the general public. This model represents a systematic representation of public preferences for scenic environments. As with the descriptive inventories, there are non-quantitative and quantitative methods for public preference models.

The most commonly used non-quantitative method is the questionnaire or verbal survey. Researchers form a questionnaire based on their perceptions and distribute the questions to a group of participants. This method is straight forward and requires little time and equipment. Questionnaires are a valuable source of quick information but accuracy is generally sacrificed for speed. Questions must be clearly and precisely stated. Open-ended questions have the advantage of allowing the expression of opinions, researchers may have overlooked. The disadvantages of open-ended questions is the lack of precision and clarity. A carefully constructed questionnaire demands an expenditure of time and money, and is an art in itself.

The other shortcoming of the survey method is the possible misinterpretation of public preferences. The flexibility of the language permits innumerable ways of expressing the same opinion. The use of various descriptions can lead to disagreements among the observers when they actually agree in essence. Also different wordings of the multiple choice questions leads to conflicting responses. For example, respondents may tell that they prefer “small forest clearings” but do not prefer “clear-cut patches in the forest”, although both of them mean the same. This makes the preparation of the questionnaire itself a very difficult process.

The verbal surveys failed due to the open ended questionnaires which resulted in conflicting responses. Also, it is difficult to translate the information into a quantitative form. An improvement over this methodology is the quantitative model which objectively analyzes verbal communications. Observers indicate their preferences for various visual attributes of environment on a scale of 1 (very low scenic quality) to 10 (very high scenic quality). These ratings can be standardized and adjusted to remove any bias. The discrepancies among the ratings of the different groups is adjusted by using a standard set of images and using a software called RMRATE to obtain the scenic beauty estimate. Although this method has considerable advantages, it is difficult to ascertain why one scene is rated higher than the other. Also it does not show the relationship of the features of the landscape to the overall scenic beauty.

The traditional approaches are neither complete nor robust. Public preference inventories represented the public preference to a scene, and the descriptive inventories the landscape factors effecting the scenic beauty; but there was no method which could relate the public perceptions to the landscape attributes. This required that scenic beauty determination involved the identification of features which relate to the public perceptions and establish their relationship to the scenic beauty estimate.

3. SCENIC BEAUTY DETERMINATION

Scenic beauty measure can be best described as a measure of an individual’s preference for the visual attributes of an image. The goal of this project was to determine the features which relate to the beauty of the forest, extract them and combine these features using suitable statistical techniques to relate them to the scenic beauty.

Scenic beauty is intimately related to the physical elements of the landscape. Contrast and variety were identified as important components of beautiful landscapes. Variety is also commonly referred to as “complexity”. Complexity is defined as the degree of variation derived from the visual qualities. Some of the factors affecting the complexity of the image were identified as color, variation of the intensity in the image, and the roughness of the image. Research at the other centers [15] focused on studying the relationship of complexity and scenic beauty. Their study showed complexity to be directly related to the scenic beauty, in that more complex areas can provide greater variety.

Studies support that perceived beauty of landscapes increases with complexity. However, too much of variety was found to result in negative response from the public [1]. Researchers were unable to find any systematic relationship between variety and esthetic preference. Thus complexity alone may not be sufficient to describe landscapes but the context of the variety and the elements comprising it may be better predictors of scenic preference. From the Complexity explains about 48 - 61% of the variance [1] of the preferences.

Color and complexity are the two main features effecting the scenic beauty. Some of the factors describing the complexity of the image are standard deviation, fractal dimension, entropy, sharpness and compression ratio. Standard deviation is the square root of the average squared deviation from the mean. Entropy, sharpness and compression ratio provide a measure of the randomness of the image and fractal dimension is a measure of the texture of the image.

Color is the most visually striking part of any image. Standard deviation and sharpness give a measure of the variation of the intensity of the pixels in the image. Entropy is a measure of the randomness within the image and fractal dimension is a measure of the texture of the image. Together these factors describe the complexity of the image. Algorithms were developed to extract each of the features. The most important component of this work is developing a technique to combine these extracted features using suitable statistical techniques to relate them to the scenic beauty.

In our study, two methods were used for relating the extracted features to the scenic beauty ratings: the classification approach and the regression analysis approach. In the classification approach, an attempt was made to classify the images as LSBE (low scenic beauty), MSBE (medium scenic beauty) or HSBE (high scenic beauty) rather than obtaining an absolute value for the SBE. The error performance is computed by the number of images misclassified. In the regression analysis, a linear combination of the extracted features is used to obtain an estimate of the scenic beauty. The correlation of the objective scenic beauty estimate to the subjective beauty estimate is found. The attempt is to obtain as close a value to the human judgement as possible.

4. ALGORITHMS

Algorithms described in this section were chosen based on their dependence to the complexity of the image. Color and complexity are the two most important features on which the scenic beauty depends. The features used to determine the complexity are entropy, standard deviation, sharpness, fractal dimension and compression ratio. The density of the trees was also selected as a feature as people typically prefer visually penetrative images [3].

4.1. Color

Color is the most visually striking feature of any image and it has a significant bearing on the scenic beauty of the image. The images in our database exist in Portable Pixel Map (PPM) format. The PPM format has each pixel represented by 24 bits, 8 bits for each of the color. The three primary colors in the image are red, green and blue. For studying the effect of colors, the distribution of the amplitude levels of each of the primary colors was computed. The minimum intensity value of each of the color was 10 and the maximum value was 255. This color range is divided into 10 bins, the center value of the first bin being 0 and the center value of the tenth bin being 255. Histograms of color intensities were generated and used as features. This gave us a total of 30 features related to color: 10 for red, 10 for green and 10 for blue. A sample distribution of a typical LSBE and HSBE image is shown in Figure 2.

Typically a LSBE image has a greater mean value of the colors as compared to a HSBE image. This indicates that most of the LSBE images are brighter than HSBE images. This trend can be accounted for by the fact that most LSBE images have trees cut and scattered on the ground. Due

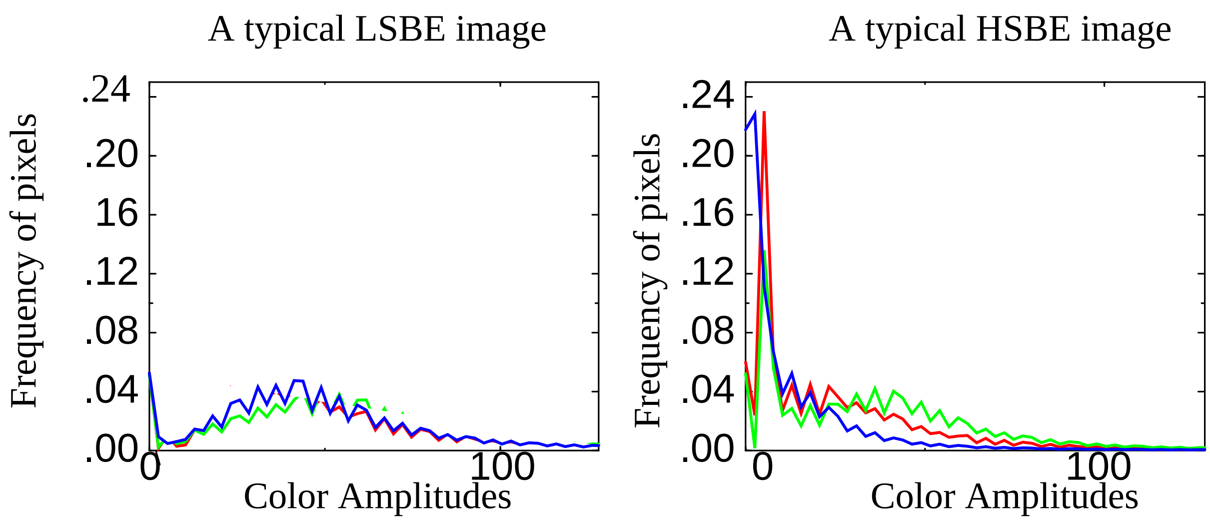


Figure 2. Typical distributions of LSBE and HSBE image

to the resulting “openness”, the image is bright but the scenic value is low due to the scattered hardwood. To study the contribution of each of the colors to the scenic value of an image, different combinations of features derived from color were evaluated. The combinations used were 1) red only, 2) green only, 3) blue only and 4) red, green and blue combined.

4.2. Edge Detection

Visual penetration is an important feature in determining the scenic beauty of an image. People prefer images that have good visual penetration, i.e. images where trees in the background can be seen with little obstruction have a high scenic beauty rating [3]. Randomly distributed trees and short bushes blocking the view of the forest are not considered as scenic. Density of trees in an image is a good measure of visual penetration. Edge detection is done to estimate the density of trees and bushes in an image. We chose the standard Canny edge detection algorithm for this purpose because it achieves a minimum localization error and error rate compared to other edge detection algorithms like Sobel edge detection and Roberts edge detection algorithms [17]. The block diagram for the Canny edge detector is shown in Figure 3.

The number of long trees and short bushes in an image is an indication of the density of trees in the image. The output of the edge detector was used to estimate this density of the trees and bushes. The edge detected output was fed to a line detector which was used to quantify the number of the lines in the image and the length of the lines. The distinction between the short bushes and the tall trees was made by assigning a threshold for the length of a line. After initial experimentation, the threshold in the line detector was fixed at 25 pixels. Any line with a length greater than the 25 pixels was considered as a long line (hence a tree) and any line whose length is less than the 25 pixels was identified as a short line (hence a short bush).

The Canny edge detection method involves smoothing of the image using a small Gaussian mask. A common Gaussian mask used is a 3 x 3 pixel map. The Gaussian mask reduced any noise present in the image. Gradients of the smoothed image in the x and y directions were obtained,

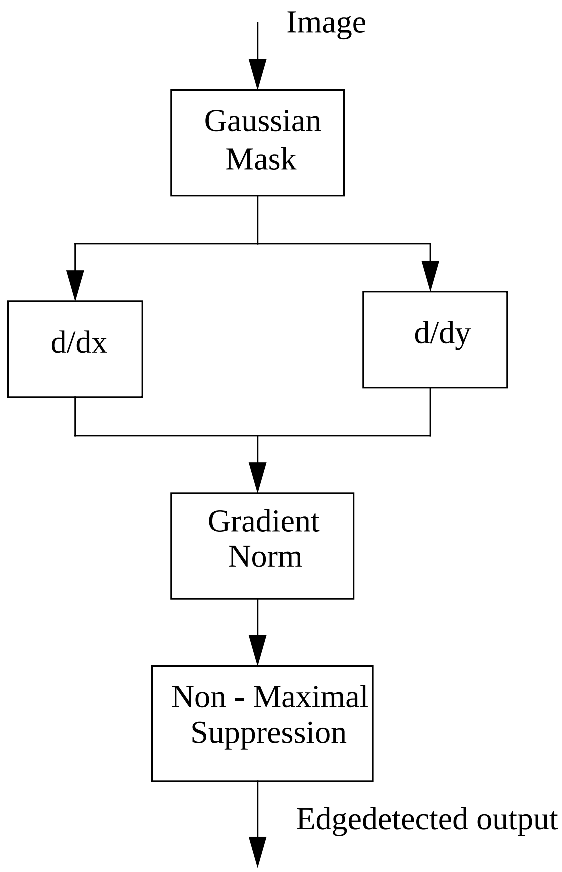


Figure 3. Block diagram of canny edge detect

and a euclidean norm of the gradients was computed. $Gradient = \sqrt{\left(\frac{d}{dx}\right)^2 + \left(\frac{d}{dy}\right)^2}$

This distinguished a pixel with an edge having a maximum value from the neighboring pixels having a smaller value. Non-maximal suppression was performed using the threshold techniques in which pixels below and above certain thresholds were kept while the rest were zeroed out. This gave the final edge detected output image. (1)

The edge detected output was then passed through a line detector in which the distinction between the long lines and short lines was made. The total number of the vertical lines, as well as the number of the short lines and the longlines were computed and used as features. The percentage of the long lines and short lines was used as a feature in evaluations. A sample image and its edge detected output are shown in Figure 4.

Figure 4. A sample image and its edge detected output.

4.3. Sharpness

Sharpness is a measure of the local variation of pixel intensities. It is computed as the frequency weighted sum of the magnitudes of difference in pixel to pixel intensity in the image. Mathematically, it is defined as

$$Sharpness = \sum_{for\,all\,m,\,n} \left(\sum_{i=-1}^{1} \sum_{j=-1}^{1} abs[x(m,n) - x(m-i,n-j)] \times f \right) \quad i,\, j \neq 0\,. \quad (2)$$

and f is the cumulative frequency of the bin in the intensity histogram. The histogram of the

Sharpness = 352.29, SBE = -122.31 Sharpness = 102.91, SBE = 84.63

Figure 5. Images with high and low values of sharpness

difference in pixel intensities gives a measure of amplitude variations in the image and hence a measure of sharpness of the image. This is found for each of the colors in the image. This process gave us three features -- sharpness of red, sharpness of green and sharpness of blue. Two images, one with the maximum sharpness in the database and one with minimum sharpness, are shown in Figure 5. The image with maximum sharpness was rated LSBE image by human subjects and the image with minimum sharpness was rated as a HSBE image.

Standard Deviation = 33.47, SBE = -4.61 Standard Deviation = 2.40, SBE = 54.07

Figure 6. Images with high and low values of standard deviation

4.4. Standard Deviation

While sharpness is a measure of local variation in pixel intensities, standard deviation is a measure of the global or overall variation of pixel intensities in the image. The standard deviation is computed as

$$STD = \sqrt{\frac{(x-\bar{x})^2}{N}}, \tag{3}$$

where x is the intensity of a specific color in the pixel and $\bar{x}$ is the mean intensity for that color in the image. The images with the maximum and minimum standard deviation in our evaluation database are shown in Figure 6. The image with a high standard deviation was classified as MSBE and that with low standard deviation as HSBE image.

The standard deviation for each of the colors was computed individually and the total standard deviation is the sum of the standard deviations of the three colors. Typically an which is bright due to penetrating sunlight has a higher standard deviation. On the other hand, an image which has less penetration by sunlight typically has a lower standard deviation. Images with high standard deviation were rated as MSBE images.

4.5. Entropy

Entropy is a measure of randomness in an image. It is represented mathematically as

$$Entropy = \sum_x p(x)\log p(x) \tag{4}$$

where x is the distribution in each of the bin and $P(x)$ is the probability of the distribution in each bin. The dynamic range of each color was divided into 10 bins and the pixel distribution was found. The entropy was computed from the probability of distribution in each of the bins. The total entropy was computed as the sum of the entropy due to each of the primary colors.

Entropy = 22.57, SBE = -122.31 Entropy = 15.77, SBE = -47.20

Figure 7. Images with high and low values of entropy

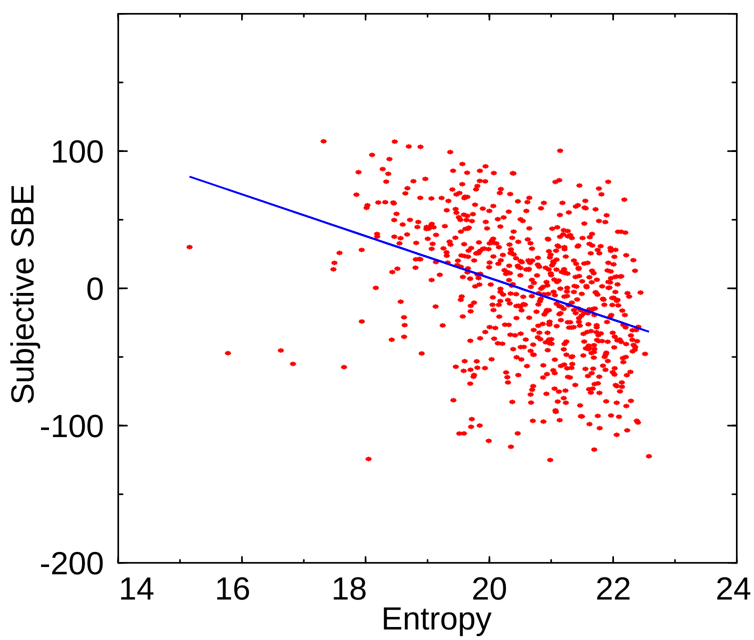


Figure 8. Scatter plot of entropy vs SBE

Typically, the more the randomness, less scenic the image. Images illustrating the effect entropy has on scenic beauty rating are shown in Figure 7. The image with the maximum entropy was classified as an LSBE image and the image with a low entropy as an MSBE image. A scatter plot of entropy vs subjective SBE is shown in Figure 8. A regression line was drawn between the entropy of an image and the subjective SBE. This gave the correlation of entropy with the subjective SBE. The correlation for this sample image was -0.373. The negative sign indicates that entropy is inversely proportional to the subjective SBE validating the observation that, the higher the randomness, the lower the scenic beauty rating.

4.6. Compression Ratio

Compression ratio of an image is another measure of the complexity of the image. Image compression is a technique which seeks to replace original pixel-related information with more compact mathematical representations. Compression ration is the ratio of the size of the original image to the size of the compressed image. High complexity images are less susceptible to compression and hence end up with a low compression ratio. JPEG coding is a widely used compression technique. It is a lossy compression technique using Huffman codes. Sample images with high and low compression ratios are shown in Figure 9.

Compression ratio = 7.77, SBE = -122.31   Compression ratio = 3.81, SBE = 84.63

Figure 9. Images with high and low values of compression ratio

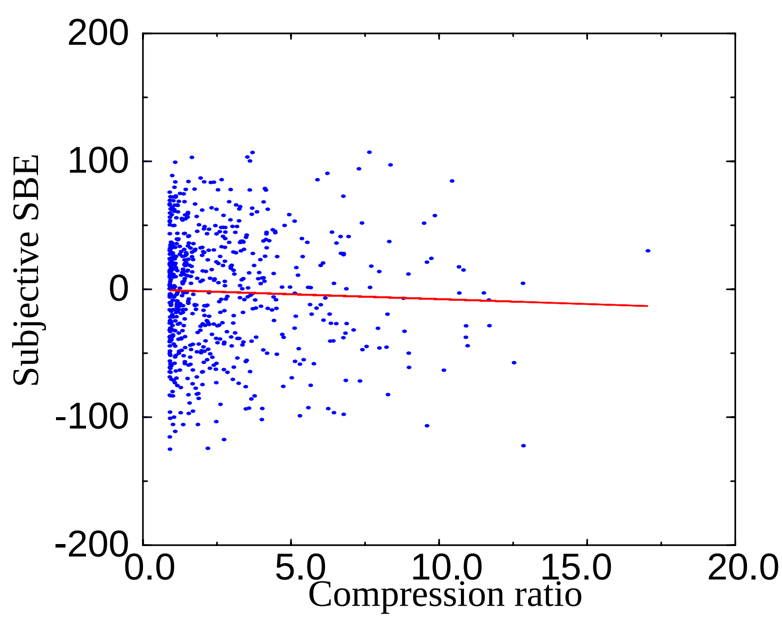


Figure 10. Scatter plot of compression ratio vs subjective SBE

The image with the high compression ration was rated LSBE and the image with the lower compression ratio was rated HSBE image. A regression fit is done with the compression ratio as the independent variable and the subjective SBE as the dependent variable. A scatter plot with the compression ratio on the x-axis and subjective SBE on the y-axis is shown in Figure 10. This plot shows that the variance in SBE values is much larger compared to the compression ratio. The regression fit is almost a flat line and the correlation is computed to be -0.038. This showed that though compression ratio is inversely proportional to the subjective SBE, the two attributes are not well correlated. This indicated that compression ratio may not be a good measure in estimating the scenic beauty of an image.

4.7. Fractal Dimension

Fractal geometry is a new language used to describe, model and analyze the complex forms of nature and fractal dimension is a measure of the texture of the image. We used a Triangular Prism Surface approach to compute the fractal dimension. It is illustrated graphically in Figure 11. In this method, a square region was chosen and was divided into four triangles with the center pixel as the common vertex for all the four triangles. In the figure this square region was represented by the region ABCD. The distance r is variable, minimum being 3 pixels. The common vertex in the figure is P. The sum of the areas of all the triangles is computed. This procedure is repeated for each of the pixel in the image and the sum of the areas is computed. This procedure is repeated for different values of r. This data can now be used to produce an area vs. distance plot, with distance r being the independent variable. A log-log plot is shown in Figure 12 with “r” on the x-axis and the area on the y-axis. The slope of the line gives a parameter “s”. The fractal dimension is related to s as,

$$Dimension, D = 2 - s \tag{5}$$

For color images, the fractal dimension for each of the color was found separately. Images with a high fractal dimension are considered to be more complex than images with a low fractal dimension. Sample images with high and low fractal dimension are shown in Figure 13.

5. STATISTICAL NORMALIZATION TECHNIQUES

The features extracted from the image are statistically combined to estimate the scenic beauty of the image. The two methods which are used for estimating the beauty of the images are classification approach and regression analysis. In the classification approach, models are built for each of the class of the images, i.e for LSBE, MSBE and HSBE and using the weighted distance measure, the image is classified into one of the class. In the regression analysis, the absolute value of the scenic beauty estimate is found.

5.1. Classification Methodology

The classification of images in the database is divided into LSBE, MSBE and HSBE according to the mean and the standard deviation of the subjective ratings. With the classification method, we classify the given test image into one of the classes using the extracted features. Model files are built for each of the class by training on a set of images. Training involves averaging the features

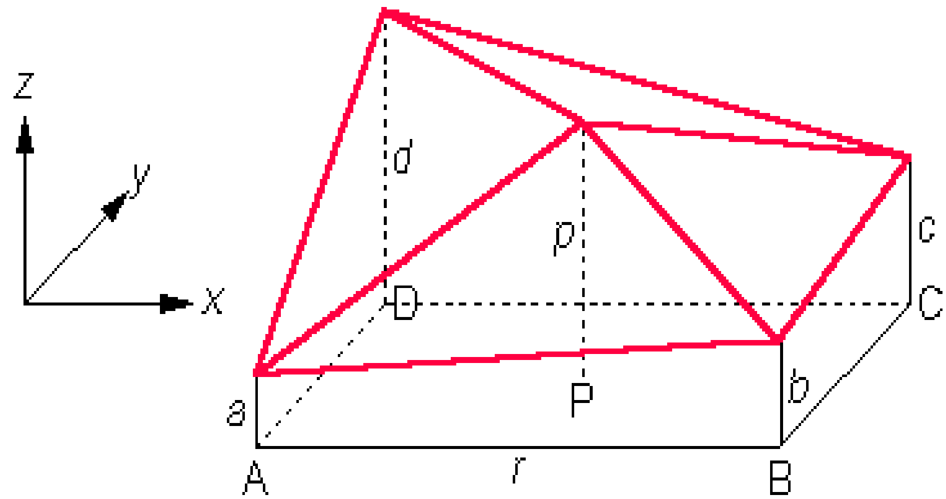


Figure 11. Triangular prism surface approach

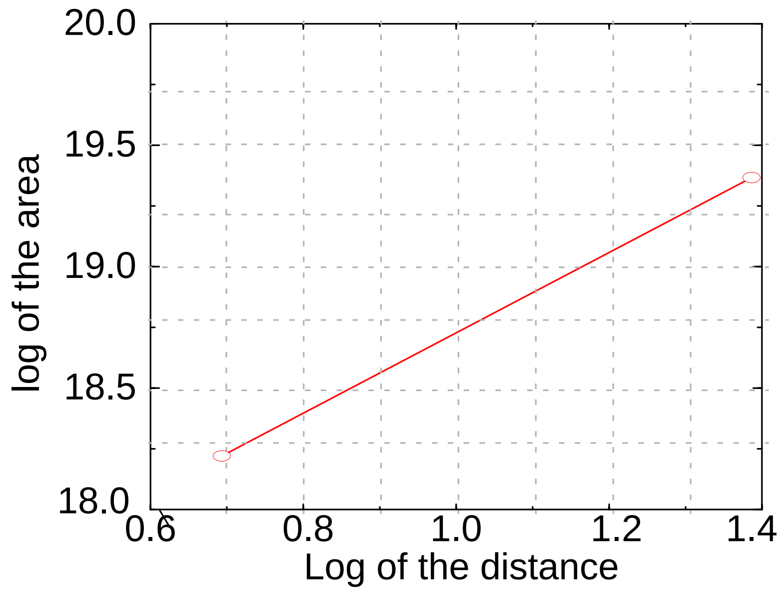


Figure 12. Plot of log of the distance vs log of the area

Fractal Dimension = 2.32, SBE = 31.20

Fractal Dimension = 0.14, SBE = -85.57

Figure 13. Image with high and low fractal dimension

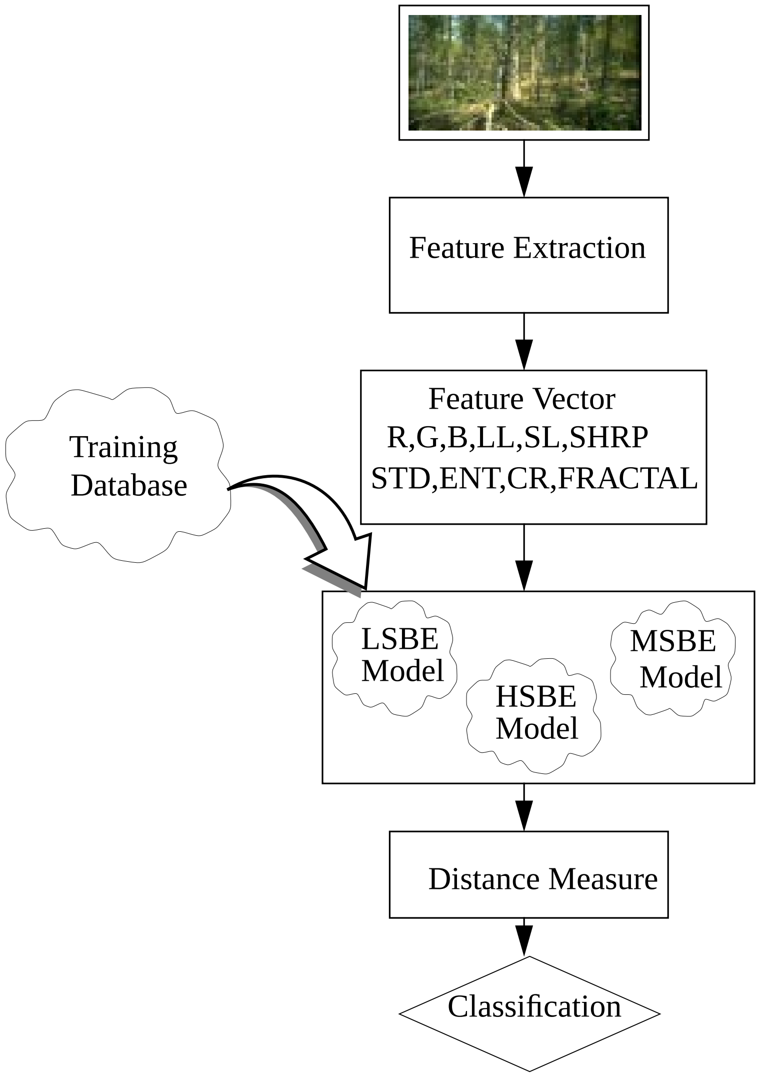


Figure 14. The classification approach using the models is shown.

of a set of images. The average feature vector along with the covariance matrix of the feature vector is written in the model file. In the test case, the distance of the feature vector of the test image is found from each of the class. The test image is assigned to the particular class to which it has the minimum distance. The distance can be found using two methods: 1) the non-weighted distance measure and 2) the weighted distance measure.

Assigning an image to a particular class is illustrated in the Figure 14. The distance of the test image from each of the classes is calculated and the image is assigned a particular class based on the minimum distance.

The euclidean distance or the RMS distance is a simple distance measure and is found by:

$$d = [x-y]^T[x-y] \tag{6}$$

x and y are the two vectors, the distance between which needs to be computed. This method is very simple but it has very serious drawback. It works only when all the features in the vector are on the same scale. In most real world problems, this is not the case. Features need to be preprocessed by whitening to get accurate results.

If we consider the features we extract from the image, some of the features are as small as less than one and some of them are as large as close to hundred. If we find the euclidean distance for such a vector, the distance will be dominated only by the larger feature in the vector. Hence this is not the true distance. We use the weighted distance measure in such cases which gives a better distance measure compared to the euclidean distance. The weighted distance is represented as

$$d = \sqrt{[x-y]^T C_x^{-1} [x-y]} \tag{7}$$

where x and y are the vectors and C_x^{-1} is the covariance matrix of vector x. The percent error in the classification approach is given by the number of images misclassified to the total number of images.

$\text{error\%} = \dfrac{\text{misclassified images}}{\text{total number of images}}$. Misclassification refers to a different classification of the test image as against the classification from the subjective scenic beauty estimate.

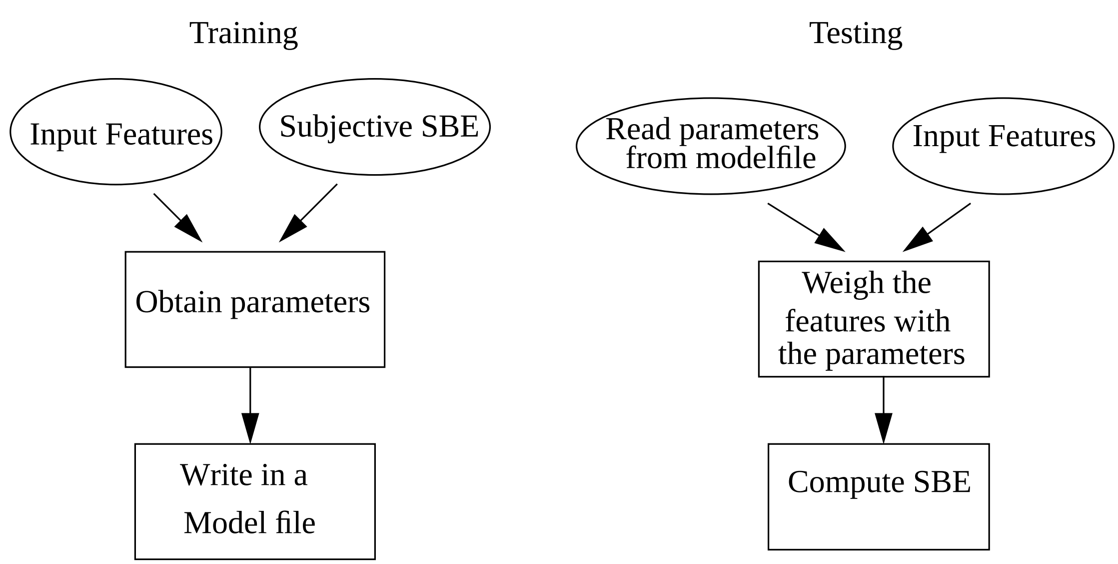


Figure 15. Summary of regression analysis

5.2. Regression Analysis

Another approach to map the extracted features to the scenic beauty is the regression analysis technique. Regression analysis is a technique in which the expected value of a dependent variable is modeled as a linear combination of a set of explanatory variables. Such a model is easy to analyze and applicable in many situations. Regression analysis is summarized in Figure 15. Consider an output y_i dependent on a set of variables $x_{1i}, x_{2i} \ldots x_{ki}$ This can be represented in a linear equation as

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots \beta_k x_{ki} + \varepsilon_i \tag{8}$$

for $1 \le i \le n$, the above equation can be written in matrix form as

$$Y = X\beta + \varepsilon \tag{9}$$

where Y is the n by 1 vector of observed values of the response variable. The feature matrix X is the n by (k+1) matrix containing the values of the input variable

$$.Y = \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ y_n \end{bmatrix} \qquad X = \begin{bmatrix} 1 & x_{11} & x_{21} & \ldots & x_{k1} \\ 1 & x_{12} & x_{22} & \ldots & x_{k2} \\ . & . & \ldots & \ldots & . \\ . & . & \ldots & \ldots & . \\ 1 & x_{1n} & x_{2n} & \ldots & x_{kn} \end{bmatrix} \tag{10}$$

As long as the input variables are not linearly related, this matrix equation can be solved to give the vector of parameter estimates. The estimate for β is chosen to minimize the mean squared error.

$$\hat{\beta} = (X^T X)^{-1} X^T Y \tag{11}$$

where T indicates transpose of the matrix. The $\hat{\beta}$ matrix has parameters or weights for each of the feature. The training images are used to get the values of β and then when testing, we can get an estimate of SBE by multiplying the feature vector with the coefficient vector.

The strength of the linear association between two processes is given by the correlation coefficient. It is used as a measure of performance for regression analysis. The value of the correlation coefficient is from -1 to +1. The closer the value to unity, irrespective of direction, the better the association between the two processes. The correlation of the derived scenic beauty estimate with the subjective scenic beauty estimate is computed. The performance is measured by the magnitude of the correlation obtained with the subjective SBE.

6. EVALUATIONS

The algorithms developed have to be evaluated to check the performance of the system in estimating the scenic beauty. Different combinations of the models were used to evaluate the algorithms. A database was developed for evaluating the algorithms. The purpose of the different combinations of the features was to identify the best combination which could predict the scenic beauty estimate with least error performance. The features were extracted from the image and the classification approach and regression analysis were used to combine these features. The database used for the evaluations and the error performances of these evaluations are described in this section.

6.1. Database

We have developed an extensive database in conjunction with the USFS to support the development of algorithms that will automatically estimate scenic quality. There are a total of 700 images in the database. The interesting feature of this database, in addition to the volume of raw data, is the inclusion of a number of measures computed by having human judgements manually assess the images. For example, subjective scenic beauty ratings on all images are available as part of the database. The images included in this database were drawn from a study spanning four years dealing with the Ouachita National Forest in Arkansas, U.S. Photographs taken under controlled conditions have been digitized using an extremely high quality scanning process, and converted into computer readable data. The database is extensively described in a separate document [16]. The database is summarized in the table below.

number of images	700 Images
Test Images	638
Baseline Images	40
Warm-up Images	20
Discarded Images	2
number of blocks	4
number of plots in each block	5
number of images in each plot	32
Seasons covered in each plot	Win, Sum, Spr, Fall
Fall	4
Summer	4
Spring	4
Winter	4
Number of files of 90-91 in each plot	16
Number of files of 94-95 in each plot	16
Images photographed at diff angles	
Number of LSBE images in the database	110
Number of MSBE images in the database	425
Number of HSBE images in the database	103

There are a total of 700 images out of which only 637 are valid images used for analysis. 40 of them are baseline slides used as standard slides and 20 are warm-up slides. The baseline slides are not used for analysis purposes. There are two slides which are discarded as the subjective SBE ratings of these images are not available.The images are photographed in the Ouachita forest range. The area is divided into four blocks for the purpose of studying, with each block divided into 5 plots. There are 32 images of each plot with each plot photographed four times in each of the four seasons. Also, the photo session was repeated in two years again in 94-95. This makes the total number of images of each plot as 32. The images were photographed at different angles.

The images in the database are divided into three classes, LSBE, MSBE and HSBE based on the mean and standard deviation of the subjective judgements. The mean SBE in the database is -2.19, and the standard deviation is 47.46. The SBE of each of the image is compared with the mean SBE and the standard deviation and the image is classified to be of one of the classes. This divides the database into 110 LSBE images, 425 MSBE image and 103 HSBE images.

6.2. Performance

The performance of the algorithm is a measure of the efficiency of the algorithm in estimating the scenic beauty of the images. A total of 45 features were extracted from the image and different combinations of these features were tried to identify the best system. The classification approach and regression analysis were used to combine the features. Various combinations of the features extracted were used for the evaluations. The total 45 features and their order which are used for evaluations are given in the table 1:

1-10	Red
11-20	Green
21-30	Blue
31	Longline
32	Shortline
33-35	Sharpness
36-38	Standard Deviation
39-41	Entropy
42	Compression Ratio
43-45	Fractal Dimension

Table 1: Table showing the order of the features used for evaluations

Different combinations of the features were tested to get the best model. All the evaluations are done on the official training and testing sets. Initially, we started with the RMS distance measure on the first training and test set. The combinations which were used are red only, green only, blue only, red, green and blue(rgb) combined, rgb with longlines, rgb with shortlines, and rgb with longlines and shortlines. The evaluations were done on the first training and test set initially.

modelfile	error% (RMS distance)	error% (weighted distance)
only red	63.75	51.87
only green	67.50	38.75
only blue	66.25	36.25
rgb	63.12	45.00
rgb+short	65.62	41.87
rgb+long	63.12	44.37
rgb+long+short	65.00	44.37

Table 2: Error performance for different model files using RMS distance

Model files were generated for each of the models by training on the LSBE, MSBE and HSBE files of the first training set. Then the images in the test set were evaluated to get the distance measure from each of the model class. The images were assigned to a particular class based on the minimum distance measure. The error is calculated from the number of images misclassified as a percentage of the total number of images tested. The evaluations for the same models were repeated with the weighted distance measure. During training, the mean vector and the covariance matrix of the vector are written in a model file. In the testing, the distance of the test vector is found from each of the class. The image is assigned to a class to which it has the minimum distance and the error is calculated from the number of images misclassified. The results for the weighted and non-weighted distance measure is given in Table 2. The weighted distance has better performance compared to the RMS distance. This proved the superiority of weighted distance over non-weighted distance measures. Hence, for further evaluation only weighted distance measures were used.

Weighted distance, being a better performance measure, was used for evaluations with the different feature sets. The first step was to run evaluations using the first training and testing set. The different models used were 1) red only, 2) green only, 3) blue only, 4) rgb combined, 5) rgb combined with shortines, 6) rgb combined with longlines, 7) rgb combined with short and longlines, 8) rgb combined with entropy, 9) rgb combined with longlines and entropy, 10) entropy combined with fractal dimension, 11) rgb combined with longlines, entropy and fractal dimension, 12) only entropy, 13) rgb combined with all the other features except the fractal dimension and finally 14) rgb combined with all the features including the fractal dimension.

Table 3 has the performance measure for both the training and testing files. The columns corresponding to “Training” are closed-loop tests in which the same data used for training is used for testing. On the other hand, the columns corresponding to “Testing” are open-loop tests. The table also shows the performance from regression analysis. For the same models, parameter estimators are found using the training images and the scenic beauty estimate for the testing images is found by weighting the feature vector of each image with the parameter estimators. The correlation is found for each of the model. The closed and open loop performance is estimated in this approach also. The “Std.err” is the mean square error obtained from the regression analysis approach.

	Training			Testing		
System	**%err**	**Corr**	**Std.err**	**%err**	**Corr**	**Std.err**
Red	49.37	0.458	41.66	51.57	0.349	46.22
Green	36.61	0.332	44.20	38.36	0.310	46.61
Blue	31.59	0.528	39.77	35.84	0.468	43.39
RGB	34.93	0.677	34.47	40.25	0.563	40.91
RGB+LL	34.10	0.683	34.20	38.99	0.566	40.86
RGB+SL	32.00	0.682	34.25	37.77	0.564	40.96
RGB+LL+SL	32.42	0.684	34.18	40.25	0.565	40.90
RGB+ENT	30.75	0.688	33.99	39.62	0.537	42.05
RGB+LL+ENT	30.12	0.693	33.77	42.13	0.600	39.56
ENT+FRACTAL	40.37	0.580	37.11	56.60	0.451	44.05
RGB+LL+ENT+FRCT	27.19	0.699	33.51	43.30	0.565	41.02
ENTROPY	49.37	0.546	39.25	49.68	0.409	45.14
RGB+ALL	25.73	0.732	31.88	42.10	0.600	39.56
RGB+ALL+FRCT	24.68	0.738	31.60	43.30	0.626	38.75

Table 3: Table showing the performance for classification approach and regression analysis

From this initial experimentation, three of the best-performing systems were selected which were then used on all the remaining training and test sets. The three systems which had the maximum correlation are RGB+LL+ENT(rgb combined with longlines and entropy), RGB+ALL(rgb combined with all the features except fractal dimension), RGB+ALL+FRACT(rgb combined with all the features including the fractal dimension). The remaining test sets were evaluated for the three above best-performing systems. The performance for all the sets and average performance is given in the table 4

From the results, the best system with classification approach is RGB+LL+ENT with an error performance of 38.47% and the best system with the regression analysis is RGB+ALL+FRACT having a correlation of 0.647. The confusion matrices for each of the above system is given in the Appendix A.

		Training			Testing		
		%err	corr	Std.err	%err	Corr	Std.err
RGB+LL +ENT	Set1	30.12	0.693	33.77	42.13	0.540	42.00
	Set2	29.16	0.667	35.38	43.03	0.629	37.53
	Set3	28.45	0.688	35.06	32.50	0.584	37.16
	Set4	24.47	0.677	34.52	36.25	0.610	38.86
		28.05	0.681	34.68	38.47	0.590	38.88
RGB+ALL	Set1	25.73	0.732	31.88	42.10	0.600	39.56
	Set2	29.79	0.712	33.34	40.50	0.677	35.21
	Set3	30.12	0.736	32.67	41.87	0.621	36.02
	Set4	19.24	0.726	32.25	34.37	0.651	37.32
		26.22	0.726	32.53	39.71	0.637	37.02
RGB+ALL +FRCT	Set1	24.68	0.738	31.60	43.30	0.626	38.75
	Set2	28.33	0.720	32.96	43.03	0.704	33.85
	Set3	33.47	0.759	31.45	45.00	0.592	37.27
	Set4	18.41	0.739	31.60	33.75	0.667	36.89
		26.22	0.739	31.90	41.27	0.647	36.69

Table 4: Table showing the performance of the best systems on all the test sets

6.3. CONCLUSIONS

We have developed an algorithm for estimating the scenic beauty rating of forestry images. This involved developing various algorithms to extract features from the images and using these features to model the subjective scenic beauty estimates of the images. The features that we chose are color, density of the trees, entropy, sharpness, standard deviation, compression ratio and fractal dimension. We also developed an extensive database in conjunction with the United States Department of Agriculture (USDA). The database is well organized with the file names explaining the block number, plot number, treatment and the year, day and month in which the image was photographed,

We have conducted various experiments using both classification approach and regression analysis to find the model which gives the best performance. The best system achieved with the classification approach was RGB+LL+ENT (red, green, blue combined with longlines and entropy) giving a classification error of 38.47%. The best system we achieved using regression analysis and all the features extracted from the image yielded a correlation of 0.65.

Specifically, we tried both weighted distance and non-weighted distance measure for the classification approach. Weighted distance had far better performance compared to the non-weighted distance measure. Sharpness, entropy,standard deviation and compression ratio were found to be inversely proportional to the scenic beauty. A higher value for any of these in the image reduced the scenic beauty. Fractal dimension was found to have a positive correlation with scenic beauty estimate and a higher value of fractal dimension indicated a higher scenic beauty rating. It was also found that compression ratio is not a very good feature for classification. This was verified by the correlation of the compression ratio with the subjective beauty estimate.

An encouraging result was that most of the misclassifications that we obtained was along the LSBE/MSBE border or along the MSBE/HSBE border. There was a very small error in the extremes of the classes; that is errors between LSBE and HSBE. The publications generated during the course of the project, the report for the project and all the software developed during the course of this project is available in the public domain at: http://www.isip.msstate.edu/resources/technology/projects/1997/sbe_imaging/.

7. FUTURE WORK

In this work we have used principal components analysis, also called the weighted distance measure, for classifying images. Principal components analysis(PCA) performance depends on the type of data we are using. If there is a large separability between the classes, principal components analysis works well. However, if this is not the case, PCA fails. Our future research efforts will be directed towards implementing Linear Discriminant Analysis(LDA), Support Vector Machines and Decision Trees, the developing classification algorithms in the field of signal processing. Linear Discriminate Analysis (LDA) is a better approach when the separability between the data is not very large. LDA tries to maximize the inter-class distance and minimize the intra-class distance so that the errors in classifying the images will be minimized.Support Vector Machines is another technique for image classification

We are also planning to add the frequency domain information in the feature extraction algorithms for object recognition. Additionally, an analysis of variance of each of the features extracted will be presented.

8. ACKNOWLEDGEMENTS

This work was supported by United States Department of Agriculture(USDA) through the United States Forest Service (USFS) Southern Research Station under cooperative agreement No. SRS-30-CA-96-049. I would like to express my appreciation to the constant guidance provided by Dr. Joseph Picone during the past one and half years for the development of this project. I would like to thank Dr. Victor Rudis for his continuous support and cooperation in this project. I would like to thank Aravind Ganapathiraju and Neeraj Deshmukh for their help during all phases of this work, particularly in the algorithm development. Finally, I would like to thank Jonathan Hamaker, Rick Duncan and other fellow members of ISIP for their support during the course of this project.

9. REFERENCES

[1] L. M. Arthur, T.C. Daniel, and R. S. Boster, “Scenic Assessment: An overview”, in *Landscape Plan*, vol. 4, pp. 109-129, 1977.

[2] G. J. Buhyoff, P. A. Miller, J. W. Roach, D. Zhou, and L. G. Fuller, “An Artificial Intelligence Methodology for Landscape Assessments”, in Artificial Intelligence Applications, vol. 8, pp. 1-14, 1994.

[3] T. A. Herrick, and V. A. Rudis, “Visitor Preference for Forest Scenery in the Ouachita National Forest,” *Proceedings of the Symposium on Ecosystem Management Research in the Ouachita Mountains: Pretreatment Conditions and Preliminary Findings*, pp. 212-222, Hot Springs, Arkansas, U.S., October 1993.

[4] S. Nicola, and J. Malouff, “Preference for Complexity in Natural Landscape Scenes”, *Perceptual and Motor Skills*, pp. 109-110, 1986.

[5] G. J. Buhyoff, and L. G. Fuller, “Explanation of Quantitative Models: A Need and an Example”, *Artificial Intelligence Applications, vol. 7*, pp. 37 - 43, 1993.

[6] G. J. Buhyoff, W. B. White, T. C. Daniel, and D. O. Hunter, “Integrated Computer Decision-Support for Forest Impact Assessment: A Conceptual Framework and Example”, *Artificial Intelligence Applications*, vol. 2, pp. 37-42, 1989.

[7] V. A. Rudis, J. H. Gramann, and T. A. Herrick, “Esthetics Evaluation”, *Proceedings of the Symposium on Ecosystem Management Research in the Ouachita Mountains: Pretreatment Conditions and Preliminary Findings*, pp. 202-211, Hot Springs, Arkansas, U.S., October 1993.

[8] V. A. Rudis, “Sampling and Modelling Visual Component Dynamics of Forested Areas,” *Proceedings of the Symposium on State-of-the-Art Methodology of Forest Inventory*, pp. 84-85, Portland, Oregan, U.S., July 1989.

[9] J. H. Gramann, and W. Yhang, *The Effect of Forest Color on the Perceived Scenic Beauty of Pine-Oak Plots in the Ouachita National Forest, Arkansas*, Ph.D. Dissertation, Texas A&M University, College Station, Texas, U.S., December 1994.

[10] T. C. Brown and T. C. Daniel, “Scaling of Ratings: Concepts and Methods,” Research Paper RM-293, U.S. Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experiment Station, Fort Collins, Colorado, U.S., 24 p., September 1990.

[11] V. Rudis, Personal Communications, January 8, 1997.

[12] N. Kalidindi, A. Lee, L. Zheng, H. Yaquin, J. Picone, and V.A. Rudis, “Scenic Beauty Estimation of Forestry Images,” *Proceedings of IEEE Southeastcon*, Roanoke, Virginia, pp. 337 - 339, Blacksburg, Virginia, USA, April 1997.

[13] A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice-Hall International Inc., Englewood Cliffs, NJ, USA, 1989.

[14] V. Rudis, James H. Gramann, and Edward J. Ruddell, "Visual Component Sampling to Monitor Scenic Beauty, Disturbances, and Understory Stand Dynamics, with Examples from Extensive-Area Forest Inventories" *internal report*.

[15] *http://www.imlab.uiuc.edu/complex/index.html#Defining.*

[16] N. Kalidindi, A. Le, and J. Picone, "Scenic Beauty Estimation Database", *prepared for Southern Forest Experiment Station*, United States Forest Service.

[17] J. Canny, "A Computational Approach to Edge Detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679-698, November 1986.

[18] V. A. Rudis, J. H. Gramann, and T. A. Herrick, "Esthetics Evaluation", *Proceedings of the Symposium on Ecosystem Management Research in the Ouachita Mountains: Pretreatment Conditions and Preliminary Findings*, pp. 202-211, Hot Springs, Arkansas, U.S., October 1993.

[19] M. Robert, B. Suresh, "Scenic Beauty Estimation Using Linear Discriminant Analysis", *DSP Project Paper*, December 1997.

[20] J. H. Gramann, W. Yhang, *The Effect of Forest Color on the Perceived Scenic Beauty of Pine-Oak Plots in the Ouachita National Forest, Arkansas*, Ph.D. Dissertation, Texas A&M University, College Station, Texas, U.S., December 1994.

APPENDIX A

Confusion matrix of the different evaluations for the testing files

system	L to L	L to M	M to H	error%
red only	14	42	2	51.57
	10	47	10	
	4	14	16	
green only	8	16	0	38.36
	15	78	16	
	5	9	12	
blue only	5	13	0	35.84
	19	82	13	
	4	8	15	
rgb	8	2	0	40.25
	14	72	13	
	6	29	15	
rgb+long	6	3	0	38.99
	16	76	13	
	6	24	15	
rgb+long+short	6	2	0	40.25
	19	77	16	
	3	24	12	
rgb+short	8	2	0	37.77
	16	77	14	
	4	24	14	
rgb+entropy	8	2	0	39.62
	17	75	15	
	3	26	13	
rgb+ll+entropy	5	2	0	42.13
	20	75	16	
	3	26	12	

Table 5: Confusion matrix for the various models of the test set 1

system	L to L	L to M	M to H	error%
entropy+fractal	3	13	1	56.6
	20	59	20	
	5	31	7	
rgb+ll+ent+fractal	10	17	1	43.3
	17	25	22	
	1	11	5	
entropy	10	21	4	49.68
	15	69	23	
	3	13	1	
rgb+all	18	27	5	42.1
	10	70	19	
	0	6	4	
rgb+all+fractal	10	17	1	43.3
	17	75	22	
	1	11	5	

Table 5: Confusion matrix for the various models of the test set 1

system	L to L	L to M	M to H	error%
rgb+ll+ent	10	22	2	43.03
	17	76	21	
	0	6	4	
rgb+all	9	11	1	40.5063
	15	67	8	
	3	26	18	
rgb+all+fractal	12	19	2	43.03
	15	69	16	
	0	16	9	

Table 6: Confusion matrix for the three best models of the test set 2

system	L to L	L to M	M to H	error%
rgb+ll+ent	4	0	0	32.5
	26	97	13	
	1	12	7	
rgb+all	2	0	0	41.875
	28	80	9	
	1	29	11	
rgb+all+fractal	0	0	0	45.00
	23	83	15	
	8	26	5	

Table 7: Confusion matrix for the three best models of the test set 3

system	L to L	L to M	M to H	error%
rgb+ll+ent	0	4	0	36.25
	21	94	20	
	2	11	8	
rgb+all	0	1	0	34.37
	23	101	24	
	0	7	4	
rgb+all+fractal	0	1	1	33.75
	23	103	24	
	0	5	3	

Table 8: Confusion matrix for the three best models of the test set 4

APPENDIX B

Confusion matrix of the different evaluations for the training files

system	L to L	L to M	M to H	error%
red only	49	117	12	49.37
	27	150	20	
	5	55	43	
green only	17	22	2	36.61
	56	259	46	
	8	41	27	
blue only	38	26	2	31.59
	41	253	37	
	2	43	36	
rgb	24	7	0	34.93
	42	233	21	
	15	82	54	
rgb+long	26	7	0	34.10
	42	236	22	
	13	79	53	
rgb+long+short	24	7	0	32.42
	46	244	20	
	11	71	55	
rgb+short	25	6	0	32.00
	42	245	20	
	14	71	55	
rgb+entropy	21	4	0	30.75
	48	256	21	
	12	62	54	
rgb+ll+entropy	22	4	0	30.12
	52	261	24	
	7	57	51	

Table 9: Confusion matrix for the various models of the training set 1

system	L to L	L to M	M to H	error%
entropy+fractal	21	40	2	40.37
	51	212	21	
	9	70	52	
rgb+ll+ent+fractal	17	0	0	27.19
	61	279	23	
	3	43	52	
entropy	24	62	1	49.37
	48	164	20	
	9	96	54	
rgb+all	64	63	5	25.73
	17	252	31	
	0	7	39	
rgb+all+fractal	63	51	5	24.68
	18	269	42	
	0	2	28	

Table 9: Confusion matrix for the various models of the training set 1

system	L to L	L to M	M to H	error%
rgb+ll+ent	20	7	0	29.16
	62	310	66	
	1	4	10	
rgb+all	9	1	0	29.79
	72	288	36	
	2	32	40	
rgb+all+fractal	56	59	9	28.33
	26	248	27	
	1	14	40	

Table 10: Confusion matrix for the three best models of the training set 2

system	L to L	L to M	M to H	error%
rgb+ll+ent	23	9	1	28.45
	52	279	42	
	4	28	40	
rgb+all	20	3	0	30.12
	51	250	19	
	8	63	64	
rgb+all+fractal	8	9	3	33.47
	58	252	22	
	13	55	58	

Table 11: Confusion matrix for the three best models of the training set 3

system	L to L	L to M	M to H	error%
rgb+ll+ent	47	18	1	24.47
	38	275	35	
	2	23	39	
rgb+all	51	11	0	19.24
	36	294	34	
	0	11	41	
rgb+all+fractal	47	7	0	18.41
	40	301	34	
	0	7	41	

Table 12: Confusion matrix for the three best models of the training set 4

The confusion matrix can be used to analyze the error performance of the algorithm. The numbers in the column of the matrix indicate the images classified as LSBE, MSBE or HSBE by the human judgements and the numbers in the row are the classification of the images as LSBE, MSBE or HSBE by the algorithm. Consider an example. The confusion matrix for the rgb+ll+ent system is shown below:

rgb+ll+ent	47	18	1	24.47
	38	275	35	
	2	23	39	

The diagonal numbers are the images which are classified correctly, i.e, an image rated as LSBE by human judgements is also classified as LSBE by the algorithm. The off diagonal numbers are the images which are misclassified. The number in the first row, second column is an image which is classified as LSBE by the algorithm but rated as MSBE by the humans. The number in the first row, third column is an image which is classified as LSBE by the algorithm and rated as HSBE by the humans. The number in the second row, first column is an image which is classified as MSBE by the algorithm and rated as LSBE by the humans. The number in the second row, third column is an image which is classified as MSBE by the algorithm and HSBE by the humans. The number in the third row, first column is an image which is classified as HSBE by the algorithm and LSBE by the humans. The number in the third row, second column is an image which is classified as HSBE by the algorithm and MSBE by the image. In the above table, the error performance is 24.47%. 361 images are classified correctly and 117 images are classified incorrectly.