Automated Coloring of Panchromatic Orthoimagery

Simon Premože William B. Thompson
Computer Science Department
University of Utah
{premoze,thompson}@cs.utah.edu

Abstract

Realistic rendering of outdoor terrain requires both that the geometry of environment be modeled accurately and that appropriate texturing be laid down on top of that geometry. We describe an approach for automatic coloring of panchromatic aerial orthoimagery. The method is able to remove shading and shadowing effects in the original image so that shading and shadowing appropriate to variable times of day and year can be added. The method we present is based on pattern recognition principles. It classifies regions in the original panchromatic orthoimagery into classes that are subsequently colored with user selected color palette. The method requires very little user interaction and is robust. The user only needs to select a few training point for each class that are later used in the pattern recognition and classification step. We also present an alternative method that is even simpler and requires no user intervention.

Keywords:

1 Introduction

Realistic rendering of outdoor terrain requires both that the geometry of environment be modeled accurately and that appropriate texturing be laid down on top of that geometry.

1.1 The nature of the data

Geo-specific rendering of terrain requires information about both the geometry and the photometry of the scene. Raw information about the geometric shape of the terrain itself is most often available as a *Digital Elevation Model* (DEM), in which elevation values are represented in a rectangular grid. The highest resolution widely available elevation data for the continental U.S. are United States Geological Survey (USGS) 7.5-Minute DEMs [9]. Elevation values with a nominal precision of 1m are provided at 30m intervals (*post spacing*) on the ground. The 7.5-Minute DEMs are created by optically scanning contour maps and then fitting an approximation surface. They are subject to a number of systematic distortions that, depending on the technology used when a particular DEM was produced, can result in the actual resolvability of ground features being far worse than the 30m post spacing might suggest.

The most direct way to render geo-specific photometry is to start with an image of the area to be rendered. Perspective effects make it difficult to register conventional aerial imagery with elevation data. As a result, an *orthorectification* process is often performed, in which the perspective image is warped to remove the effects of lens projection, camera orientation, and terrain. The result is an image that is effectively in a scaled orthographic projection. The USGS provides 1m resolution panchromatic orthoimagery for much of the continental U.S. No comparable source for color orthoimagery exists. Aerial survey companies can produce such imagery on a custom basis, but the cost is significant. Satellite images are often used to render terrain. While not true color imagery as that term is commonly used, multi-spectral satellite data can be converted to an RGB format that closely approximates perceptual color. In addition, much work has gone into the classification of multi-spectral satellite data to determine properties such as vegetation cover. Unfortunately, the resolution of available multi-spectral satellite data is at best on the order of 20m on the ground [7, 1].

If actual imagery of the terrain being rendered is to be used, the available choices are usually limited to false-colored multi-spectral satellite data of limited resolution or USGS high-resolution panchromatic orthoimagery. Since visual realism in terrain rendering depends in part on high resolution texturing, it is important to explore whether or not panchromatic orthoimagery can be effectively utilized. Can realistic color be generated? Can we get enough information about ground cover to add detail not resolved in the imagery? Can we remove shadows and shading effects so as to simulate views at times other than when the original imagery was acquired?

1.2 Coloring of Grayscale Images

Values in grayscale images vary only along one dimension (intensity or luminance). The task of coloring a grayscale image involves assigning a quantity that varies in three dimensions (e.g., red and green and blue channels). A mapping between a color and luminance is not unique, because different colors can have the same luminance (intensity) but different hue and saturation. Colorization process of grayscale images is therefore ambiguous in nature and requires some amount of human interaction. In subsequent sections, we describe two methods of coloring grayscale orthoimages. The first method is based on the statistical pattern recognition. The second method employs transformation of images in a decorrelated color space where the statistical distribution between the source and target images is matched.

2 Statistical Pattern Recognition

We briefly review statistical pattern recognition. More detailed review can be found in Jain *et al.* [4]. Statistical pattern recognition is used to establish boundaries between patterns. A *pattern* is represented by a set of attributed or features. There are three steps in recognition:

- **Preprocessing** Noise and extraneous data is removed from the input data and pattern is normalized and segmented from the background.

- **Learning** For each pattern in the input data, appropriate features representing input patterns are found. The classifier is trained to partition the feature space of the input data.

- **Classification** Based on the measured features from the learning stage, the classifier assigns the input pattern to one of the pattern classes.

A given pattern has d features represented as a $d-$dimensional feature vector $\mathbf{x} = (x_1, x_2, \ldots, x_d)$. A statistical pattern recognizer assigns a given pattern to one of the c categories $\omega_1, \omega_2, \ldots, \omega_c$ based on a feature vectors. Features are assumed to have a probability density function [4] dependent on the pattern class. A pattern vector $\mathbf{x}$ belonging to a class ω_i can be viewed as an observation chosen randomly from the conditional probability function $p(\mathbf{x}|\omega_i)$. Many rules are available that define the decision boundary between different classes: the Bayes decision rule, the maximum likelihood rule, the Neyman-Pearson rule. According to the optimal Bayes decision rule, an input pattern vector $\mathbf{x}$ is assigned to a class ω_i for which the conditional risk

$$R(\omega_i|\mathbf{x}) = \sum_{j=1}^{c} L(\omega_i, \omega_j) \cdot P(\omega_j|\mathbf{x}) \tag{1}$$

is minimum. $P(\omega_j|\mathbf{x})$ is the posterior probability and $L(\omega_i, \omega_j)$ is the loss incurred in deciding ω_i when the true class is ω_j. If the loss function $L(\omega_i, \omega_j)$ is binary:

$$L(\omega_i, \omega_j) = \begin{cases} 0, & \text{if } i = j \\ 1, & \text{if } i \neq j \end{cases} \tag{2}$$

then the conditional risk $R(\omega_i|\mathbf{x})$ from equation 1 is the conditional probability of misclassification. For this particular loss function $L(\omega_i, \omega_j)$, the Bayes decision rule assigns an input pattern $\mathbf{x}$ to class ω_i if:

$$P(\omega_i|\mathbf{x}) > P(\omega_j|\mathbf{x}) \text{for all } j \neq i. \tag{3}$$

If all of the class-conditional densities are specified, then the maximum *a posteriori* rule can be used in a classifier. Unfortunately, the class-conditional densities are usually not known in practice and must be learned from the training patterns. Sometimes the form of the class-conditional densities is known (e.g. Gaussian), but some important parameters of the distribution such as mean vectors or covariance matrices are unknown. In this case, the unknown parameters in the distribution are replaced with the estimated values. If the form of the class-densities is not known, we must either estimate the density function or directly construct the decision boundary. Statistical pattern recognition requires learning in the training stage of classification. Each training pattern is explicitly labeled with the category to which the pattern belongs. The performance of a classifier greatly depends on the number of available training samples and specific values at those sample points.

3 Orthoimagery

3.1 Normalizing and classifying orthoimages

Orthoimages are produced from conventional aerial photographs and are subject to all of the vagrancies of the photographic process. Though care is taken to use images taken when the sun angle is high, shadows still occur. This is particularly so in images of alpine terrain due to the steep slopes that are often present. To determine surface type at each location in an orthoimage, it is desirable to first reduce those brightness effects in the image that are due to shading rather than surface reflectivity. In order to render a view with a simulated sun angle different from the actual sun angle when the image was required, this same shading normalization must be accompanied by a process that removes the existing shadows.

3.2 Removing shading effects

To use aerial imagery in the rendering of terrain as it would appear at different times of day, we need to minimize the luminance variability in the source imagery that is due to illumination effects at the time the imagery was acquired. If we had a way to recover surface albedo from luminance, this would also aid in determining what sort of surface cover was present at a given location in the image. Given Lambertian surfaces, a known distribution of illumination, and surface orientation at every point, it is straightforward to determine surface reflectances. In practice, we know none of these properties. Surface reflectance is far from Lambertian, illumination depends not only on sun angle but also on complex, weather dependent variations in sky brightness, and DEMs provide only low resolution information about surface orientation.

Nevertheless, shading effects can be reduced by applying a normalization based on the cosine of the angle between the approximate surface orientation, as specified in a DEM, and an estimate of the direct illumination direction. Sun angle is often provided with satellite data. For USGS orthoimages, it must be estimated from the imagery. Computer vision shape-from-shading methods can be used to solve this problem [2]. If shadows are present and one or more matches can be established between points on shadow generating contours and the corresponding point on the trailing edge of the shadow, then the direction of direct illumination can be inferred from the DEM-specified elevations of the corresponding points.

3.3 Classifying orthoimages

Figure 2 shows a 480m by 340m section of an orthoimage of an area of the Rocky Mountains. Included within the image are regions of pine trees, brush, talus, rock cliffs, and snow (Figure 1). Though not visible at the resolution with which the image is printed here, the pine trees are surrounded by an understory consisting of dirt, grass, and shrub. Portions of talus, cliff, and snow are in shadow. Each of these classes of surface cover has a distinct coloration. Given the panchromatic brightness at each pixel and the corresponding surface type, it is straightforward to produce a relatively accurate color version of the image.

Image brightness can yield a rough categorization of these regions: pine is dark, talus a mid-gray, and snow is bright. A quantitative examination of image values, however, quickly demonstrates that thresholding cannot adequately separate the classes of interest, no matter how carefully the thresholds are chosen. Computer vision techniques based on 2-D shape analysis are not likely to succeed either, given the complexity of the images. Instead, we have successfully used a pattern classification approach similar to that used to classify multi-spectral data.

For each pixel in the deshadowed orthoimage, we computed eight features:

1. pixel brightness
2. average neighborhood brightness
3. minimum neighborhood brightness
4. maximum neighborhood brightness
5. elevation
6. slope
7. aspect
8. angle to southern occluder

Features 2–4 allow consideration of brightness within a local context. Features 5 and 6 are computed by interpolating 30m DEM values. Feature 7 measures the direction a given point on a slope is facing, and important determinant of vegetation cover. Feature 8 measure the angle from a given point to the southern skyline. Larger values increase the likelihood that the point will be in shadow when the image was acquired.

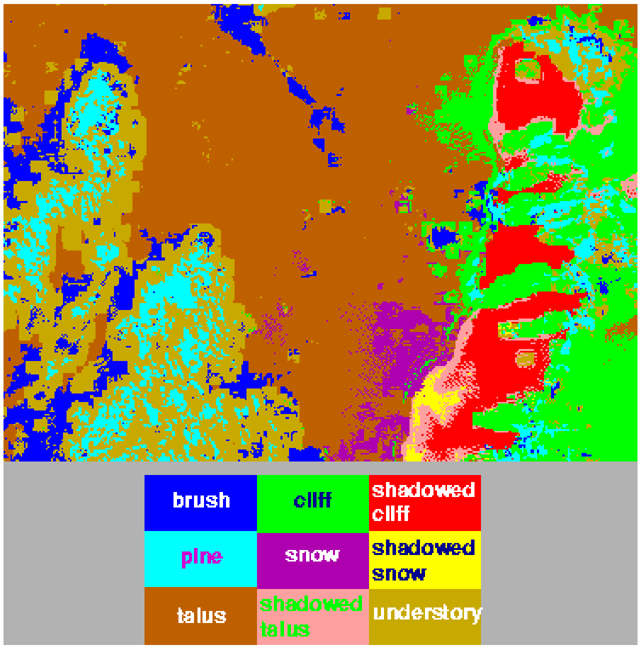

Figure 1: *Classification classes*

A simple normal distribution, maximum likelihood Bayes classifier [3] proved sufficient, avoiding the need for complex training procedures, hand tuning of parameters, or other manual adjustments. This classifier assumes that the values of each feature for each class are generated by a normally distributed random process, possibly involving correlation between different features. Population statistics, $P(\mathbf{x}|C_k)$, are computed for feature values, represented as a vector $mathbfx$, given that the features came from a particular class, C_k. Since the feature values arising from a given class are assumed to be normally distributed, their statistics are completely characterized by a mean vector, μ_k, and a covariance matrix, Φ_k. It is easily shown that given the *a priori* likelihood of each class, $P(C_k)$, the minimum error classifier is achieved by assigning the class C_k such that:

$$P(C_k|\mathbf{x}) \geq P(C_j|\mathbf{x}) \forall j \neq k. \tag{4}$$

Bayes law is used to convert this to a discriminant function formulation, in which C_k is chosen maximizing:

$$\begin{aligned} g_k \quad = \quad & -\frac{1}{2}\mathbf{x}^t[\Phi_k]^{-1}\mathbf{x} + \mathbf{x}^t[\Phi_k]^{-1}\mu_k \\ & -\frac{1}{2}\mu_k^t[\Phi_k]^{-1}\mu_k + \ln P(C_k) - \frac{1}{2}\ln|\Phi_k|. \end{aligned} \tag{5}$$

For each class listed in Figure 1, several hundred image locations were selected manually to form a training set, in a process requiring only a few minutes of time. Statistics on the distributions of feature values for the training set were determined and used to form the discriminant functions for a maximum likelihood Bayes classifier [8]. This classifier was then used to categorize each pixel location in the full orthoimage.

Classification results are shown in Figure 3. While ground-truth validation has not been done, spot checking of the results corresponds closely with what would be expected from a careful examination of the orthoimage. It is important to note that the classification was accomplished with no hand tuning of parameters or other manual adjustments, other than the selection of training samples.

Figure 2: *480m by 340m section of an orthoimage of the Wastach Mountains*

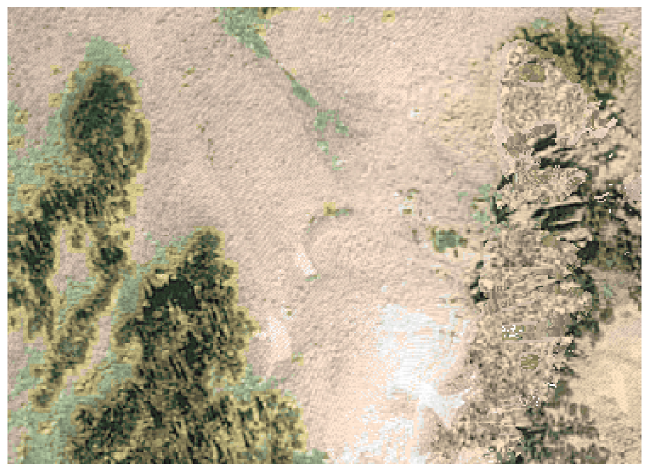

Figure 3: *Classification results with coloring from Figure 1.*

3.4 Removing shadows

While the aerial imagery used to generate orthoimages is chosen to minimize shadows, shadowing is still present. As would be expected, the severity of this problem increases with the ruggedness of the terrain. These shadows need to be removed and replaced by simulated shadows resulting from a different illumination direction if we want to use the imagery to texture a terrain scene for a date/time different from when the image was taken. Given accurate information about the direction of incident direct illumination and high resolution elevation data, expected shadow locations could be easily computed. Given that we have neither, another approach is needed.

The maximum likelihood classifier does a good job of identifying shadow areas and can even categorize different surface covers within the shadowed regions. This can be used to remove the photometric effects of shadowing, even when the direction of illumination is not known. For purposes of visual rendering, it is enough to renormalize shadowed portions of the orthoimage to have a brightness distribution statistically similar to unshadowed regions of the same surface type. In practice, it appears to be enough to standardize the mean and standard deviations of the shadowed regions. For some orthoimages, dynamic range compression in darker shadow regions results in quantization noise that is exacerbated by mean-

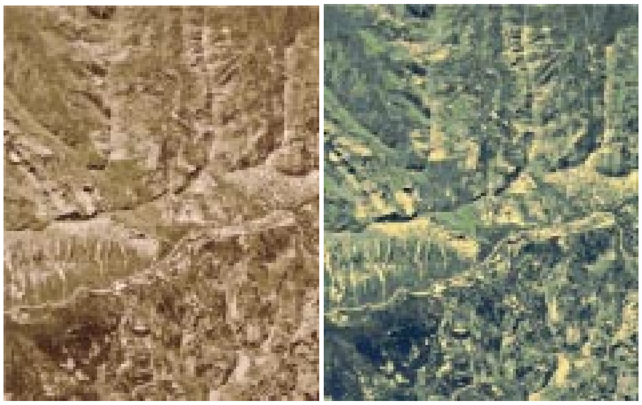

Figure 4: *Color Transfer to Grayscale Orthoimage*. Fully automatic statistical method is used to transfer color from a sample source image to a grayscale target image by making statistical color distribution in the target image equal to the distribution in the source image. Different source images were applied to the same grayscale target image.

variance normalization. This problem can be largely eliminated by adding a spatially correlated dither.

Often, shadow boundaries in orthoimages exhibit a penumbra-like effect, though at a scale much larger than the shadow penumbra that would be generated by a light source the angular extent of the sun. The causes of this phenomenon are not clear, but are likely due to a combination of interreflection, variations in sky brightness, and photographic dodging done as a pre-processing step in orthophoto preparation. Whatever the cause, these shadow fringes are visually distracting and can generate ground type misclassifications. Fortunately, it is an easy matter to replace dark pixels near classified shadow regions with lighter pixels slightly farther away, largely eliminating the problem.

4 Color Transfer to Grayscale Images

Recently, Ruderman *et al.* developed a color space, called $l\alpha\beta$, which minimizes correlation between channels for many "natural" scenes [6]. Reinhard *et al.* [5] used this color space to transfer color from one color image to another. The basic idea of the method is to transform the three-dimensional color distribution between the target and source images such that the two distributions match.

A relatively simple extension to this algorithm can be employed to transfer color to grayscale images. A grayscale image only has intensity (or luminance) distribution, therefore only luminance distributions can be matched between two images. A single luminance value is ambiguous because it could represent different parts of an image. Therefore the statistics within some neighborhood is considered to match luminance distributions. When a pixel is matched the color is transferred to a grayscale image, but the luminance of the pixel is retained. The final colors are assigned to each pixel in the original grayscale image by matching each pixel in the source image with the pixel in a target image using some distance metric (L_2 distance). Each pixel is therefore determined only by matching it to other pixels in the same image. More details about this algorithm can be found in Reinhard *et al.* [5] and Welsh *et al.* [10]. Some preliminary results of transferring color from color images to panchromatic images are shown in figure 4.

5 Conclusion

Automatic color transfer to grayscale images is an underconstrained problem that does not have optimal solution. We described an approach for semi-automatic coloring of panchromatic aerial orthoimagery based on pattern recognition and classification techniques. The method is able to remove shading and shadowing effects in the original image so that shading and shadowing appropriate to variable times of day and year can be added. Automatic methods of color transfer are possible based on statistical analysis of color distributions, however they often do not provide the desired results due to the underconstrained problem. It would be beneficial to try to combine semi-automatic pattern classification method with fully automatic statistical color transfer method using Bayesian statistics and learning algorithms.

Acknowledgement We would like to thank Tomihisa Welsh for his help with image colorization.

References

[1] Aerial-Images, Inc. http://www.aerial-images.com. 615 Hillsborough Street, Raleigh, NC 27603, USA.

[2] M. J. Brooks and B. K. P. Horn. Shape and source from shading. In B. K. P. Horn and M. J. Brooks, editors, *Shape from Shading*. MIT Press, Cambridge, MA, 1989.

[3] L. Devroye. Automatic pattern recognition: A study of the probability of error. *IEEE Transactions on Pattern Analysis and MAchine Intelligence*, 10(4):530–543, 1988.

[4] Anil K. Jain, Robert P. W. Duin, and Jianchang Mao. "statistical pattern recognition: A review". *IEEE Transactions on Pattern Analysis and MAchine Intelligence*, 22(1):4–37, January 2000.

[5] Erik Reinhard, Michael Ashikhmin, Bruce Gooch, and Peter Shirley. Color transfer between images. *IEEE Computer Graphics and Applications*, pages 34–40, "September/October" 2001.

[6] D L Ruderman, T W Cronin, and C-C Chiao. Statistics of cone responses to natural images: Implications for visual coding. *J. Opt. Soc. Am. A*, 15(8):2036–2045, 1998.

[7] SPOT Image Corporation. http://www.spot.com. 1897 Preston White Drive, Reston, VA 20191, USA.

[8] J. T. Tou and R. C. Gonzalez. *Pattern Recognition Principles*. Addison-Wesley, Reading, MA, 1974.

[9] U.S. Geological Survey. http://www.usgs.gov. 807 National Center, Reston, VA 20192, USA.

[10] Tomihisa Welsh, Michael Ashikhmin, and Klaus Mueller. Transferring color to grayscale images. In *Proceedings of SIGGRAPH 2002*, 2002.