ELECTRONICS AND ELECTRICAL ENGINEERING
ISSN 1392-1215 2006. No. 8(72)
ELEKTRONIKA IR ELEKTROTECHNIKA

T170 ELECTRONICS
ELEKTRONIKA

Multilayer Transformation of Different Resolution for Colour Image Analysis and Segmentation

V. Paukštaitis, A. Dosinas
Departament of Applied Electronics, Kaunas University of Technology,
Studentu str. 50, LT-51368 Kaunas, Lithuania, tel.: 300270, e-mail: kat0102@ktu.lt

Introduction

Recent growth of computing power and price cutting of image acquisition hardware has quickened the development of colour image processing methods. Grayscale image pixel has two space coordinates and luminance in addition to them, colour images have two more colour components, thus making colour images more complex than grayscale ones. Following that, colour image processing becomes advanced, too.

Common automated image recognition tasks are presented in Fig. 1. In the begining, preprocessing is performed (enhancement in the aspect of further processing), the image segmentation follows it and only then the separated image objects are recognized. It may happen so, that the image objects are not recognized correctly during the first iteration. Thus, the recognition information is backed to change the segmentation or preprocessing tasks' parameters, consequently, the new iteration is performed.

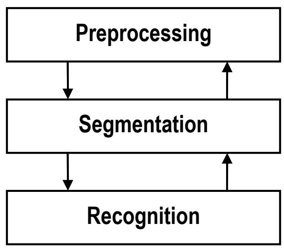

Fig. 1. Image recognition process

The main segmentation task is to isolate the recognizable object from the background or other objects in the image [1]

Formally, segmentation is defined as the partitioning, according to homogeneity predicate $P(\mathbf{S}_i)$, of image pixel set $\mathbf{F}$ into connected subsets or regions $(\mathbf{S}_1, \mathbf{S}_2, ..., \mathbf{S}_n)$ [2], that

$$\bigcup_{i=1}^{n} \mathbf{S}_i = \mathbf{F}\text{, kai } \mathbf{S}_i \cap \mathbf{S}_j = \varnothing \text{ ir } i \neq j. \qquad (1)$$

Homogeneity predicate $P(\mathbf{S}_i) = 1 \forall S_i$ and $P(\mathbf{S}_i \cup \mathbf{S}_j) = 0$, where $i \neq j$, and $\mathbf{S}_i$ and $\mathbf{S}_j$ are neighbours.

Many of image segmentation methods are developed using simple methodic and this resulted in the whole expanding bunch of segmentation methods for colour, grayscale or binary images. Although, the segmentation is psichovisual problem and general solution is hard to obtain [3], accordingly most of methods do not take the psichovisual aspects in concern.

First methods were simple and the knowledge about human visual system was not incorporated. Recent segmentation methods already apply the modern tidings about the biological system consciousness and the segmentation results become nearer to human observed ones.

The simpler segmentation methods are applied when the scene parameters are known or when the scene is prepared for the acquisition of simpler image composition [4]. These are mostly industrial applications. Recently, the need for image recognition in apriori unknown or variable surroundings (various medical images, aeronautic pictures and etc.) market is increasing. Thus, this publication is dedicated to analyzing questions concerning the shaping of the segmentation method incorporating human visual properties.

Colour Image Segmentation

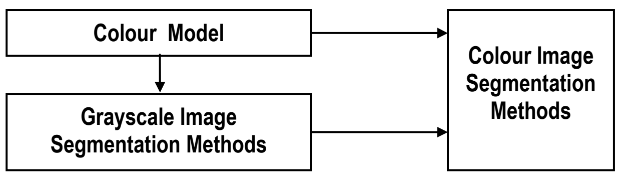

Fig. 2. Design of colour image segmentation methods

One of the simplest ways of making the colour image segmentation method is to use known grayscale image segmentation methods with applied colour model. Such methodic is presented in Fig. 2.

As it was mentioned, there are many methods created for image segmentation. According to mathematical methods applied, the created segmentation methods can be divided into such eight groups [2]:

- thresholding methods,
- feature space clustering methods,
- region-based methods,
- edge-detection methods,
- fuzzy-logic methods,
- artificial neural networks methods,
- physics-based methods,
- combinations of listed methods.

Thresholding methods are based on finding minimums in histograms and treating those minimums as the thresholds to partition the image. While, the colour images' histogram is 3D, such histogram is clustered and the cluster bounds treated as homogeneity predicate's bounds. The fuzzy-logic algorithms often are applied for the clustering.

The heart of region-based methods is sprinkle and growing of the initial regions, called seeds. The growing is performed according to homogeneity predicate. The image can be divided by some rules, but also according to homogeneity predicate. The edge-detection methods are used to find object edges and then apply them to the seed growing algorithm as the growing barrier.

Artificial neural networks are based on adaptive learning of special approximation algoritm features. This time, the extraction of features, selected during the experiments with the problem-specific images, is one of the prominent tasks. Then, the extracted features are supplied to the trained neural network, which indeed judges about the homogeneity of the given image part.

Physics based segmentation methods applies object's physical description, such as geometric form, reflectance and etc. These methods can precisely evaluate the object segmented, but they are very complex.

Mostly, the segmentation algorithms are very dependent to application field and every has it's own drawbacks [5], that's why it's very hard to concern method suitability to other image types. Methods are mostly frail to shadows, continiuous intensity and colour transitions in the image, variuos noises and etc. Consequently, the problem arises: how objectively compare the segmentation method work results from such a variety of the colour image segmentation methods.

Spatial Image Transformation

Human visual system (HVS) analyzes images flexibly. The main features of such system are central accurate vision and less accurate periferal vision, differential stimulus spread according to local illumination in the eye retina, stereoscopic vision, estimation of moving objects, some fractal positioning of image sensing units and interconnection between analythical neurons and etc. Such properties makes the humans to segment images properly. All the methods described above

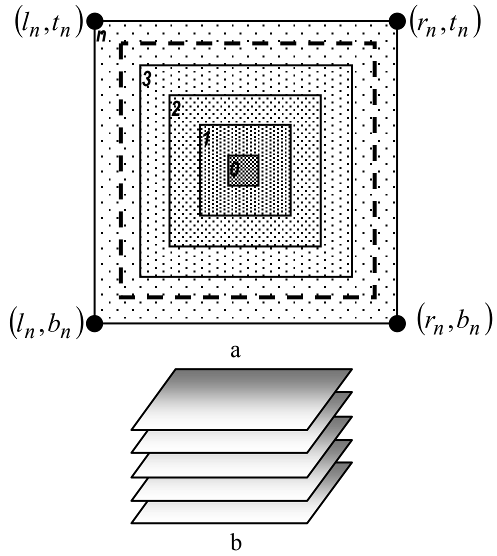

Fig. 3. The graphical interpretation of retinal transformation: a - the transformation bounds of the image, b - obtainable multilayered image

doesn't exploit the HVS features or incorporates them partially. However, for the segmentation quality or even image quality assessment problem's solution it is preferable to use methods which act as a human observer. The offered segmentation method applies some of the HVS featues.

In such case, the initial image is projected in such way that the resolution in the image centre is utmost (analogy to projection onto retina) and decreasing distinctively from the centre. To that end, the n-layered image is composed with layers involving increasing image space which in rectangular image case is defined by following variable system:

$$C_n(x_c, y_c, W, V) = \begin{cases} t_n = y_c - (2n+1)W/2, \\ l_n = x_c - (2n+1)V/2, \\ b_n = y_c + (2n+1)W/2, \\ r_n = x_c + (2n+1)V/2, \end{cases} \qquad (2)$$

where V and W are the 0-th layer rectangle image width and height, respectively, x_c, y_c - the centre coordinates of 0-th layer rectangle and t_n, l_n, b_n, r_n - n-th layer rectagle sides' parameters (Fig. 3a).

For the simplicity reasons, layers are presented in rectagular form, but they also could have other shape, for example, hexagonal form image layers.

After the layer boundaries calculation step (2), the cropped initial image $\mathbf{F}$ is defined as

$$\mathbf{h}_n = \lfloor \mathbf{F}_{ij} \rfloor, \qquad (3)$$

where $i = l_n, l_n + 1, ..., r_n$, $j = t_n, t_n + 1, ..., b_n$.

The neighbouring pixels are integrated to one pixel in the cropped image, consequently reducing the resolution by the rule:

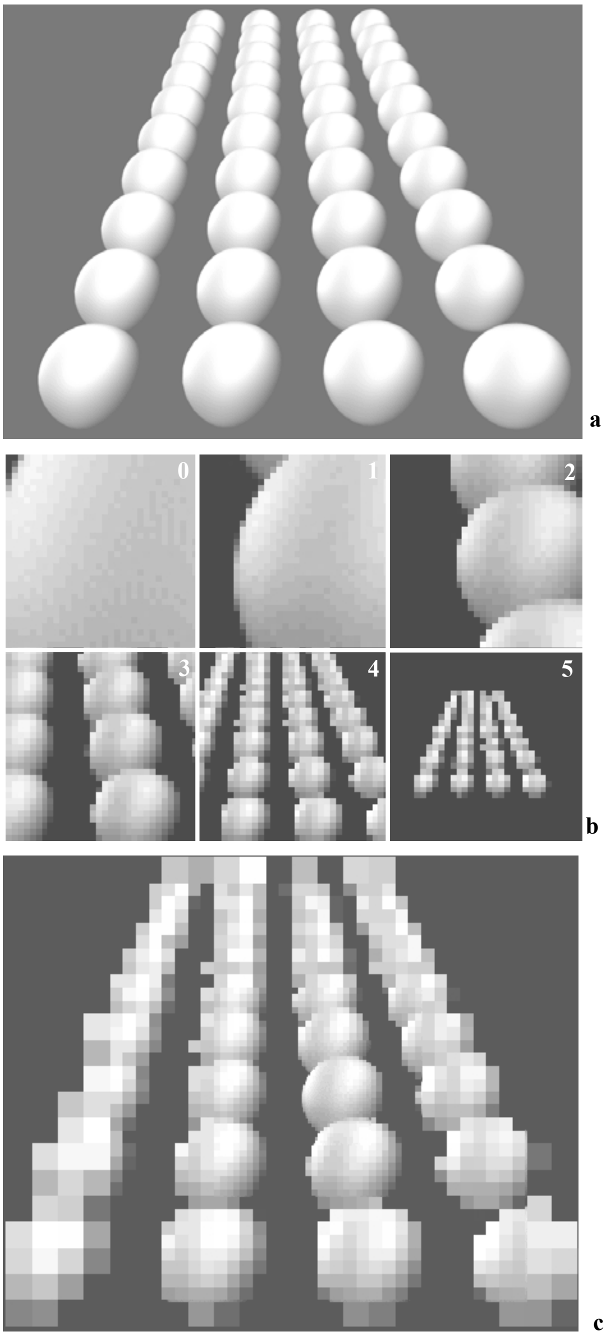

Fig. 4. Transformation example: a - initial image, b – obtained 6 layers of 32x32 size, c - layers back-mapped to initial space

$$p_{rc} = \sum_{i=1}^{M}\sum_{j=1}^{N} \mathbf{h}_n\left(i + (c-1)M,\, j + (r-1)N\right)\cdot \mathbf{q}_n(i,j), \qquad (5)$$

where $M = 2^{-n}V$, $N = 2^{-n}W$, $\mathbf{q}_n$ – weight matrix, r, c - matrix element row and column index respectively.

In the case of rectangular layer shape, every layer is the pixel set expressed in matrix form:

$$\mathbf{g}_n = \lfloor p_{ij} \rfloor, \qquad (6)$$

where $i = 1,2,...,W$, $j = 1,2,...,V$.

Then the whole transformation is the set of all obtained layers:

$$\mathbf{G} = \{\mathbf{g}_1,..,\mathbf{g}_K\}, \qquad (7)$$

where K - the number of obtained layers.

Every such image's *n*-th layer's resolution

$$R(n) = 2^{-n}, \qquad (8)$$

and increasing the *n*, $R^2(n)$ parameter shows the times the cropped image $\mathbf{h}_n$ is reduced. Because of $R(n)$ reduction the noise levels are moderated, but the spatial recogntion accuracy is also decreased.

The image transformed by such algorithm can be hierarchicaly, i.e. layer-by-layer segmented. The central part of the image has high resolution and can be incorporated in the detail analysis and the obtained results forwarded to higher layer analysis. This property allows to investigate objects‘ spread in some way or the continuity of the object boundaries. The image of pending layer, when $n > 0$, always include the image of lower layer, consequently the general analysis results of the pending layer and the segmentation results of the lower layer can be corrected. This way the image can be analyzed globally and locally.

The centre of 0-th layer can be moved to other parts of image during the analysis. Then the centres of other analyzed layers change, too. For this reason, analysis algorithm must incorporate a mechanism, which will prognoze the image sites of interest, i.e. sites where the centres should be moved.

The practical implementation of presented transformation is presented in Fig. 4. For simplicity reasons, all weights in $\mathbf{q_n}$ were considered equal to one. The central rectangle parameter taken as follows: *W*=32, *V*=32, $x_c = (W/2) + 46$, $y_c = V/2$, the size of initial image (Fig. 4 a) is 800x600.

Colour Space Model

The color image acquisition devices mostly output signals representing red (R), green (B), blue (B) colours. These R, G, B values can be directly used to display colour images in RGB colour. Unfortunately, different image formation devices can form different component values for the same image. RGB colour space is mostly used in television systems and image cameras and image monitors. But the RGB has the drawback, that is, all components are changing if the luminance is changing. Futhermore, the equal distances in RGB colour space does not guarantee equal perceptual colour differences. This makes the colour camparison in RGB colour space inconvenient. That's why the RGB colour space model is transformed into XYZ colour space [6]:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0{,}412453 & 0{,}35758 & 0{,}180423 \\ 0{,}212671 & 0{,}715160 & 0{,}072169 \\ 0{,}019334 & 0{,}119193 & 0{,}950227 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}. \qquad (9)$$

In this equation, the RGB vector is multiplied by experimentally concluded illuminant D65 matrix. Illuminant outlines the colour lighted on the photographed object. XYZ transformation is applied in other colour spaces where obtained colour differences correspond to the colour differences humans do understand. The example of such model is simplified CIE L*u*v* transformation

$$\begin{cases} L^* = 116 \cdot \sqrt[3]{Y/Y_0} - 16, \\ u^* = 13 \cdot L^* \left(\left(4X/\left(X + 15Y + 3Z \right) \right) - u_0 \right), \\ v^* = 13 \cdot L^* \left(\left(6Y/\left(X + 15Y + 3Z \right) \right) - v_0 \right), \end{cases} \qquad (10)$$

where the parameters with index 0 are the values corresponding the white light.

Feature Extraction

The segmentation can be interpreted as the recognition of image part because homogeneity predicate $P(S_i)$ does not limit the application of equation (1) just for colour components. The $P(S_i)$ attribute can be expressed in some parameter set, which outlines the image homogeneity, for example, the uniformity in some part of image. While it is needed to find the general case of segmentation operation, it is needed to find the feature set, which can objectively and fully describe the initail image.

Because of the HVS flexibility, it is considered, that it has such feature set and in this part of system models the visual cortex's V3 and V3A function, which majority of cells in them are responsive to lines of specific orientation, but they, too, are not interested in the colour of the oriented line [7]. Technically, the linear Gabor filter is applied [8]. The Gabor kernel is defined as follows:

$$Q_{\lambda,\sigma,\theta,\varphi}(x,y) = e^{-\frac{\tilde{x}^2 + \gamma^2 \tilde{y}^2}{2\sigma^2}} \cos\left(2\pi \frac{\tilde{x}}{\lambda} + \varphi \right), \qquad (11)$$

where the initial coordinates $(x,y) \in \Re^2$ are rotated by the θ angle $\tilde{x} = x\cos\theta + y\sin\theta$, $\tilde{y} = -x\sin\theta + y\cos\theta$, λ – spatial frequency, φ – phase bias, σ – dispersion, γ – factor, showing the diffuseness in y direction.

After the Gabor filtering, the threshold is considered for the particular type of images and filtered images are binarized, i.e. segmented:

$$B_\theta(x,y) = \begin{cases} 0, & I_G^\theta(x,y) < T, \\ 1, & I_G^\theta(x,y) \geq T, \end{cases} \qquad (12)$$

where $I_G^\theta(x,y)$ – processed image by Gabor filter with θ angle, T – threshold.

The two binarized images of different angular processing are then combined together applying logic AND operation:

$$\mathbf{K}_{\theta 1,\theta 2} = \mathbf{B}_{\theta 1} \cup \mathbf{B}_{\theta 2}, \qquad (13)$$

where $\mathbf{K}_{\theta 1,\theta 2}$ – joint image of $\theta 1$ and $\theta 2$ angles, $\mathbf{B}_{\theta 1}$, $\mathbf{B}_{\theta 2}$ – thresholded images of $\theta 1$ and $\theta 2$ angles. Such joint images are presented in the Fig. 5f.

The obtained joint images $\mathbf{K}$ are sequentially scanned and labeled, referencing to 4-neighbour environment of the pending pixel. Then, the segments having the longest and shortest axes approximately equal (considered circular or similar form shape segments) separated and the center of mass found:

$$x_k = \sum_{i=1}^{M_p} x_{ip} \Big/ M_p \,, \; y_k = \sum_{i=1}^{M_p} y_{ik} \Big/ M_p \,, \qquad (14)$$

where $p = 1,2,...P$, P – number of segments in labeled image, M_p – area of p-th segment, x_{ip}, y_{ip} – the values of the coordinates of the points belonging to p-th segment.

These cenrtes are used as spatial features for classifying the image part by some standard or reference image.

Practical Application

The proposed method has been applied to real scene images, that were specially selected such having the contamination of natural noises and the scene wasn't known apriori. One of the images is shown in Fig. 5a. This image has been transformed by (6) while the centre coordinates of the layers were selected manually and the image presented in Fig. 5b has been obtained. It is seen that the site of interest (this time, the registration plate) is represented sharply and it's surround is pointless. This way the inormation amount fordwarded to processing is much lower. The seven layers obtained form tranformation (6) are presented in Fig. 5c.

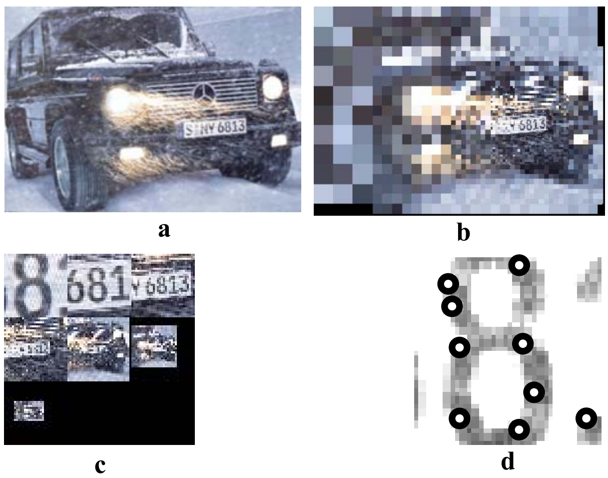

Fig. 5. Application example: a - real initial image, b - transformed initial image, c - the layers of the transformed image, d – $\mathbf{K}_{0°,90°}$ centres of the 0-th layer

On the next stage, the 0-th layer has been selected where the character "8" is clearly seen. The performance of the feature extraction algorithm is demonstrated applying high quality "8" character shown in Fig. 6a. This high quality image is considered the standard. Actually, the image presented in Fig 6c. is the initial standard image filtered through linear vertically oriented Gabor filter, selecting it's parameters to $\lambda = 0{,}15$, $\sigma = 0{,}5$, $\theta = 0°$, $\varphi = 0°$, $\gamma = 0{,}5$. Selecting the threshold T experimentally, this image has been

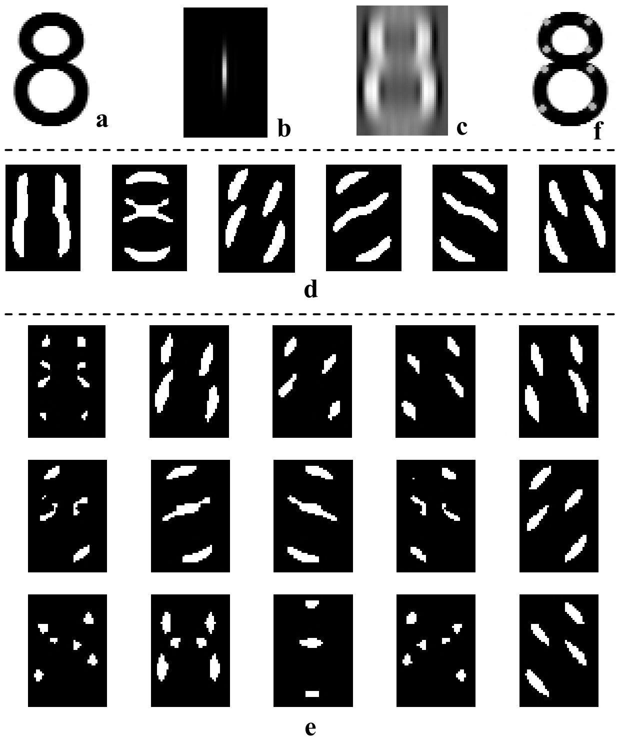

Fig. 6. Steps of feature extraction: a) reference image, b) Gabor kernel, c) filtered reference image, d) thresholded images, e) joint images, f) extracted features

segmented by (12). Filtering the initial image through Gabor filter with orientation angles of $0°, 90°, 30°, 60°, 120°, 150°$ and then segmenting them, the obtained images are shown in Fig 6d. The combinations by (13) of such orientationally filtered images are presented in fig. 6f.

The analyzed image's reference segments obtained by the feature extraction algorithm are all layed in Fig. 6e.

The image luminance matrix, i.e. L* component's matrix of the CIE L*u*v* colour model, has been applied for the spatial feature extraction.

As it is seen from the given application example, if the site of interest is pointed out then the extracted features are obtained in the character's locus and that concludes that the method works and partially emulates human visual system functions, that is to see the small details in the centre and large ones and to separate the lines of particular orientation and length, i.e. simulate the V3 function of the human visual cortex. This way, the image can be segmented not only applying the parameters of the segmentation method for the solution of the particular problem, but utilizing the spatial and colour features of the objects of that problem as the psichovisual information.

Conclusions

The segmentation methods of colour images are often developed advancing the grayscale image segmentation algorithms with application of the particular colour model. That's why, the amount of the segmentation methods of the colour images are increasing and the fittness of the segmentation method to the final goal is the tender question, which can be answered by applying the properties of the human visual system.

The introduced image processing transformation is based on the human retina receptors' distribution principle, when the central part has a high resolution, and in the peripheria the resolution is decreased distinctively from the designated centre. This lets the decreasement of the amount of information quantity and to analyze the image locally and globally.

The segmentation can often be interpreted as the recognition of specifical image site. The introduced segmentation method is based on the V3 function of the human visual cortex.

References

1. **Duda R. O., Hart P. E., Stork D. G**. Pattern Classification: Second Edition. – John Wiley & Sons, 2001.
2. **Cheng H. D. et al**. Color Image Segmentation: Advances and Prospects // Pattern Recognition, 2001. – Vol. 34. – P. 2259–2281.
3. **Fu K. S., Mui J. K**. A Survey on Image Segmentation // Pattern Recognition, 1981.– Vol. 13. – P. 3–16.
4. **Vaitkūnas M., Dosinas A., Bartkevičius V**. Spalvų spektrų įtakos kineskopo testinio vaizdo parametrų vertinimui modeliavimas // Elektronika ir elektrotechnika. – Kaunas: Technologija, 2003. – Nr. 4(46). – P. 34-38.
5. **Bačauskienė M., Verikas A.** The Evidence Theory Based Post-Processing of Colour Images // Informatica. – 2004. – Vol. 15. – No. 3. – P. 315–328.
6. **Verikas A., Malmquist K., Bačauskienė M., Bergman L.** Monitoring the De-Inking Process Through Neural Network-Based Colour Image Analysis // Neural Computing and Application. – 2000. – Vol. 9. – P. 142–151.
7. **Zeki S. M**. The Third Visual Complex of Rhesus Monkey Prestriate Cortex // Physiology Journal. – 1978.– No. 277.– P. 245–272.
8. **Grigorescu C., Petkov N., Westenberg M. A.** Contour Detection Based on Nonclassical Receptive Field Inhibition // IEEE Transactions on Image Processing. – 2003. – Vol. 12. – No. 7. – P. 729–739.

Submitted for publication 2006 06 29

V. Paukštaitis, A. Dosinas. Multilayer Transformation of Different Resolution for Colour Image Analysis and Segmentation // Electronics and Electrical Engineering. – Kaunas: Technologija, 2006. – No. 8(72). – P. 49–54.

The segmentation methods of colour images are often developed advancing the grayscale segmentation methods with application of the particular colour model. The short chracteristics and disadvantages of these methods are described shortly, in addition with common segmentation problems arising within their application. The publication presents the image segmentation problem analysis from psichovisual position offering an image transformation based on human visual system property to analyze the image sharply in the centre and pointlessly in the peripheria of the retina. Such transformation resolves the image to hierarchical layers with the different resolution dependentendly from the selected central location in the image. Colour analysis is performed through CIE L*u*v* colour space model. The spatial features for the segmentation are extracted by simulation of the human visual cortex's V3 function, which detects the lines of particular length and orientation. Such properties were implemented using the linear Gabor filter, then thresholding the obtained results and finding the centres of the reference segments. The application example demonstrates the performance of the transformation and feature extraction algorithm using the L* component's matrix of the colour model. Ill. 6, bibl. 8 (in Lithuanian; summaries in English, Russian and Lithuanian).

В. Паукштайтис, А. Досинас. Многослойная трансформация с различной разрешающей способностью для анализа и сегментации цветных изображений // Электроника и электротехника. – Каунас: Технология, 2006.– № 8(72). – С. 49–54.

Развитие методов сегментации цветных изображений очень часто основывается на методах сегментации полутоновых черно-белых изображений с использованием одной или другой цветовой модели. Приводятся краткие характеристики этих методов, их недостатки и общие проблемы, возникающие с использованием этих методов. В статье представляется анализ сегментации цветных изображений с психовизуальных позиций, предлагаеться трансформация изображения, основанная на свойстве зрительной системы человека анализировать изображение детально в центральной части сетчатки и менее детально – в периферийных областях. Такая трансформация, взависимости от выбранного в изображении места такой центральной части, разлагает изображение на иерархические слои, имеющие различную разрешающую способность. Цветовой анализ моделируется при помощи цветовой модели CIE L*u*v*. Пространстенные признаки сегмнтации формируется при помощи моделирования функции V3 визуальной коры мозга человека, выделять линии определённой длины и ориентации. Эти функции реализованы при помощи линейного фильтра Габора с преминением последующей пороговой операции сегментации и нахождения центров координат опорых сегментов. Работоспособность предложенной трансформации и алгорифма выделения признаков приводится примером применения, в котором используется компонента L* цветовой модели изображения. Ил. 6, библ. 8 (на английском языке; рефераты на английском, русском и литовском яз.).

V. Paukštaitis, A. Dosinas. Daugiasluoksnė skirtingos skiriamosios gebos transformacija spalvotiems vaizdams analizuoti ir segmentuoti // Elektronika ir elektrotechnika. – Kaunas: Technologija, 2006. – Nr. 8(72). – P. 49–54.

Spalvotų vaizdų segmentavimo metodai dažniausiai tobulinami remiantis nespalvotų vaizdų segmentavimo metodais, pritaikius jiems vienokį ar kitokį spalvų modelį. Pateikiamos trumpos šių metodų charakteristikos, jų trūkumai ir bendros segmentavimo problemos, iškylančios taikant tokius metodus. Straipsnyje pateikiama spalvotų vaizdų segmentavimo procedūros analizė iš psichovizualinės pozicijos, siūloma vaizdo transformacija, paremta žmogaus regos ypatybe analizuoti vaizdą tinklainės centre detaliai, o periferijoje – ne taip detaliai. Tokia transformacija priklausomai nuo parinktos centrinės vietos išskaido vaizdą į skirtingos skiriamosios gebos hierarchinius sluoksnius. Spalvų analizė modeliuojama CIE L*u*v* spalvų modeliu. Erdviniai segmentavimo požymiai formuojami modeliuojant žmogaus vizualinės smegenų žievės V3 funkciją išskirti tam tikro ilgio ir orientacijos linijas. Tai daroma naudojant tiesinį Gaboro filtrą, segmentuojant gautus signalus slenksčiu ir surandant atraminių segmentų centro koordinates. Pasiūlytos transformacijos ir požymių išskyrimo algoritmo veikimas demonstruojamas taikomuoju pavyzdžiu, panaudojant spalvų modelio L* dedamosios matricą. Il. 6, bibl. 8 (lietuvių kalba; santraukos anglų, rusų ir lietuvių k.).