Available online at: https://ijact.in

Date of Submission	19/03/2019
Date of Acceptance	29/03/2019
Date of Publication	13/04/2019
Page numbers	3092-3096 (5 Pages)

ISSN:2320-0790

An International Journal of Advanced Computer Technology

IMAGE GENERATION BY USING K-MEANS CLUSTERING TECHNIQUE

Mariam A. Ali[1]

College of Administration and Economics, Department of Financial and Banking Sciences
University of Baghdad

Abstract: This paper presents a new approach for image generation by applying k-means algorithm. The K-means clustering algorithm is one of the most widely used algorithm in the literature, and many authors successfully compare their new proposal with the results achieved by the k-Means. Our research proposes a color-based image generation method that uses K-means clustering technique which is an iterative technique used to partition an image into k clusters. At first the pixels of source image are clustered into k partitions based on their color, where the clustering process is accomplished. Then the clustered colors are merged to generate the target image. After applying K-means algorithm for different values of k (no. of colors), Mean Square Error (MSE) and processing time are evaluated for different types of image. This approach thus provides a feasible new solution for image generation which may be helpful in image compression.

Keywords: K-means Algorithm, Clustering, colorization.

I. Introduction

Most popular reasons to generate an image from another image : choose the colors and intensity which increase the visual appeal of an image such as old photo; they help to make a scientific illustration more attractive.[1]Additionally, selection of the suitable feature(s) is the key starting point to building image processing and machine vision systems. In many ways, the success or failure of an algorithm depends greatly on an appropriately designed feature selection and representation. In the computer vision community, we can generally classify features as low level, intermediate level and high level. Low level deals with pixel level features such as neighborhood shape and size and pixel neighborhood statistics (e.g., mean, standard deviation, and Euclidian distance). High level deals with abstract concepts (such as types and semantic meaning of image's objects), and intermediate level deals with something in between. In pixel level, RGB channels of the image are used, traditionally, as the feature vector. Instead, it requires an appropriate definition of "similarity" to measure the relationship between source of property image and target image. In this case, RGB channel alone may contain data of color to match between the images as one may wish to process a target image giving a source image with completely different colors.

Although various image processing concepts are in common use in the computer graphics[2], an application of image generation using K-means clustering algorithm is relatively a recent field of research.

II. RELATED WORK

Several techniques and methods are published for image generation. Readers please refer to recent papers. Some image generation works are object transfer of Yanqiu et al [3], image histogram and HSV of Shamik S. et al. But the classical full search of Welsh et al, and the Antypole strategy of Di Blasi and Reforgiato [4] are color transfer. The main concept of these image generation techniques is to exploits textural information. For example, the work of Welsh et al, which is inspired by the color transfer [2] examines the luminance values in the neighborhood of each pixel in the target image and add to its luminance the chromatic information of a pixel from a source image with best neighborhoods matching. This technique works well on images were differently colored regions give rise to distinct textures. Di Blasi and Reforgiato [4] propose a new strategy to improve Welsh et al work, where Antipole clustering strategy is adopted as an efficient data structure for fast color retrieving. Their approach provides a way to speed up the searching process but at the expense of increasing implementation complexity. Yanqiu et al proposed to generate an embedded image automatically by inserting an interior object with partial deformation into an outer object image where the color of interior object is opposite to the outer image. Bezier Curves had been used to generate a 2D virtual secure image from Arabic text in any document and evaluated from read text in the document[5].Ammar A. et al [2]proposed new method for image generating rely on luminance matching techniques between the source and target pixel neighborhood, and transfer the best pixels of neighborhood matching, i.e., luminance and chromatic information from source to target image. Additionally, B.A. Attea et al presented a genetic algorithm for generating a painting that is close to a given input image [6]. Genetic algorithms have been developed that modifies the traditional uniform crossover to spread out vital genes at the expense of lethal genes rather than exchanging them between matting parents.

Many K-means clustering techniques in image generation and quantization have been proposed in the literature. Some of them are discussed below.

The developed properties of HSV color space and k-means clustering algorithm have been used for image segmentation of the image into objects and color histogram generation by Shamik S. et al [7].Omran et al[8] used Particle Swarm Optimization(PSO) to quantize color images and performed clustering of the color space . The K-means clustering algorithm is then applied and each pixel is assigned to the cluster with the closest centroids which refined then by applying PSO.

III. THE PROPOSED IMAGES GENERATION TECHNIQUE

In this research, K-means clustering technique has been used to generate target image from source one by using k colors from the source. K-means clusters the pixels in image into specified number of clusters (colors) are unsupervised learning mechanism such that the distance between clusters is minimized. Result of this step is that we get clusters with their centers, pixels contained in them and distance of each pixel from the center of the contained cluster.

A. RGB Color Space

The name of the model comes from the initials of the three additive primary colors, red, green, and blue. In RGB color model each color appear as the combination of three primary colors: red , green and blue. To generate or modify any color within the RGB color cube requires modification of all three components (red, green and blue colors) to be of equal pixel depth and display resolution. White color is full intensity for each component and the resultants of zero intensity for each component gives black. When the intensities of red, green and blue are the same, the result is a shade of gray; lighter and darker depending on the intensity.[9]

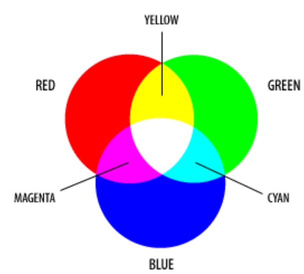


Fig (1) RGB Color Model

B. K-Means Clustering Algorithm

k-means clustering is a method of vector quantization and a one of the most widely used methods for data clustering [10] that solve the well known clustering problem.

Given a set of data X= (x1, x2, …, xn) ∈ RD, the aim of k-means clustering is to partition X into k (≤n) sets S = {S1, S2, …, Sk} so as to minimize the sum of distance functions(in this case a squared error function)of each point in the cluster to the K center. In other words, its objective is to find:

$$\arg\min_{\mathbf{S}} \sum_{i=1}^{k} \sum_{\mathbf{x} \in S_i} \|\mathbf{x} - \boldsymbol{\mu}_i\|^2$$

where μi is the mean of points in Si.

Algorithm: K-means algorithm for generating target image by using k-colors from source image.

Input: 1- source image

2- Target image

3- k : number of colors

Output: colored target image.

Procedure:

1. Load source image and resize it.

2. Initialize k-means by sample k colors randomly from the resizing source image.

k1 , k2 , k3 , ….. kn.

Repeat

For each pixel in the source image calculate its nearest mean

M(i) =Min(j) │P(i) – k(j) │2

Update the values of means only if there are pixels assigned to them

$$x = \frac{\sum \text{in } 1(\text{M(i)} = \text{ j}) \text{ P(i)}}{\sum \text{in } 1(\text{M(i)} = \text{ j})}$$

Load target image and search the best pixels and neighborhood pixels.

Transfer pixels(colors) from source image to the target image by replacing each of target's pixels with the nearest of the centroid colors which are found from the source image. Then save and display new image.

IV. THE EXPERIMENTAL RESULTS

In general , there is no good objective criterion available for measuring the perceived image similarity. However, there are a number of common error measurements [7]. One of the error metrics used to compare the various image techniques is the Mean Square Error (MSE). The MSE is the cumulative squared error between the output image and the original image and it measures the average amount of difference between pixels of output image and original target image. If the MSE is small, the output image closely resembled the original.

The mathematical formula for MSE is defined as:

$$\text{MSE} = \frac{1}{MN} \sum_{y=1}^{M} \sum_{x=1}^{N} \left[I(x,y) - I'(x,y) \right]^2$$

where I(x,y) is the original image, I'(x,y) is the generated image.

All experimental results obtained with the K-means algorithm are reported in this section.

Figure(1) shows some examples of generation images which are obtained with the K-means method.

V. RESULTS AND DISCUSSIONS

In order to generate colored target image from anther colored image, the k-means method had been implemented and MSE and PSNR areapplied on the result images.

This proposed method was implemented using MATLAB version R2009a on Lenovo Intel(R) core ™ i5 , 2.5 GHZ, 4GB RAM.

Clusters of target image is generated in previous step contain information about cluster center. Cluster from source image is match with cluster from target image on base of minimum distance and the color features of source image which is calculated between centers of the clusters in RGB color space. In this case we used k-colors from source image to generate the target image.

This generation is done by matching the pixel (3 distances) from center of cluster. The best matched color is transferred to target image pixel based on closest distance of corresponding clusters. So, RGB values are copied from pixels of source image cluster to pixels of target image clusters.

VI. CONCLUSION

In this paper, we have presented a new algorithm for image generation based K-means clustering algorithm. For larger values of k, the algorithm gives good results and takes long time , whereas, for smaller values of k, the inverse is right. So , the advantage of K-means algorithm is simple and quite efficient and gives good compressed image quality . On the hand, this algorithm is slow, especially when the number of colors in the source image is large.

In general, K-means algorithm achieves high coding performance comparing with other image generation algorithms. In order to achieve low computing cost and high image coding performance, multiple reduced images are generated.

However, K-means has not received much respect in the color quantization, colorization and generation literature because of its high computational requirements and sensitivity to initialization in spite of its popular approach in general purpose clustering algorithm.

VII. References

[1] Imran, M.(2011). Coloring grayscale images, Master Thesis Computer Engineering, Nr:E4104.

[2] Hassan, A. A. and Aboud, A. D.(2008)" Automated approach for color image generation, Journal of Engineering, Volume 14 Issue: 1, pp2164-2172.

[3] Rrinhard E, ,Ashikhmin M., Gooch B and Shirley P. (2001), Color transfer between images, IEEE Computer Graphics and Applications, 21(5), 34-41.

[4] Blasi G. Di and ReforgiatoR. D.(2003). Fast colorization of gray images, In proceedings of Eurographics Italian Chapter.

[5] Anmar A. M.,(2014), "2D Virtual Image Generation for Document Security Using Bezier Cubic Splines", Eng. & Tech. Journal, Vol. 32, , No.6 Part (B), **ISSN**: 16816900 24120758 pp1074-1083.

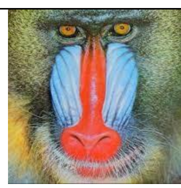
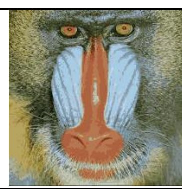
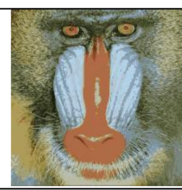
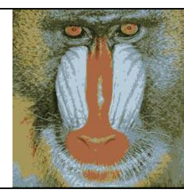
[6] Attea, B.A. &Aboud, A.D. (2005) A genetic approach for automated image generation: greyscale image generation, Journal of Engineering, 2(11), 393-403

[7] Sural,S., Qian,G. and Pramanik,S. (2002) Segmentation and histogram generation using the HSV color space image retrieval, IEEE ICIP, DOI:10.1109/ICIP.2002.1040019, 589-592.

[8] Mohamed G. Omran , Andries P. Engelbrecht and Ayed Salman (2005),"A Color Image Quantization Algorithm Based on Particle Swarm Optimization, Informatica 29, 261-269.

[9] https://en.wikipedia.org/wiki/RGB_color_model

[10]Guojun Gan, Chaoqun Ma, Jianhong Wu, (2007), Data Clustering: Theory, Algorithms, and Applications, SIAM. Society for Industrial and Applied Mathematics Philadelphia, Pennsylvania, American Statistical Association Alexandria, Virginia.

Source Image	**Target Image**	**Result Images**			
		K=64	**K=32**	**K=16**	**K=8**
					
204x204	204x204	MSE=45.48 Time =170.10 sec.	MSE=55.19 Time =170.09 sec.	MSE=73.46 Time =159.59 sec.	MSE=84.28 Time =159.56 sec.

Source Image	**Target Image**	**Result Images**			
		K=64	**K=32**	**K=16**	**K=8**
					
259x194	259x194	MSE=27.44 Time =181.64 sec.	MSE=37.52 Time =156.28 sec.	MSE=48.70 Time =146.91 sec.	MSE=58.52 Time =113.08 sec.

Source Image	**Target Image**	**Result Images**			
		K=64	**K=32**	**K=16**	**K=8**

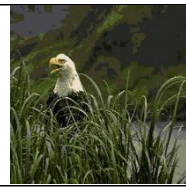
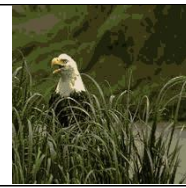
					
204x204	204x204	MSE=57.39 Time =175.30 sec.	MSE=77.92 Time =173.43 sec.	MSE=97.56 Time =167.22 sec.	MSE=106.21 Time =94.75 sec.

Figure: 1