

# Studies on Colour Image Segmentation Method Based on Finite Left Truncated Bivariate Gaussian Mixture Model with K-Means

P.SRINIVASA RAO

*Received: 27 August 2011 Accepted: 18 September 2011 Published: 30 September 2011*

## Abstract

Colour Image segmentation is one of the prime requisites for computer vision and analysis. Much work has been reported in literature regarding colour image segmentation under HSI colour space and Gaussian mixture model (GMM). Since the Hue and Saturation values of the pixel in the image are non-negative. And may not be meso-kurtic, it is needed left truncate the Gaussian variate and is used to represent these two features of the colour image. The effect of truncation can not be ignored in developing the model based colour image segmentation. Hence in this paper a left truncated bivariate Gaussian mixture model is utilized to segment the colour image. The correlation between Hue and Saturation plays a predominant role in segmenting the colour images which is observed through experimental results. The expectation maximization (EM) algorithm is used for estimating model parameters. The number of image segments can be initialization of the model parameters are done with K-means algorithm. The performance of the proposed algorithm is studied by calculating the segmentation performance techniques like probabilistic rand index (PRI), global consistency error (GCE) and variation of information (VOI). The utility of the estimated joint probability density function of feature vector of the image is demonstrated through image retrievals. The image quality measures obtained for six images taken from Berkeley image dataset reveals that the proposed algorithm outperforms the existing algorithms in image segmentation and retrievals.

**Index terms**— Image Segmentation, Hue, Saturation, Finite Left Truncated Bivariate Gaussian distribution, K-means algorithm, Image Quality Metrics, EM- algorithm. groups. It is an important technology for image processing and understanding. The structural characteristics of objects and surfaces in an image can be determined by segmenting the image using image domain properties. One of the major advantages of image segmentation is denoising. Denoising is the process of removing unwanted noise from the image. Segmentation specifically attempts to separate structure from noise on a local scale. It is one of the most important steps in computer vision and analysis.

For the last three decades lot of work has been reported in literature regarding image segmentation methods (Lucchese L. et al (2001), Srinivas Y. and Srinivas Rao K. (2007), Majid Fakheri et al (2010), Siddhartha Bhattacharyya (2011)). The image segmentation methods can be divided into two categories depending upon the type of image. The images can be broadly categorized into two types namely, gray level images and colour images. A gray level image is usually characterized by pixel intensity (Farag A..A.. et al (2004), Seshashayee M. et al (2011), Srinivas Yerramalle et al (2010)). But in colour images the colour is a perceptual phenomenon related to human response to different wavelengths in the visible electro-magnetic spectrum. In colour images the features that represent the image pixel are highly influenced by three feature descriptions namely, intensity, colour and texture. Among these features colour is the most important one in segmenting the colour images since intensity and texture features also be embedded in colour features. (Fesharaki and Hellestrand (1992), Kato Z. et al (2006), Kang Feng et al (2009), Kaikuo Xu et al (2011)). A better colour space than the RGB space in representing the colours of human perception is the HSI space, in which the colour information is represented





169 k? is ( ) ( ) ( ) ( ) 2 2 1 2 2 2 1 2 2 2 1 2 2 2 2 2 **22 1 2** ( ) **1 1 1 (1 1 1 ( , ; ) = ( ) 0 1 k s k s k k s s k k k**  
 170 k s k s k k k k k k k N l x y t x y s x y D F E ? ? μ μ ? ? ? ? μ μ ? ? ? ? ? ? ? ? ? ? ? ? ? + ? ? ?  
 171 ? ? ? ? ? ? ? ? ? ? ? ? ? ? ? ? ? ? ? ? ? ? ? ? ? + ? ? + + ? + + = ? ? ? ? ? ? ? ? ? ? ? ? ? ? ? ? ? ? (17)

173 )k s s l t x y ?

[illegible][illegible]

Solving equations (12), (23), (24), (15), (16) and (17-20)). This method performs well, if the sample size is large, and the computation time is heavily increased. When the sample size is small, some small regions may not be sampled. To overcome this problem, we use K-means algorithm to divide the whole image into various homogeneous regions. In K-means algorithm the centroids of the clusters are recomputed as soon as pixel joins the cluster. The initial values of  $i$  can be taken as  $i = K - 1$ .

, where, K is the number of image regions obtained from the K-means algorithm (Rose H. Turi (2001)). K-means algorithm uses an iterative procedure that minimizes the sum of distances from each object to its cluster centroid, over all clusters. This procedure consists of the following steps.

1) Randomly choose K data points from the whole dataset as initial clusters. These data points represent initial cluster centroids. 2) Calculate Euclidean distance of each data point from each cluster centre and assign the data points to its nearest cluster centre. 3) Calculate new cluster centre so that squared error distance of each cluster should be minimum. 4) Repeat step 2 and 3 until clustering centers do not change. 5) Stop the process.

The efficiency of the EM-algorithm in estimating the parameters is heavily dependent on the number of regions in the image. The number of mixture components taken for K-means algorithm is obtained,

197 **3 MEANS**

In the above algorithm, the cluster centers are only updated once all points have been allocated to their closed cluster centre. The advantage of K -means are that it is a very simple method, and it is based on intuition about the nature of a cluster, which is that the within cluster error should be as small as possible. The disadvantage of this method is that the number of clusters must be supplied as a parameter, leading to the user having to decide what the best number of clusters for the image is ?? ? for each image region and with the method of moments given by Bengt Muthen (1990) for Truncated Bivariate Normal Distribution with initial parameters as?  $i = 1/K$  for  $i=1,2,...,K$   $\mu = 1k$

$\bar{x}$  is the  $k$  th region sample mean of the Hue angle. Substituting these values as the initial estimates, we obtain the refined estimates of the parameters by using the EM-algorithm.

V.

## 208 4 SEGMENTATION ALGORITHM

[illegible]216 **5 EXPERIMENTAL RESULTS**

To demonstrate the utility of the image segmentation algorithm developed in this paper, an experiment is conducted with six images taken from Berkeley images dataset (<http://www.eecs.berkeley.edu/step4>) Assign each pixel into the corresponding  $j$ th region (segment) according to the maximum likelihood of the  $j$ th component  $L_j$ .

© 2011 Global Journals Inc. (US) Global Journal of Computer Science and Technology Volume XI Issue XVIII Version I

4© 2011 Global Journals Inc. (US) Global Journal of Computer Science and Technology Volume XI Issue XVIII Version I 28



Figure 1:

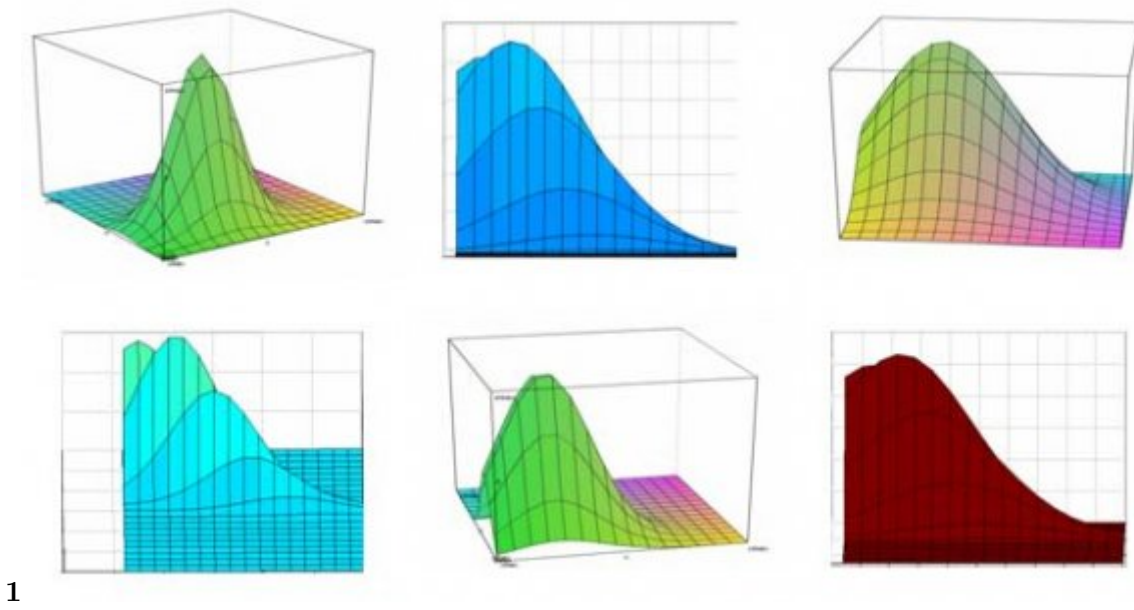


Figure 2: 1

| IMAGE                | OSTRICH  | POT      | TOWER    | BEARS    | DEER     | BIRD     |
|----------------------|----------|----------|----------|----------|----------|----------|
| <b>Estimate of K</b> | <b>2</b> | <b>3</b> | <b>4</b> | <b>2</b> | <b>4</b> | <b>3</b> |

Figure 3:

## 5 EXPERIMENTAL RESULTS

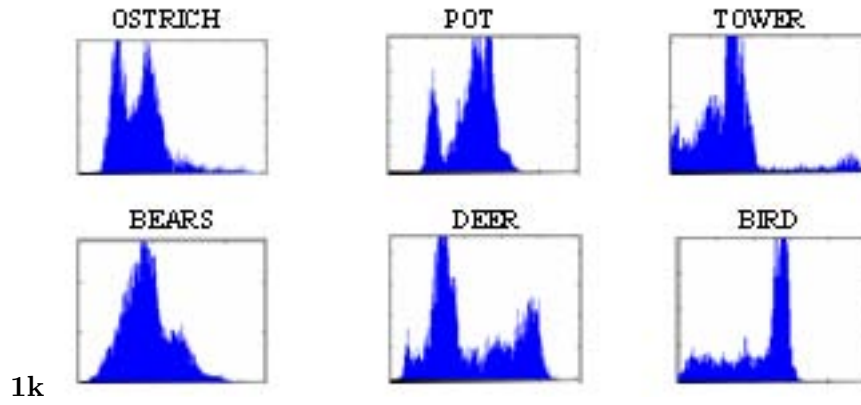


Figure 4:  $2k \mu = 1k y$

| Table-2a                                             |                                                |        |                                                |        |
|------------------------------------------------------|------------------------------------------------|--------|------------------------------------------------|--------|
| Estimated Values of The Parameters For OSTRICH Image |                                                |        |                                                |        |
| Number of image regions ( $K=2$ )                    |                                                |        |                                                |        |
| Parameters                                           | Estimation of Initial Parameters by $K$ -means |        | Estimation of Final Parameters by EM-Algorithm |        |
|                                                      | Regions( $i$ )                                 |        | Regions( $i$ )                                 |        |
|                                                      | 1                                              | 2      | 1                                              | 2      |
| $\alpha_i$                                           | 1/2                                            | 1/2    | 0.2627                                         | 0.7373 |
| $\mu_{1i}$                                           | 0.1781                                         | 0.1940 | 0.2054                                         | 0.2798 |
| $\mu_{2i}$                                           | 0.3321                                         | 0.7613 | 0.2505                                         | 0.7775 |
| $\sigma_{1i}^2$                                      | 0.0016                                         | 0.0004 | 0.0287                                         | 0.0772 |
| $\sigma_{2i}^2$                                      | 0.0126                                         | 0.0207 | 0.0747                                         | 0.0768 |
| $\rho_i$                                             | -0.4310                                        | 0.6996 | -0.6163                                        | 0.2840 |

Figure 5: Step 1 )  $2 \text{ } 1k ?$  ,  $2 \text{ } 2k ?$  ,  $k ?$  and  $k ?$  for  $i= 1 , 2$  ,

| Table-2b                                         |                                                |        |         |                                                |        |        |
|--------------------------------------------------|------------------------------------------------|--------|---------|------------------------------------------------|--------|--------|
| Estimated Values of The Parameters For POT Image |                                                |        |         |                                                |        |        |
| Number of image regions ( $K=3$ )                |                                                |        |         |                                                |        |        |
| Parameters                                       | Estimation of Initial Parameters by $K$ -means |        |         | Estimation of Final Parameters by EM-Algorithm |        |        |
|                                                  | Regions( $i$ )                                 |        |         | Regions( $i$ )                                 |        |        |
|                                                  | 1                                              | 2      | 3       | 1                                              | 2      | 3      |
| $\alpha_i$                                       | 1/3                                            | 1/3    | 1/3     | 0.4888                                         | 0.2019 | 0.3093 |
| $\mu_{1i}$                                       | 0.5532                                         | 0.4946 | 0.1517  | 0.5505                                         | 0.5223 | 0.3089 |
| $\mu_{2i}$                                       | 0.2168                                         | 0.1125 | 0.1219  | 0.1958                                         | 0.0810 | 0.1106 |
| $\sigma_{1i}^2$                                  | 0.0004                                         | 0.0027 | 0.0029  | 0.0248                                         | 0.0269 | 0.6202 |
| $\sigma_{2i}^2$                                  | 0.0008                                         | 0.0018 | 0.0035  | 0.0358                                         | 0.0328 | 0.0469 |
| $\rho_i$                                         | 0.1666                                         | 0.3570 | -0.7230 | -0.4604                                        | 0.9867 | 0.1373 |

Figure 6:

| Table-2c                                           |                                                |         |        |         |                                                |        |         |        |
|----------------------------------------------------|------------------------------------------------|---------|--------|---------|------------------------------------------------|--------|---------|--------|
| Estimated Values of The Parameters For TOWER Image |                                                |         |        |         |                                                |        |         |        |
| Number of Image Regions ( $K=4$ )                  |                                                |         |        |         |                                                |        |         |        |
| Parameters                                         | Estimation of Initial Parameters by $K$ -means |         |        |         | Estimation of Final Parameters by EM-Algorithm |        |         |        |
|                                                    | Regions( $i$ )                                 |         |        |         | Regions( $i$ )                                 |        |         |        |
|                                                    | 1                                              | 2       | 3      | 4       | 1                                              | 2      | 3       | 4      |
| $\alpha_i$                                         | 1/4                                            | 1/4     | 1/4    | 1/4     | 0.1999                                         | 0.1523 | 0.1872  | 0.4606 |
| $\mu_{1i}$                                         | 0.1519                                         | 0.5699  | 0.1505 | 0.5738  | 0.4011                                         | 0.6276 | 0.1616  | 0.5743 |
| $\mu_{2i}$                                         | 0.1937                                         | 0.2789  | 0.6176 | 0.7724  | 0.2408                                         | 0.3519 | 0.5047  | 0.7721 |
| $\sigma_{1i}^2$                                    | 0.0033                                         | 0.0073  | 0.0011 | 0.0006  | 0.1268                                         | 0.1045 | 0.0043  | 0.0213 |
| $\sigma_{2i}^2$                                    | 0.0148                                         | 0.0148  | 0.0291 | 0.0059  | 0.1055                                         | 0.1696 | 0.6936  | 0.0539 |
| $\rho_i$                                           | 0.1561                                         | -0.0259 | 0.0386 | -0.1086 | 0.9061                                         | 0.7744 | -0.1902 | 0.1944 |

Figure 7: Figure 2 : 2 1i ? , 2 2i?

| Table-2d                                           |                                                |         |                                                |        |
|----------------------------------------------------|------------------------------------------------|---------|------------------------------------------------|--------|
| Estimated Values of The Parameters For BEARS Image |                                                |         |                                                |        |
| Number of Image Regions ( $K=2$ )                  |                                                |         |                                                |        |
| Parameters                                         | Estimation of Initial Parameters by $K$ -means |         | Estimation of Final Parameters by EM-Algorithm |        |
|                                                    | Regions( $i$ )                                 |         | Regions( $i$ )                                 |        |
|                                                    | 1                                              | 2       | 1                                              | 2      |
| $\alpha_i$                                         | 1/2                                            | 1/2     | 0.4531                                         | 0.5469 |
| $\mu_{1i}$                                         | 0.4787                                         | 0.2364  | 0.4867                                         | 0.4067 |
| $\mu_{2i}$                                         | 0.4532                                         | 0.2600  | 0.4171                                         | 0.3375 |
| $\sigma_{1i}^2$                                    | 0.0027                                         | 0.0154  | 0.0560                                         | 0.0667 |
| $\sigma_{2i}^2$                                    | 0.0129                                         | 0.0170  | 0.0786                                         | 0.1439 |
| $\rho_i$                                           | 0.2044                                         | -0.6378 | 0.1263                                         | 0.6274 |

Figure 8:

Figure 9:

and  $k$  ? .

?  $1k \mu$   
 $, 2k \mu$  ,  
 $2 1k ?$  ,  
 $2 ? 2k$

by plotting the histogram of the pixel intensities of the whole image. The mixing parameter  $k$  ? and the region parameters  $1k \mu$  ,  $2k \mu$  ,  $2 1k ?$  ,  $2 2k ?$  ,  $k$  ? are unknown as prior. A commonly used method in initializing parameters is by drawing a random sample in the entire image data (McLachlan G. and Peel D. (

Figure 10:



Substituting the final estimates of the model parameters, the probability density function of the feature vector of each image are estimated. Using the estimated probability density functions and the image segmentation algorithm given in section V, the image segmentation is done for each of the six images under consideration. The original and segmented images are shown in Figure ??.

## .1 PERFORMANCE EVALUATION

After conducting the experiment with the image segmentation algorithm developed in this paper, its performance is studied. The comparison is based on three performance measures namely, Probabilistic Rand Index (PRI) given by Unnikrishnan R. and et.al ??2007), the Variation of Information (VOI) given by Meila M. (2005), and Global Consistency error (GCE) given by Martin D. and et al (2001). The objective of the segmentation methods are based on regional similarity measures in relations to their local neighborhood.

The performance of developed algorithm using finite left truncated bivariate Gaussian mixture model (FLTBGMM) is studied by computing the segmentation performance measures namely, PRI, GCE and VOI for the six images under study. The computed values of the performance measures for the developed algorithm and the earlier existing finite Gaussian mixture model (GMM) with K-means algorithm are presented in Table ?? for a comparative study. ??, it is observed that all the image quality metrics for the images are meeting the standard criteria. This implies that using the proposed algorithm the images are retrieved accurately. A comparative study of proposed algorithm with that of algorithm based on finite Gaussian mixture model (GMM) and Finite left truncated bivariate Gaussian mixture model with K-means reveals that the mean square error of the proposed model is less than that of the finite GMM and FLTBGMM. Based on all other quality metrics also it is observed that the performance of the proposed model in retrieving the images is better than the finite Gaussian mixture model.

## .2 VIII.

## .3 CONCLUSION

In this paper we introduce a novel and new colour image segmentation method based on left truncated bivariate Gaussian mixture model. Here it is assumed that the colour image is characterized by HSI colour space, in which the Hue and Saturation values are non negative. they are characterized by left truncated Bivariate Gaussian mixture model. The left truncated bivariate Gaussian distribution includes the Bivariate Gaussian distribution is a limiting case when the truncation points tends to infinite. It also includes several platy, meso, lefty and skewed distributions as particular cases for different values of the parameters. The model parameters are estimated by using EMalgorithm. The initialization and the number of image segments are determined through K-means algorithm and moment method of estimation. The segmentation algorithm is developed with component maximum likelihood. The experimentation with six colour images reveals that this algorithm outperforms the existing algorithms in both image segmentation and image retrievals. The image quality metrics also supported the utility of the proposed algorithm. It is possible to develop image segmentation algorithm with finite mixture of doubly truncated multivariate Gaussian distribution with more image features which require further investigations.

[Bhattacharyya ()] ‘A Brief Survey of Color Image Preprocessing and Segmentation Techniques’. Siddhartha Bhattacharyya . *Journal of Pattern Recognition Research* 2011. p. .

[Martin et al. ()] ‘A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics’. D Martin , C Fowlkes , D Tal , MalikJ . *proc. 8th Int. Conference Computer vision*, (8th Int. Conference Computer vision) 2001. 2 p. .

[Kato and Ting-Chuen ()] ‘A Markov random field image segmentation model for color textured images’. Z Kato , Pong Ting-Chuen . *Image and Computing Vision* 2006. 24 (10) p. .

[Paulinas and Usinskas ()] ‘A survey of genentic algorithms applications for image enhancement and segmentation’. Mantas Paulinas , Andrius Usinskas . *Information Technology and control* 2007. 36 (3) p. .

[Xu et al. ()] ‘An MDL Approach to Color Image Segmentation’. Kaikuo Xu , Hongwei Zhang , Tianyun Yan , Wei Wei , Shaomin , Wen Qiang . *International Conference on Multimedia and Signal Processing*, 2011. 2 p. .

[Rose and Turi ()] *Cluster Based Image Segmentation*, H Rose , Turi . 2001. Australia. Monash University (phd Thesis)

[Sujaritha and Annadurai ()] ‘Color image segmentation using Adaptive Spatial Gaussian Mixture Model’. M Sujaritha , S Annadurai . *International journal of signal processing* 2010. 6 (1) p. .

[Lucchese and Mitra ()] ‘Color image segmentation: A state-of art survey’. L Lucchese , S K Mitra . *Proc. Indian National Science Academy (INSA-A)*, (Indian National Science Academy (INSA-A)) 2001. 67 p. .

[Lee et al. ()] ‘Color shift model-based image enhancement for digital multifocusing based on a multiple color-filter aperture camera’. E Lee , W Kang , S Kim , J Paik . *IEEE Trans. On Consumer Electronics* 2010. (2) p. .

## 5 EXPERIMENTAL RESULTS

---

- [Ferri and Vidal ()] ‘Colour image segmentation and labelling through multieditcondensing’. F Ferri , E Vidal . *pattern Recognition Letters* 1992. 13 (8) p. .
- [Patra et al. ()] ‘Combining GLCM Features and Markov Random Field Model for Colour Textured Image Segmentation’. Dipti Patra , J Mridula , K Kumar . *Int. Conf. on Devices and Communications (ICDeCom)*, 2011. p. .
- [Meila ()] ‘Comparing Clustering -An axiomatic view’. M Meila . *proc.22nd Int. Conf. Machine Learning*, (.22nd Int. Conf. Machine Learning) 2005. p. .
- [Johnson et al. ()] *Continuous Univariate Distributions” Volume-I*, Norman L Johnson , Samuel Kortz , Balakrishnan . 1994. New York: John Wiley and Sons Publications.
- [Majid Fakheri et al. ()] ‘EM segmentation algorithm for colour image retrieval’. Majid Fakheri , T Sedghi , M C Amirani . *6th Iranian Conference on Machine Vision and Image Processing*, 2010. p. .
- [Feng et al. ()] ‘Flame Color Image Segmentation Based on Neural Network’. Kang Feng , Wang Yaming , Zhao Yun . *International Forum on Computer Science-Technology and Applications* 2009. p. .
- [Global Journal of Computer Science and Technology Volume XI Issue XVIII Version I 30] *Global Journal of Computer Science and Technology Volume XI Issue XVIII Version I 30*,
- [Eskicioglu and Fisher ()] ‘Image Quality Measures and their Performance’. A M Eskicioglu , P S Fisher . *IEEE Transactions On comm* 1995. 43 (12) p. .
- [Seshashayee et al. ()] ‘Image segmentation based on a finite generalized new symmetric mixture model with K-means’. M Seshashayee , K Srinivasa Rao , Srinivasa Satyanarayana Ch , P Rao . *Int. J. Computer Science Issues* 2011. 3 (8) p. .
- [Farnoosh et al. ()] ‘Image Segmentation using Gaussian Mixture Models’. Rahman Farnoosh , Gholamhossein Yari , Behnam Zarpak . *IUST International Journal of Engineering Science* 2008. No.1-2, 2008. 19 p. .
- [Shital Raut Raghuvanshi et al. ()] ‘Image Segmentation-A state-of-Art Survey for Prediction’. M Shital Raut Raghuvanshi , R Dharaskar , A Raut . *Advanced Computer control, ICACC’09. International Conference*, 2009. p. .
- [Mclanchlan and Krishnan ()] G Mclanchlan , T Krishnan . *The EM Algorithm and Extensions*, (New York) 1997. 1997. John Wiley and Sons.
- [Mclanchlan and Peel ()] A Mclanchlan , G Peel , D . *The EM Algorithm For Parameter Estimations*, (New York) 2000. 2000. John Wileyand Sons.
- [Muthen ()] ‘Moments of the censored and truncated bivariatenormal distribution’. Bengt Muthen . *British Journal of Mathematical and Statistical psychology* 1990. (43) p. .
- [Farag et al. ()] ‘Precise Image Segmentation by Iterative EM-Based Approximation of Empirical Grey Level Distributions with Linear Combinations of Gaussians’. A A Farag , A El-Baz , G Gimelfarb . *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW’04)*, (the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW’04)) 2004.
- [Fesharaki and Hellestrand ()] *Realtime color image segmentation*, M N Fesharaki , G R Hellestrand . SCS&E 9316. 1992. Australia. Univ. of New South Wales (Technical Report)
- [Haralick ()] ‘Survey: Image segmentation Techniques’. Shapiro Haralick . *CVGIP* 1985. 29 p. .
- [Sangwine and Horne ()] *The Colour Image Processing Hand Book*, S J Sangwine , R E N Horne . 1998.
- [Unnikrishnan et al. ()] ‘Toward objective evaluation of image segmentation algorithms’. R Unnikrishnan , C Pantofaru , M Hernbert . *IEEE Trans. Pattern Annl.Mach.Intell* 2007. 29 (6) p. .
- [Srinivas et al. ()] ‘Unsupervised Image Segmentation Based on Finite Generalized Gaussian Mixture Model With Hierarchical Clustering’. Y Srinivas , K Srinivasa Rao , Prasad Reddy , PV G D . *International journal for Computational vision and Biomechanics* 2010. 3 (1) p. .
- [SrinivasY ()] ‘Unsupervised image segmentation using finite doubly truncated gaussian mixture model and Hierarchical clustering’. SrinivasY , Srinivas . *Journal of Current Science* 2007. 93 (4) p. .