Image Sorting Using Centroidal Voronoi Tessellation

Charles Wesley Cowan (cwcowan), Alonzo Javier Benavides (abenavid)

9/4/2007

Abstract

In this study, the research team will be developing a computational framework which will autonomously sort any set of images based on their content. Most systems rely on some degree of human-computer interaction, typically based on user labeling, tagging. The researchers will attempt to eliminate any form of user classification; instead, the proposed system will rely solely on the computational models of the images themselves. This is important because there is a multitude of images on the internet that are not labeled or still waiting to be labeled. By creating a standardized framework to interpret and process images, we will be able to eliminate any ambiguity introduced in user tagging as well as the time constraints imposed by human-interaction.

The central tool in this approach will be Centroidal Voronoi Tessellation (CVT). Normal Voronoi Tessellation is a partitioning of a given space based on some distance or energy function. Every point in the space is partitioned amongst a set of 'generators,' points representative of every point in a given partition. In Centroidal Voronoi Tessellation, the generator of each partition is chosen specifically to minimize the energy function between it and every point in that partition. If the energy of the entire system is defined as the sum of the energies between generators and their associated points for every point in the space, the CVT absolutely minimizes this total energy function for the whole system. The goal of this project is to represent images in the form of CVTs in a comparable form by using an appropriate energy function.

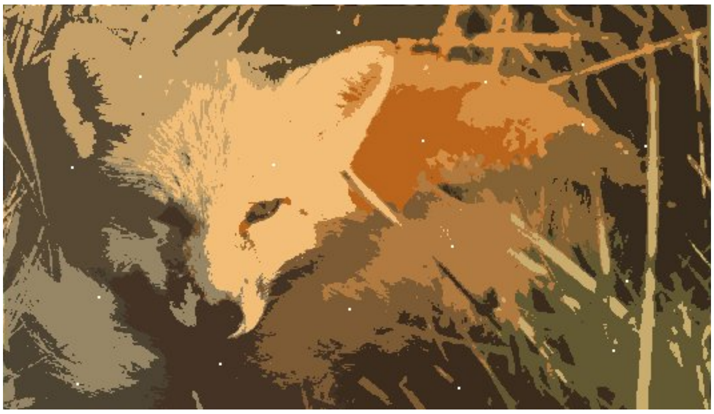
Currently, user labeling and the use of tabboo words represents the majority of image classification techniques. With the proposed system, it will no longer be necessary to worry about the inconsistency of human judgement and the time-constraint of user labeling. The use of a computational framework will greatly reduce the number of variables involved in image searching, sorting, and labeling. A standardized procedure of quantifying the content of images will lead to the optimization of resources for image-based projects.

Data Analysis

Consider the image below. As an example of the CVT breakdown of an image, this image will be partitioned among a set of generators to demonstrate the represenative properties of the CVT.

Selecting 16 generators spaced evenly across the image, the pixels of the image are grouped to the generator they are nearest to in terms of colour and distance. The effect is the image below, where each pixel is rendered in the average colour of the partition to which it belongs. The positions of the generators of each pixel are rendered as white dots.

However, that is simply a voronoi tessellation. Using Lloyd's method, we can itterately determine the centroids of each partition, reposition the generators, and reparttion the image. The ultimate effect is that the generators become located at the centroids of each partition, and the color/position distance between any pixel and its associated generator is minimize. Notice the way the partitions correlate to related shapes and colours, caputuring the significant features of the picture.



Of course, the degree to which the CVT captures the picture depends to a large extent on how many generators there are. The above images, using 16 generators, are a fair approximation. Below, using 400 generators, the accuracy of the representation is dramatically increased.

Once all pictures are broken down in this way, the information encapsulated in the partitioning can be used to compare pictures based on their content.

Background Literature

The method for processing each individual image is will explained in Centroidal Voronoi Tessellation Algorithms for Image Processing, by Qiang Du. His research paper can be found at the following address:

www.math.psu.edu/qdu/Res/Pre/dgjwimage.pdf

The goal of this project is to extend these kind of results to large collections of images.

Materials and Methods

For this research project, there are several problems that need to be addressed. The main material for this project will be a sizable database of images on which to experiment. Secondly, there are a number of issues involved with the CVT process itself, the most important of which is the weight function used in the partitioning. The optimal definition of this weight function for the stated goals of the project must be determined. Once images are broken down into their CVT partitions, the population of each partition must be analyzed for relevant information to the image. The sorting process is then based on performing a CVT partitioning of the image CVTs themselves. For this to be effective, the optimal information from the image CVT to use in the second level partitioning must be determined as well, as well as the optimal weight function to use on that information.

The basic logic of the proposed method is as follows: for each image in a given database, the pixels of the image are partitioned into a finite number of groups based on a predetermined energy function. Once a given image is partitioned in this way, each partition can be surveyed for relevant information to the image as a whole. Such information might include average color, position, and overlap between partitions. This information, in total, becomes representative of the image as a whole, as though in a compressed form. This reduced representation can then be compared to similarly structured representations of other pictures. To preform this comparison, a second partitioning, also based on CVTs, is performed on the image representations themselves, using an appropriately defined weight function. In doing this, the representations will be grouped among themselves in terms of similarity of structure and content. In the end, the original database of images be separated to form a set of smaller groups of images that are most like the other images in its group. By using an appropriate weight function, any given image will resemble every image in its group better than all the other images that were in the original database.

Conclusion

The success of this project will lead to automating the way images are processed and sorted. By completing the aforementioned computational framework, there will be no need to individually label all the images in a collection. In finding and labeling the generators of a database, all the images in their final partition can be classified under the same category as the name given to the generator. The simplicity and purely computational approach of this method will help advance the way the computer science community handles large quantities of images.