ISSN: 0976-9102 (ONLINE)
DOI: 10.21917/ijivp.2021.0340

SHADOW DETECTION FROM AERIAL IMAGERY WITH MORPHOLOGICAL PREPROCESSING AND PIXEL CLUSTERING METHODS

Ashwini M. Deshpande[1], Manali Gaikwad[2], Saudamini Patki[3], Aditee Rathi[4] and Sampa Roy[5]

[1,2,3,4]Department of Electronics and Telecommunication Engineering, Cummins College of Engineering for Women, India
[5]Signal and Image Processing Group, Space Applications Centre, India

Abstract

Building extraction from aerial imagery facilitates many geo-specialized tasks like Urban Planning, Map Generation and Disaster Management. Well planned cities ensure good sanitation, lesser pollution and hence, a better standard of living for its citizens. This is essential for developing countries which face a major crisis of urban migration and space crunch, and where planned cities would be a move towards smart living. The objective of this work is to segment building footprints from aerial images. Traditional pixel clustering algorithms like K-means, Color Quantization (CQ) and Gaussian Mixture Model (GMM) are implemented with inclusion of preprocessing steps for improved performance. These techniques are compared based on performance and time taken. The number of clusters/components are selected on the basis of Silhouette Score and Akaike Information Criterion/ Bayesian Information Criterion (AIC/BIC). A commonly encountered problem in building segmentation is misclassification of pixels due to shadows. This challenge is dealt by masking shadows using morphological operations as a part of preprocessing.

Keywords:

Shadow Detection, K-Means, Colour Quantization, Gaussian Mixture Model, AIC/BIC

1. INTRODUCTION

Semantic segmentation is the process of assigning a class label to each pixel within an image. Segmentation of buildings is a fundamental step in providing information to computers related to extent of development and construction in an area, understanding infrastructure requirements and identifying potential land use. In this way it helps in training them for providing better assistance in map making and urban planning. In a country like India where there is a major crisis of urban migration and space crunch, planned cities are a move towards smart living.

The objective of this work is to segment buildings from complex scenes which comprise man-made and natural objects, such as roads, vehicles, waterbodies, and vegetation. Here, the scope pertains to segmenting aerial images of cityscapes. Aerial images are captured from a bird's eye view perspective by a suspended camera like a drone, aircraft, or helicopter. The dataset chosen to work on is SpaceNet buildings (Vegas) made available by AWS [1]. This dataset is chosen for work as it shows prominent building and tree shadows.

State-of-the art methods such as pixel clustering methods works on the group of pixels surrounding the pixel at the center considering a smaller region. The neighboring pixels of the pixel in center facilitates to find similarity or change in brightness using cues such as edges and/or lines.

Pixel clustering methods are unsupervised algorithms which help fragment the image into segments constituting similar pixel properties like pixel intensity, color, texture, and pixel proximity. Traditional pixel clustering methods such as K-means, Color Quantization and Gaussian Mixture Model are used for building segmentation in this work. The performance of these traditional algorithms is compared on the aerial images with an objective of segmenting building footprints. Parameters like silhouette score and Akaike Information Criterion/ Bayesian Information Criterion (AIC/BIC) were used to optimize the model.

Precise extraction of buildings is limited sometimes by shadow artifacts. To increase the accuracy of segmentation, the shadow region is masked by morphological operations before applying the pixel clustering methods. Morphological operations proposed enable to mask the shadow regions detected and help to improve the accuracy of classification between the urban features like roads and buildings.

2. LITERATURE SURVEY

Automatic detection and classification of urban features like roads and buildings from high resolution satellite/aerial imagery using computer vision approaches is one of the challenging tasks. Techniques proposed by many researchers in literature are either based on image processing techniques such as edge-based, contour-based detection or have used supervised learning techniques.

Farnoosh et al. [2] have proposed using GMM with EM which uses Bayesian MAP estimation. A sequence of prior and posterior probabilities is made for maximum posterior estimation to construct a labelled image. This approach works well for relatively simple images, but in case of complex images parameter initialization plays a crucial role.

Valenzuela et al. [3] have suggested a method utilizing adaptive initialization, deterministic subsampling and efficient core-set construction to attain high speed and quality quantization.

Sattarov et al. [4] explored different cases to find out impact of initialization and preprocessing on the clustering process.

Celebi [5] worked on developing a variant of k-means that utilizes data reduction, sample weighing, and accelerates nearest neighbour search.

Sudhir Singh et al. [6] have studied the time complexity of k-means clustering and nearest neighbour clustering algorithm to improve the compactness of clusters.

Chinki et al. [7] worked on colour based image segmentation by applying *K*-means. This technique partitions the image into *k* clusters. It produces accurate segmentation when applied to images defined by homogeneous regions with respect to texture and colour with no local constraints applied to impose spatial continuity.

Building extraction using principal contours of buildings and radiometric behavior of buildings in dense urban areas was proposed by Peng et al. [8].

An active contour-based model was developed by Ahmadi et al. [9] for building boundary extraction from high resolution aerial images. Their method deals with complex images, however radiometric signature similarity between building roofs and background is challenging for this method.

Kaiser et al. [10] worked on making a more efficient segmentation model by training on crowdsourced noisy data. It was noted that this approach helped predict labels for unseen cities accurately and consumed 30% less time. However, it was observed that the results showed a tendency to miss parts of road, had blurry building outlines, shadow and orthorectification artifacts.

3. METHODOLOGY

A commonly encountered problem in segmentation of aerial imagery is misclassification of pixels due to shadow regions. To overcome this challenge, the shadow regions are identified, and masked using morphological operations. These operations aim to remove any kind of imperfections by accounting for the structure and form of that image. As stated by Statella et al. [11] shadow regions are characterized as low intensity, low contrast regions which tend to behave as minimum in digital images.

The Fig.1 shows the proposed preprocessing steps employed on input images before applying pixel clustering methods. First, opening operation is carried out on the original image to smoothen the patches of interest. Opening operation basically tends to eliminate bright pixels from the edges of foreground. The size and shape of kernel used for morphological operation depends upon the object of interest. Upon experimentation, it was found that square kernels of size (3,3) work best for these images.

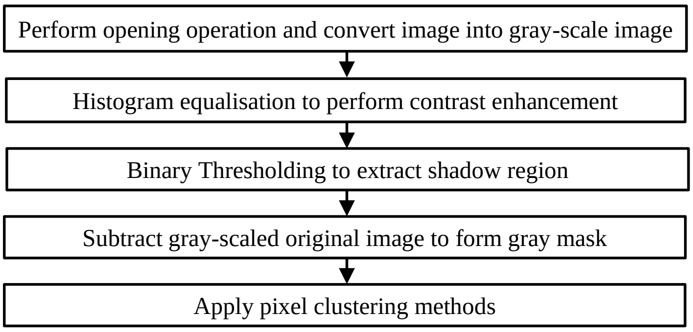


Fig.1. Preprocessing steps

Next, the resultant image after opening operation is converted into grayscale and to improve contrast of the image, histogram equalization is carried out. Histogram equalization is done by spreading out the most frequent intensity values. This enables low local contrast regions to appear as high contrast. This is followed by inverse binary thresholding which gives highlighted shadow region. In inverse binary thresholding a gray scale image is converted into black and white image by setting pixels with intensity greater than threshold to black and pixels with intensity less than threshold to white.

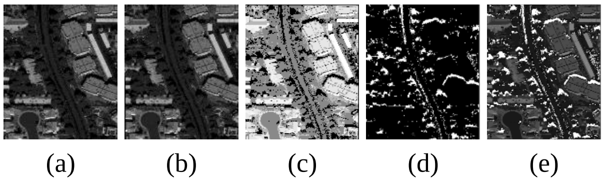


Fig.2. Steps for preprocessing (a) Input image (marked with building shadow) (b) Image after opening operation performed (c) Image after applying histogram equalization (d) Binary thresholded image (e) Result after being subtracted to the original image (marked with masked shadow)

The original image is subtracted by weights with the resulting image to give an image with shadow region covered by gray mask. This mask clearly segregates the gray shadow region from the darker building pixels and prevents interference of shadow pixels in classification of building pixels. The Fig.2 shows the stepwise results obtained during preprocessing of input image. (After preprocessing) The pixels of the preprocessed image are then clustered using K-means, color quantization and Gaussian Mixture Model.

The K-means clustering method works by agglomerating pixels in such a way that the pixels inside each cluster are very identical to each other, yet very different from the other clusters [12]. If the pixels have very less difference between them, they belong to the same cluster. The distance of point from the cluster is found using Eq.(1).

$$J = \sum_{i=1}^{m}\sum_{k=1}^{K} w_{ik} \left\| x^i - \mu_k \right\|^2 \tag{1}$$

Here, μ_k refers to centroid of x^i cluster. Weight w_{ik} =1 if the data point x^i belongs to cluster k, in any other case it is equal to 0. Steps involved in this clustering method are shown in Fig.3.

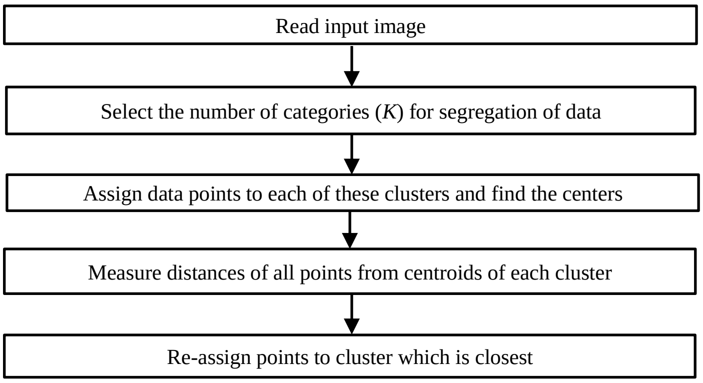


Fig.3. Steps involved in K-means clustering algorithm

Color Quantization (CQ) achieves image segmentation by reducing the number of colors in the image [13]. This is carried out by mapping three-channel 8-bit colors of an image to a limited color palette. Thus, it prevents loss of information and preserves visual appearance of the image.

This pixel mapping is carried out with the help of a codebook vector, which is a list of numbers that have same attributes as training data [14]. Predictions are made after searching through codebook vector for K (size of codebook) similar instances. For codebook size K and input vector dimension L, ($\log_2 K$) bits are required to specify which of the code vectors are selected.

The rate of vector quantizer with L dimensional input vector and a codebook of size K is given by Eq.(2).

$$(\log_2 K)/L \tag{2}$$

The pixels are shuffled prior to passing through the color quantization model for optimization. For determining the likelihood of input with a particular instance, Euclidean distance is calculated.

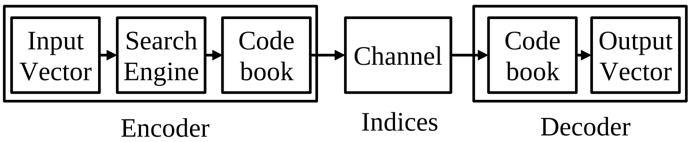


Fig.4. Colour Quantization scheme for pixel clustering

From the input image, the Gaussian Mixture Model (GMM) computes a set of Gaussian distributions, each mapped to a particular cluster or component [2]. Gaussian distribution representing a cluster/component can be defined completely by its mean and variance as shown in Eq.(3).

$$f\left(x\middle|\mu,\sigma^2\right)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{3}$$

where μ is the mean and σ^2 is the variance. The Gaussian mixture model can be represented as:

$$P(x)=\sum_{k=1}^{K}\pi_k N\left(x\middle|\mu_k,\Sigma_k\right) \tag{4}$$

where, μ is the mean, π is the combination of weights used and K is the number of clusters. The responsibility by each Gaussian for every data point is given by:

$$\gamma(z_{nk})=\frac{\pi_k N\left(x_n\middle|\mu_k,\Sigma_k\right)}{\sum_{j=1}^{K}\pi_j N\left(x_n\middle|\mu_j,\Sigma_j\right)} \tag{5}$$

This is obtained according to Bayes' rule. Every data point/pixel is then checked against the Gaussian distribution it fits and assigned the label of that component. As covariance in the GMM model is flexible, the distributions of clusters need not be spherical. This means that instead of assigning a data point to one fixed cluster, the model gives the probability of how that data point belongs to one cluster more than the other. Hence, GMM does not give hard-assignments like k-means and color quantization clustering methods. The Fig.5 shows the steps involved in GMM based clustering method.

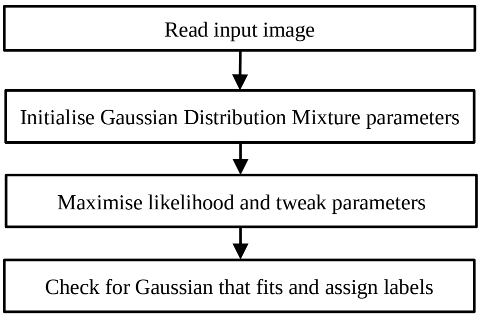


Fig.5. GMM algorithm steps for pixel clustering

As the efficiency of these unsupervised methods depends on the correct initialization of parameters (e.g. mean, covariance, and weights of distribution for GMM), we use Expectation-Maximization (EM) algorithm to set them.

The EM algorithm consists of an E-step and an M-step. In the E-step, the parameters are randomly initialized. In the M-step they are updated in such a way that the log likelihood is maximized. For optimizing the distance-based clustering models, the number of clusters was chosen by computing the silhouette score and AIC/BIC parameters.

The silhouette score gives a measure of how well defined the distribution of clusters is by taking into account the intra-cluster and mean nearest cluster (Euclidian) distances. It can be calculated using Eq.(6).

$$\text{Euclidean Distance} = (p-q)/(\max(p,q)) \tag{6}$$

where p is mean nearest cluster distance and q is intra-cluster distance. The input images to calculate the silhouette score are preprocessed to remove shadows using morphological operations. The score was evaluated for 2 to 10 clusters. It was observed that for SpaceNet [1] dataset, 3 or 4 clusters showed maximum silhouette score i.e., 0.2105 for 3 clusters and 0.2158 for 4 clusters.

AIC/BIC parameters take into consideration the likelihood function and help in model selection by penalizing features which result in overfitting. BIC penalizes complex models more heavily as compared to AIC. The elbow graph obtained by plotting AIC/BIC for the images without pre-processing is shown in Fig.6.

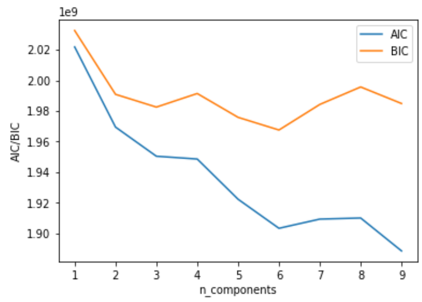


Fig.6. AIC/BIC parameters for images without pre-processing

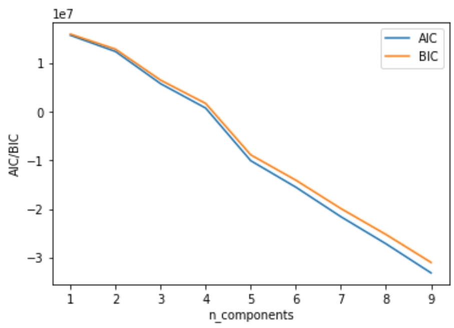


Fig.7. AIC/BIC parameters for images with pre-processing

From the Fig.7, we observe multiple elbows obtained at 2, 3, 4 and 6 components for images without pre-processing. The results obtained on selecting number of components as 2 and 4 are compared, as both the parameters i.e., silhouette score and AIC/BIC have required values. Similarly, for pre-processed images elbows are attained at 4 and 5. The results obtained on selecting number of components as 4 and 5 are compared, as both silhouette score and AIC/BIC have required values.

4. EXPERIMENTAL RESULTS

This section describes experiments on SpaceNet dataset [1], conducted to test the time consumed and extent of building extraction. The number of images selected to experiment are 10 of size: 650×650. These images are of panchromatic nature from AOI 2 Vegas.

The results without applying the morphological filters using the 3 approaches are presented in Fig.8, Fig.9 and Fig.10. The shadows in original image are masked using morphological operations as shown in Fig.2(a) and Fig.2(e). Then, the pre-processed image is given as input to the 3 approaches and the results are shown in Fig.11 and Fig.12.

The outcomes of processing the image using K-means (without pre-processing) are shown in Fig.8(a) and Fig.8(b). The results for fitting K-means model with 4 and 5 clusters (with pre-processing) show that the buildings have been extracted fairly well and the building boundaries are seen clearly as seen in Fig 11(a) and Fig.12(a). It can be observed that by applying the K-means method with 2 and 3 clusters the buildings and roads are not segmented properly.

For Colour Quantization good performance is shown for 5 clusters (with preprocessing) as seen in Fig 12(b). For clusters 2, 3 and 4 the performance is poor. Hence, Colour Quantization requires more clusters than the other methods to show good performance.

The outcomes of Gaussian Mixture Model in both cases are good. However time required is more. It can discriminate shadows even without preprocessing.

Limitation of using K-means clustering method is that it is much slower especially if the training is performed with all the pixels of the original image. We can overcome this by choosing random set of pixels in an image. Another limitation of K-means is that it learns from all samples at every iteration. Hence it is time consuming.

Colour Quantization model takes the least time to generate results. However, trade-off between accuracy and time taken is observed. From the outputs it is seen that colour quantization performs better for n=5 clusters. For n=5 clusters building footprints can be clearly distinguished, but minute details are not visible. In case of vector quantization both storage space and time needed to perform quantization grow faster than exponentially with the number of dimensions. Since there is no proper structure to codebook, algorithm must go through each and every signal vector of codebook to find the closest one it belongs to. Hence vector quantization is limited to small vectors. We can improve the performance of colour quantization using K-means as mentioned in [14].

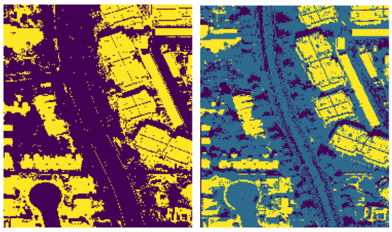
Although results obtained upon fitting K-means and GMM are similar, where GMM slightly limits noise, it consumes the most time out of the three methods. GMM has a flexible covariance. This means that instead of hard assigning data points to one cluster, the GMM output shows that it is more likely for a data point to belong to one cluster than other.

Unlike k-means, GMM does not assume clusters to be of any specific geometry and does not bias cluster sizes to circular structure. However, as the number of parameters increases, it becomes more difficult to correctly initialize them and optimize and interpret the clustering. GMM can be used with a variational model as proposed in [15] to overcome the limitation of colour inhomogeneity in image segmentation.

Hence it should be noted that where time is of essence, Colour Quantization should be the preferred method. If detail is of more important, then K-means should be preferred. Comparison of the three methods with respect to training time is given in Table.1.

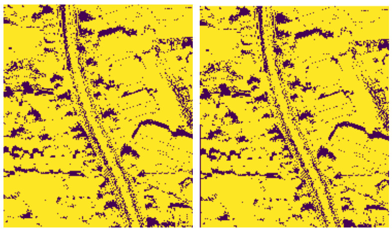
Table.1. Comparison of training times of different methods

Method used	Number of Clusters	Time taken (sec)
K-means	4	0.167
	5	0.126
Color Quantization	4	0.054
	5	0.063
GMM	4	0.277
	5	0.853



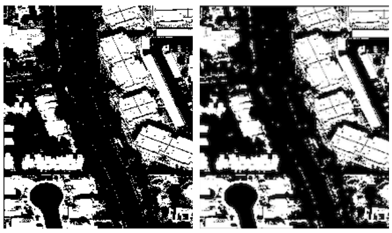
(a) (b)

Fig.8. Results without morphological operations by K-means method for (a) 2 clusters (b) 3 clusters



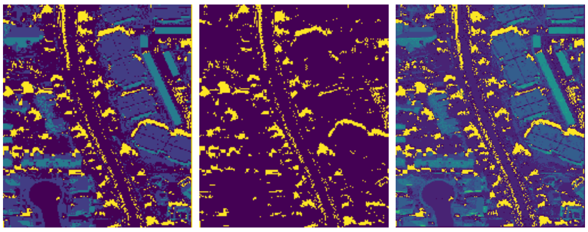
(a) (b)

Fig.9. Results without morphological operations by colour quantization method for (a) 2 clusters (b) 3 clusters



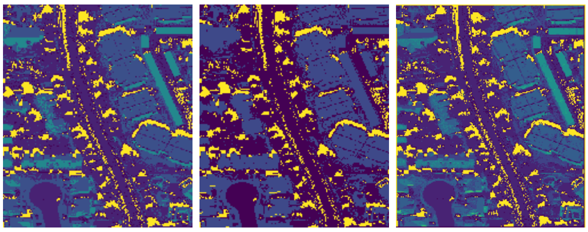
(a) (b)

Fig.10. Results without morphological operations by GMM for (a) 2 clusters (b) 3 clusters



(a) (b) (c)

Fig.11. Results with morphological operations for 4 clusters/components (a) K-means (b) Colour Quantization (c) GMM



(a) (b) (c)

Fig.12. Results with morphological operation for 5 clusters/components (a) K-means (b) Colour Quantization (c) GMM

5. CONCLUSION

Upon comparison amongst three state of art techniques (K-means, Colour Quantization, GMM) it is observed that, Colour Quantization takes the least training time while preserving less detail, whereas K-means and GMM need more training time and give a detailed result image. Hence colour Quantization is good when only spatial context of building is required, and time is a constraint. On the other hand, K-means shows slightly more noise compared to GMM but preserves all details and takes less time.

In complex scenes it is observed that similar objects may have different spectral intensities, and different objects may have same spectral intensity, and this leads to misclassification of pixels. Occlusions due to towering buildings and bridges will also prevent accurate labelling of pixels and interpreting contextual information precisely. Comprehension of instances of similar objects and object boundaries has scope for further optimization.

When working with satellite images it is important to note that as 70% of satellite images contain cloud (and cloud shadow), these regions can be masked using morphological operations, or the underlying region can be predicted with the help of neural networks, to achieve better inference.

REFERENCES

[1] SpaceNet on AWS, Available at https://spacenetchallenge.github.io/datasets/spacenetBuildings-V2summary.html, Accessed at 2020.

[2] Rahman Farnoosh, Gholamhossein Yari and Behnam Zarpak, "Image Segmentation using Gaussian Mixture Models", *International Journal of Engineering Science*, Vol. 19, No. 2, pp. 1-16, 2008.

[3] German Valenzuela, M. Emre Celebi and Gerald Schaefer, "Colour Quantization using Coreset Sampling", *Proceedings of IEEE International Conference on Systems, Man and Cybernetics*, pp. 1-12, 2018.

[4] Timur Sattarov and Mircea-Serban Pavel, "On the Importance of Preprocessing and Initialization in K-Means", *Proceedings of IEEE International Conference on Data Mining and its Applications*, pp. 1-12, 2015.

[5] M. Emre Celebi, "Effective Initialization of K-means for Color Quantization", *Proceedings of IEEE International Conference on Image Processing*, pp. 1649-1652, 2009.

[6] Sudhir Singh and Nasib Singh Gill, "Analysis and Study of K-Means Clustering Algorithm", *International Journal of Engineering Research and Technology*, Vol. 2, No. 7, pp. 2546-2551, 2013.

[7] Chinki, Chandhok, Soni Chaturvedi and Aleefia Khurshid, "An Approach to Image Segmentation using K-means Clustering Algorithm", *International Journal of Artificial Intelligence and Neural Network*, Vol. 2, No. 2, pp. 1-13, 2012.

[8] J. Peng, D. Zhang and Y. Liu, "An Improved Snake Model for Building Detection from Urban Aerial Images", *Pattern Recognition Letters*, Vol. 26, No. 5, pp. 587-595, 2005.

[9] S. Ahmadi, M.V. Zoej, H. Ebadi, H.A. Moghaddam and A. Mohammadzadeh, "Automatic Urban Building Boundary Extraction from High Resolution Aerial Images using an Innovative Model of Active Contours", *International Journal of Applied Earth Observation and Geoinformation*, Vol. 12, No. 3, pp. 150-157, 2010.

[10] P. Kaiser, J.D. Wegner, A. Lucchi, M. Jaggi, T. Hofmann and K. Schindler, "Learning Aerial Image Segmentation from Online Maps", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 55, No. 11, pp. 6054-6068, 2017.

[11] Thiago Statella and Erivaldo Antônio da Silva, "Shadows and Clouds Detection in High Resolution Images using Mathematical Morphology", *Proceedings of International Annual Conference on Imaging and Geospatial Information Society*, pp. 1-7, 2018.

[12] Tapas Kanungo and Nathan S. Natanyahu, "An Efficient K-means Clustering Algorithm: Analysis and Implementation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 7, pp. 1-14, 2002.

[13] Henryk Palus, "Color Image Quantization by the K-Means Algorithm", in Workshop Farbbildverarbeitung, 7-8 October 2004.

[14] M. Emre Celebi, "Improving the Performance of K-Means for Color Quantization", *Image and Vision Computing*, Vol. 29, No. 4, pp. 260-271, 2011.

[15] Liu Jun and Zhang Haili, "Image Segmentation using a Local GMM in a Variational Framework", *Journal of Mathematical Imaging and Vision*, Vol. 46, No. 2, pp. 1-15, 2013.