

<https://doi.org/10.7236/IIBC.2017.17.2.147>

IIBC 2017-2-22

## 퍼지 유사성 기반 슈퍼-픽셀 생성

### Super-Pixels Generation based on Fuzzy Similarity

김용길\*, 문경일\*\*

Yong-Gil Kim\*, Kyung-II Moon\*\*

요 약 최근에는 슈퍼-픽셀 (super-pixel)은 컴퓨터 발전 응용에 널리 사용되고 있다. 슈퍼 픽셀 알고리즘은 픽셀을 지각적으로 실행이 가능한 영역으로 변환하여 그리드 픽셀의 경직된 특징을 줄일 수 있다. 특히, 슈퍼 픽셀은 깊이 추정, 골격 작업, 마디 라벨링 및 기능 국소화 등에 사용된다. 그러나 이러한 작업을 수행하기 위해 우수한 슈퍼 픽셀 파티션을 생성하는 것은 쉽지 않다. 특히 슈퍼 픽셀은 비합, 지속, 폐쇄, 지각 불변과 같은 형태 측면을 고려할 때보다 의미 있는 특징을 만족시키지는 못한다. 본 논문에서는 단순 선형 반복 클러스터링과 퍼지 클러스터링 개념을 결합한 고급 알고리즘을 제안한다. 단순 선형 반복 클러스터링 기술은 이미지 경계, 속도, 메모리 효율이 기존 방법보다 높다. 그것은 형태 측면의 맥락에서 슈퍼 픽셀 형태에 대해 양호하게 작거나 규칙적인 특성을 제안하는 것은 아니다. 퍼지 유사성 측정은 제한된 크기와 이웃을 고려하여 합리적인 그래프를 제공한다. 보다 작고 규칙적인 픽셀을 얻으며 부분적으로 관련된 특징을 추출 할 수 있다. 시뮬레이션은 퍼지 유사성 기반 슈퍼 픽셀 생성은 사람의 이미지를 분해하는 방식으로 자연적 특징을 대표적으로 나타낸다.

**Abstract** In recent years, Super-pixels have become very popular for use in computer vision applications. Super-pixel algorithm transforms pixels into perceptually feasible regions to reduce stiff features of grid pixel. In particular, super-pixels are useful to depth estimation, skeleton works, body labeling, and feature localization, etc. But, it is not easy to generate a good super-pixel partition for doing these tasks. Especially, super-pixels do not satisfy more meaningful features in view of the gestalt aspects such as non-sum, continuation, closure, perceptual constancy. In this paper, we suggest an advanced algorithm which combines simple linear iterative clustering with fuzzy clustering concepts. Simple linear iterative clustering technique has high adherence to image boundaries, speed, memory efficient than conventional methods. But, it does not suggest good compact and regular property to the super-pixel shapes in context of gestalt aspects. Fuzzy similarity measures provide a reasonable graph in view of bounded size and few neighbors. Thus, more compact and regular pixels are obtained, and can extract locally relevant features. Simulation shows that fuzzy similarity based super-pixel building represents natural features as the manner in which humans decompose images.

**Key Words** : Super-pixel, Fuzzy clustering, Gestalt aspects, Canny-Fuzzy Segmentation

#### I. Introduction

Ren and Malik discussed a preprocessing method

that builds an image into super-pixels, and mostly reduced the image complexity of further processing stage<sup>[16]</sup>. Hoiem, et al. showed very useful for depth

\*정회원, 조선이공대학교 컴퓨터보안과

\*\*정회원, 호남대학교 공과대학 컴퓨터공학과(교신저자)

접수일자 2017년 1월 15일, 수정완료 2017년 3월 15일

게재확정일자 2017년 4월 7일

Received: 15 January, 2017 / Revised: 15 March, 2017 /

Accepted: 7 April, 2017

\*\*Corresponding Author: kimoon@honam.ac.kr.

department of computer engineering, Honam university, korea

computing<sup>[8]</sup>. Sun et al. represented the cutout contour to real object boundaries with effect in case of low contrast edges<sup>[13]</sup>. He, et al. introduced a special type of the standardized segmentation method, using a special factor to transform an over-segmented image into super-pixels having an rough size<sup>[7]</sup>. Levinshtein, et al. introduced a super-pixel segmentation having multi-resolution to estimate medial point, and suggested a perceptual function to divide close medial objects likely to belong to the same medial branch<sup>[11]</sup>. Mori proposed the restricted common locations to lie at centers of super-pixels in view of human body. Joint positions reduce the state space, and give good masks for half-limbs features through pixel aggregation. Fulkerson, et al. suggested a classifier histogram of local features obtained in each super-pixel. It aggregates the feature histograms for each super-pixel, and the graph operations are enhanced on a random field for the super-pixel. Recently, Mohammad, et al. suggested a super-pixel method that adjusts the depth map through the closed boundaries. It showed actually good results<sup>[9]</sup>.

A geometric super-pixel method reserves local image boundary, and prevents under-segmentation due to compact limitations<sup>[3]</sup>. This method is very fast, and can be applied to megapixel sized images with high intensities. It shows good quality results on several complex images, and has less under-segmentation error than ones without the compact bounds. However, since it has a user-defined similarity measure between pixels just as any region segmentation, the method is general. Selection of the similarity measure depends on the related task. Recognition based segmentation emphasizes the objects in the image. Thus, segmentation task is done perfectly, but low level features such as color, texture and edges do not accept as contextual segmentation works. Unfortunately, most super-pixel methods are not easy to use, and have not high quality segmentations. They often have poor quality segmentation and inconsistent size and shape, etc.

In this study, we present a simple fuzzy clustering method to obtain a local pixel clustering in the 5-D space defined by the L, a, b values of CIELAB color space and xy-pixel coordinates. The CIELAB coordinates are related to the lightness of the color and its position between yellow and blue. The CIELAB color is appropriate for the color images with various noises and canny method is very good from various edge detection methods. Recently, clustering with swarm-based algorithms has been shown to give good results in many real-world applications<sup>[3]</sup>. Ant based clustering method shows good results for the segmentation of brain images<sup>[10]</sup>. Edge extraction from a digital image showed very interesting detection result. Ant based clustering using CIELAB color produces very natural outputs in view of human segmented image.

A fuzzy similarity measure provide more compact aspects and regular property to super-pixel shapes, and is uniformly used grayscale images as well as color ones. In particular, fuzzy similarity measure is simple to work and easily applied in practice. It uses the super-pixels as the clustering objects pixels, and it can extend the clustering details and remove the noise influence. In the preprocessing stage, an image is divided into super-pixels, and a further parsing step is used for the areas with larger gray changes over maximum entropy. Subsequently, designed fuzzy clustering is carried out to the fuzzy possibility of each super-pixel, and an iterative method is used to redefine their classifications. Experiment shows that the fuzzy clustering is significantly more efficient than competing methods, while producing segmentations of better quality as measured by standard boundary claim and under-segmentation error measures.

## II. Background

In graph based algorithm, each pixel is regarded as a node in a graph, and edge weight between two nodes

is proportional to the similarity between the pixels. Super-pixel segments are derived by effectively minimizing a cost function defined on the graph. In particular, since the standard cuts divide recursively a given graph with contour and texture features, they minimize entirely a cost function defined on the edges at the partitioned boundaries<sup>[17]</sup>. It is based on entropy theory and the image boundaries by the super-pixels to obtain support masks<sup>[15]</sup>. But it is computationally expensive for large images. Fezenszwalb and Huttenlocher presented another graph-based segmentation to obtain the super-pixels<sup>[5]</sup>. It represents an aggregative clustering of pixels on a graph, in which each segment is the minimum spanning tree of basic pixels. It is applied to depth estimation<sup>[8]</sup>, and is quite fast in practice as compared to watersheds<sup>[17]</sup>. However, it does not give a precise control on many super-pixels or their compact features. A super-pixel grid can be used by finding optimal path that divide the image in  $xy$ -directives<sup>[14]</sup>. It is possible to control the size, number, and compact property of the super-pixels, but, the quality and speed are dependent to previous boundary maps. Graph-based models like as conditional random fields can obtain fast speed increases when converting from pixel based graphs to super-pixels<sup>[6]</sup>, but irregular super-pixels can reduce the performance.

If the super-pixels are irregular, local features are less meaningful in case of using scale-invariant transform. Also, the results for several super-pixels are doubtful. This effect can be shown when we compare the performance of SLIC super-pixels to existing methods for two works: object class recognition and medical image segmentation<sup>[1]</sup>. Gradient ascents adjust the clusters from the previous loop to obtain better segmentation in each step. Mean-shift searches a mode that builds image segments recursively by going to the gradual kernel center for all objects in the feature space. The super-pixels building can be large or small on the input kernel factors. However, there is no direct control method over the number, size, or compact property of the super-pixels. Quick-shift segmentation

also searches a mode like mean-shift, and is faster in practice<sup>[4][18]</sup>. Each point is moved toward the nearest neighbor ones that updates Parzen probability density function in the feature space. It is not iterative method and does not allow one to adjust the size of super-pixels exactly. Quick-shift based super-pixel has been used in applications such as object localization and motion segmentation<sup>[2][6]</sup>. General watershed techniques perform gradient ascent from local minima in the image space in order to obtain watersheds<sup>[12]</sup>. Vincent and Soille used a fast way based on pixel queue. Lazy Snapping is applied to graph cuts on the super-pixels by this algorithm<sup>[13][19]</sup>.

SLIC segmentation builds super-pixels through pixel clustering using color similarity and proximity in the image plane<sup>[1][22]</sup>. It is done in 5-dimensional labxy space, where lab denotes the pixel color vector in CIELAB color space. CIELAB is considered as perceptual equality for small color distances, and  $xy$ -plane is the pixel location. In SLIC, a distance measure is simply sum of two distances which one is the distance in CIELAB color space and the other is the standardized distance in  $xy$ -plane. It starts by sampling a desired number of nearly equal sized super-pixels and moving ones to the seed locations that is the lowest gradient one. It does not locate them at an edge, and reduces the possibility of choosing a noisy pixel. It is a special case of  $k$ -means clustering selected for super-pixels generation. Super-pixel segmentation must have high boundary claims and low under-segmentation error. Usually, the boundary claim is measured as the proportion that ground truth edges include in a pixel of the least one super-pixel boundary. Under-segmentation is measured as the error which segments an image with respect to a ground truth. This error is calculated in terms of picking out the segment results placed over ground truth segmentations. Thus, it sets a detracted value to the super-pixel that does not nearly fit a ground truth segment. Given the ground truth segments  $g_1, g_2, \dots, g_M$  and the super-pixel results  $s_1, s_2, \dots, s_L$ , let  $o_{ij}$  be the

intersection or overlap error of the super-pixel  $s_j$  with respect to the ground truth segment  $g_i$ . Then, the under-segmentation error for a ground truth segment  $g_i$  is computed as:

$$U_e = \left\{ \sum_{i=1}^M \sum_{s_c} |s_j| - N \right\} / N \quad (1)$$

where  $|\times|$  is the segment size in pixels,  $N$  the size of the image in pixels, and  $s_c$  is corresponding to  $s_j$  given  $o_{ij}$  greater than the min number of pixels for overlapping. The min number of pixels for overlap is set to be 5% of  $|s_j|$  for small errors in ground truth segmentation.  $U_e$  is computed for each image of the ground truth and then averaged. Usually, experiments of SLIC segmentation represent the good super-pixels at a lower computational cost, and the segmentation quality outperforms the other methods in terms of the boundary claim and under-segmentation error<sup>[1]</sup>.

There are biased segmentations toward small regions<sup>[16]</sup>. Main reason is that a good segmentation  $s$  is a simple sum over each segment and the number of segments does not have statistically normal distribution. A prior distribution having segment size  $|s|$  is required to segment an image into approximately equal size and to control over the scale of the segmentation. Ren and Malik represented an empirical distribution of  $|s|$  with respect to the human segmented images as a log-normal prior distribution<sup>[16]</sup>. Then, an objective function for human segmented image is defined as:

$$J^*(s) = J(s) + \sum_{s \in S} \left[ -\frac{(\log |s| - \mu_s)^2}{2\sigma_s^2} \right] \quad (2)$$

where  $J(s)$  is the classifier function and  $J^*(s)$  denotes the sum of the classifier function over segments. This is corresponding to the optimization of  $J$  in the segmentation space  $S$ .

For the super pixels to be useful, they must be fast, easy to use, and produce high quality segmentations. Unfortunately, most super pixel methods do not satisfy all these requirements. Especially, negative figures of grouping must be constructed by randomly matching human segmentation. Natural segmented images are

used as positive figures. Also, in a preprocessing stage, since an image is segmented over into super pixels, a variety of features must be derived from the classical gestalt aspects, including contour, texture, brightness and good continuation. The gestalt aspects represent several concepts to guide our perceptual features: non-sum, similarity, simplicity, good continuation, closure, perceptual constancy and inconstancy. Non-sum means that whole is not sum of the image pixels. The appearance of figures depends closely on the organization of pixels than on each pixel. Similarity represents to group features at many pixels.

Goodness is closely related to symmetry, simplicity, smoothness, and regularity. Smooth contours are preferred over abrupt spatial changes. Some pixels that have a continuous pattern are to group together. Closure is the tendency to perceive complete figures even though certain information is vanished. Visual constancy is corresponding to see known objects as the same ones regardless of the perspective, distance, or lighting angles. To get the object constancy, the human sometimes compensates excessively, recognizing the features that do not exist. If we discern some kind of pattern by aggregating the perception pixels into larger figures, then each figure may itself contain many features. Such hierarchical structure can persist for several levels.

Fuzzy similarity measures are used to compare different kinds of objects such as images. Their definitions are based on proximity measures, operations on fuzzy sets etc. which makes different propositions of properties of similarity measures. The relevant properties we consider for the fuzzy similarity measures depend on the usefulness within the domain of research but they are not considered as complete. Sridevi and Nadarajan presented a new fuzzy similarity measure for generalized fuzzy sets<sup>[20]</sup>. The proposed measure worked successfully in situations where the generalized fuzzy sets have same center of gravity points and overcome the drawbacks of the existing methods. We see that giving a fuzzy definition for

distances between points of fuzzy sets very much improves the similarity measure than the geometric distances adopted by earlier methods. The measure greatly reduces the influence of inaccurate measures and provides a very intuitive quantification.

### III. Combining super-pixel and fuzzy clustering

Since the super-pixel uses spatial information, the related algorithm should have good performance in noise reduction. Also, it must be possible to trace the edge information of the original figure during the local consistency enhancement. The basic segmented area has some image features such as shape boundary and region contours. These features are not connected with the single pixel. Thus, the super-pixel can enhance the segmentation accuracy and processing time. Also, each gray level of the pixel is very similar in the super-pixels, and the inhomogeneous intensity exists in the super-pixels. But, the ambiguous feature is difficult to segment images well. Therefore, fuzzy clustering technique can solve this problem to a certain extent.

#### 1. Fuzzy Partitioning

Fuzzy partition can be regarded as a generalization of crisp partition, it allows  $\mu_{ik}$  to obtain real values in  $[0, 1]$ . A matrix  $U = [\mu_{ik}]$  represents the fuzzy partitions, and its condition is given by the following:

$$\sum_{k=1}^c \mu_{ik} = 1, 0 < \sum_{i=1}^N \mu_{ik} < N \quad (3)$$

Here  $i$ -th column of  $U$  contains the possibility of  $i$ -th fuzzy subset of the image. Equation 3 constrains the sum of each column to 1, and thus the total possibility of each pixel in the image equals one. The distribution of possibility among the fuzzy subsets is not constrained. The fuzzy clustering is based on the minimization of an objective function defined as the following:

$$J = \sum_{i=1}^c \sum_{k=1}^N \mu_{ik}^m (p_k - v_i)^T A_i (p_k - v_i), v_i \in R^n \quad (4)$$

Here  $v_i$  denotes a cluster center to be determined. Equation 3 can be regarded as a measure of the total variance of image pixel from the cluster center.

The minimization of equation 4 belongs to a nonlinear optimization problem to solve by using grouped coordinate minimization, simulated annealing and genetic algorithm, etc. The typical method is a simple Picard iteration through the first-order conditions for stationary points of equation 4, known as Fuzzy  $c$ -means algorithm. The stationary points of equation 4 are obtained by joining the equation 3 to  $J$  by Lagrange multipliers. Each cluster center has its own matrix  $A_i$ , which satisfies the following norm:

$$D_{ik} = (p_k - v_i)^T A_i (p_k - v_i) \quad (5)$$

Here  $A_i$  is used as an optimization variable, thus, allowing each center to adapt the distance norm to the local spatial structure.

Assume that  $A$  denote a  $c$ -tuple of the norm-based matrices,  $A = (A_1, A_2, \dots, A_c)$ . Then, the objective function is defined as the following:

$$J = \sum_{i=1}^c \sum_{k=1}^N \mu_{ik}^m D_{ik} \quad (6)$$

The objective function cannot be directly minimized with respect to  $A_i$ .  $A_i$  should be restricted in some way to obtain a feasible solution. Doing this is to restrict the determinant  $A_i$ . Allowing the matrix  $A_i$  to vary with fixed determinant value,  $s_i$  is corresponding to optimize the cluster center while its volume has constant value. Using the Lagrange multiplier, the expression for  $A_i$  is the following:

$$A_i = (s_i |C_i|)^{1/N} C_i^{-1} \quad (7)$$

Here  $C_i$  denotes a kind of fuzzy covariance matrix of  $i$ -th cluster center defined as the following:

$$C_i = \sum_{k=1}^N w_{ik} (p_k - v_i)(p_k - v_i)^T, w_{ik} = \mu_{ik}^m / \sum_{k=1}^N \mu_{ik}^m \quad (8)$$

There are two main approaches to determining the feasible number of clusters in the pixels. One is compatible cluster merging. This method starts with a large number of clusters, and reduces successively this number through cluster merging that is compatible to some criteria. The other method is a kind of goodness fit test, in which the pixels are grouped for different values of  $c$ , and use a validity measure to assess the goodness of the partitions. Though various fit measures have been introduced in the literatures, none of them is perfect by oneself. In the case of our super-pixel, classification entropy and partition index are comparatively appropriate. Classification entropy measures the vague property of the cluster partition, which is similar with the partition coefficient.

$$C_E = -\frac{1}{N} \sum_{i=1}^c \sum_{k=1}^N \mu_{ik} \log(\mu_{ik}) \quad (9)$$

Partition index measures the ratio of the sum of compactness and separation of the cluster centers. It uses a sum of individual cluster validities regularized through partition through the fuzzy cardinality of each cluster.

$$P_I = \sum_{i=1}^c \left[ \sum_{k=1}^N \mu_{ik}^m / N_i \cdot \|p_k - v_i\|^2 / \sum_{j=1}^c \|v_j - v_i\|^2 \right] \quad (10)$$

## 2. Fuzzy SLIC Super-pixel Algorithm

Assume that  $K$  is a number of approximately uniform-sized super-pixels. For an image with  $N$  pixels, the approximate size of each super-pixel is  $N = K$  pixels. For approximately equal-sized super-pixels, there would be a super-pixel center at every grid interval  $S = (N/K)^{1/2}$ . We choose  $K$  super-pixel cluster center vectors  $c_k = [l_k, a_k, b_k, x_k, y_k]^T$  ( $k = 1, \dots, K$ ) at regular grid interval  $S$ . Since the spatial extent of any super-pixel is approximately  $S^2$  (the approximate area of a super-pixel), we can assume that pixels related to this cluster center exist within a  $(2S) \times (2S)$  area around the super-pixel center on the  $xy$ -plane. It is corresponding to the study area for the pixels nearest

to each cluster center. Euclidean distances in CIELAB space are perceptually meaningful for small distances ( $m$  in Eq. 11).

If the spatial distances of pixels exceed this color distance, then they have larger pixel color similarities, and then give the results in super-pixels that do not reserve area boundaries in the image plane. A variable  $m$  is chosen in  $d_s$  to maintain the compact property of a super-pixel. The greater  $m$ , the more spatial proximity is emphasized and the more compact the cluster. So, we use a distance measure  $d_s$  defined as the followings:

$$\begin{aligned} d_s(k, i) &= d_{lab}(k, i) + (m/S) d_{xy}(k, i) \\ d_{lab}(k, i) &= \sqrt{(l_k - l_i)^2 + (a_k - a_i)^2 + (b_k - b_i)^2} \\ d_{xy}(k, i) &= \sqrt{(x_k - x_i)^2 + (y_k - y_i)^2} \end{aligned} \quad (11)$$

Then, our fuzzy clustering is the iterative minimization of the following objective function:

$$P_I = \sum_{i=1}^c \left[ \sum_{k=1}^N \mu_{ik}^m / N_i \cdot \|p_k - v_i\|^2 / \sum_{j=1}^c \|v_j - v_i\|^2 \right] \quad (12)$$

where  $\{y_i, i = 1, 2, \dots, N\}$  denote an image with  $N$  pixels to be divided into  $K$  super-pixel clusters,  $\{c_k, k = 1, 2, \dots, K\}$  every super-pixel cluster centers, and  $C = (c_1, c_2, \dots, c_K)$  the cluster center matrix. The vector  $U_k = (u_{1k}, u_{2k}, \dots, u_{ik})^T$  is the membership of the  $k$ -th pixel in  $i$  clusters,  $u_{ik}$  ( $u_{ik} \in [0, 1]$ ) the membership of the  $k$ th pixel in the  $i$ -th cluster, the  $U = (U_1, U_2, \dots, U_k)$  the membership matrix,  $p$  the membership function index that controls the fuzziness of resulting partitions, and  $\|\times\|^2$  is a norm metric which usually uses Euclidean measure. If the membership matrix is not used, super-pixel segmentation steps are:

- 1) Initialize the cluster center matrix  $C$  by sampling pixels at grid steps  $S$ .
- 2) Perturb cluster centers in an  $n \times n$  neighborhood matrix to the lowest gradient position. The radients  $G$  of image is computed as:  $G(x, y) = \|L(x+1, y) - L(x-1, y)\|^2 + \|L(x, y+1) - L(x, y-1)\|^2$ , where

$L(x, y)$  is the *lab* vector corresponding to the pixel at position  $(x, y)$ , and  $\|\times\|$  is the  $L_2$  norm.

- 3) Repeat step 4 through 6 until the error E is not greater than a given threshold.
- 4) For each cluster center  $c_k$ , perform the step 5:
- 5) Assign the optimal matching pixels from a  $(2S) \times (2S)$  neighborhood matrix around the cluster center according to the distance measure (Eq. 2).
- 6) Compute new cluster centers and residual error E.
- 7) In case that the membership matrix is available, step 3 is replaced as in the case that the value of  $J$  is less than the setting threshold. Steps 4 through 6 are replaced as the membership updates process. The membership function is updated as

$$\tilde{u}_{ik} = \left( \frac{1}{\|y_i - d_s(k, i)\|^2} \right)^{1/(p-1)} / \sum_{j=1}^K \left( \frac{1}{\|y_i - d_s(k, j)\|^2} \right)^{1/(p-1)} \quad (13)$$

The cluster centers are updated as

$$\tilde{d}_s = \sum_{i=1}^N \tilde{u}_{ik}^p y_i / \sum_{i=1}^N \tilde{u}_{ik}^p \quad (14)$$

#### IV. Experimental validation

The super-pixels should have the features with perceptually consistency, i.e. all pixels in a super-pixel are most likely uniform in color and texture, etc. Also, because super-pixels are the results of an over-segmentation, most structures of the image must be conserved. This experiment compares the simple SLIC method and our fuzzy clustering mixing method, in case that it is difficult to obtain good segmentations. First, Fig 1 (a) is an image from<sup>[1]</sup>, (b) is a negative image, (c) a mid-range stretched image, and (d) a log-transformed one. These works are corresponding to image enhancement in view of the classical gestalt views, including contour, texture, brightness and continuation. These works include negative function, power law transform, log transform, averaging, etc. (see Fig 2). If we want to segment an image in a manner similar to human segmentation, then we should

consider the manner in which human partitions the image.

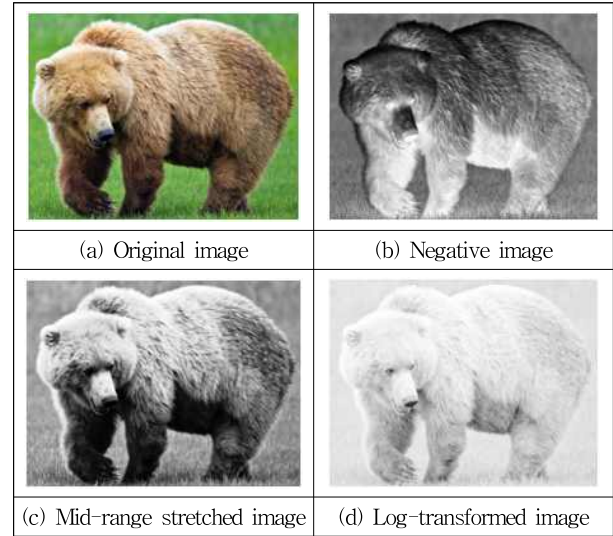


그림 1. 전형적인 형태 측면에서 이미지 향상  
 Fig. 1. Image Enhancement in view of typical gestalt aspects

The real images are generally exclusive entities whose alignment in a scene leads to an image with spatially coherent pixel groups. Each pixel group starts from a single object, and may not be visually coherent. “Ground truth” segmentations for images composed of a wide variety of natural scenes are usually obtained by human beings. However, the segmentation results by different humans for certain image are not identical. Even if two observers have the same perceptual approach to an image, they may segment at different levels. It implies that we must consider a consistency approach that does not distinguish these differences. Canny algorithm and intuitionistic fuzzy set for image segmentation have been successfully applied. Some wrong edges occur in canny algorithm, while using intuitionistic fuzzy approach for image edge detection, we can effectively segment the image pixels without affecting the integrity of the further image analysis.

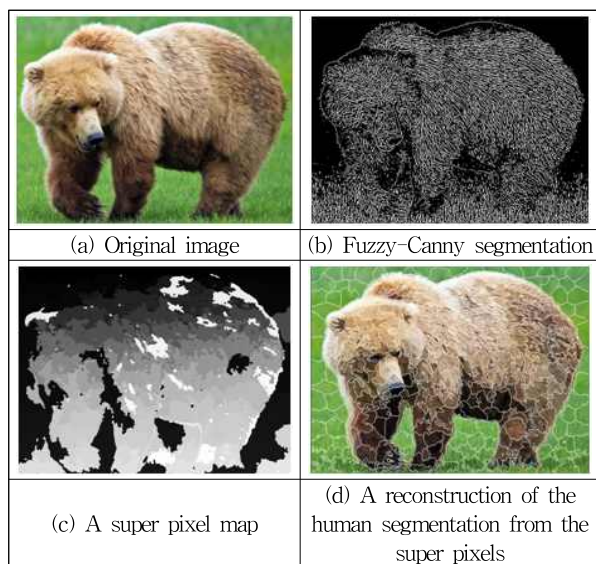


그림 2. 잘된 세분화 사이 분류  
 Fig. 2. Classification between Good Segmentations

There are image segmentation effects in a manner similar to human segmentation. Fig 2 illustrates the basic approaches. Fig 2 (b) is a result by fuzzy-Canny segmentation. It has an effect such as a human marked segmentation. Double threshold along with canny edge detector with respect to edge detection was used to identify the small objects in the image. Here, threshold plays an important role that extracts the clear image from unclear picture. Because of the uncertainties and dynamic property in various aspects of image processing, fuzzy processing is desirable. These uncertainties include additive and non-additive noise in low level image processing, imprecision in the underlying assumptions of the algorithms, and ambiguous interpretation during high level processing. For the usual edge detection represents features as intensity ridges. Through increasing the threshold greater than 50 and contrast can be enhanced. We used 16 fuzzy edge templates that get the possible direction of edges in the image, and then computed the divergence between the origin image and the 16 fuzzy templates. Fig 2 (c) is a super-pixel map, and (d) an image reconstruction of the human segmentation from the super pixels. The segmentation in (d) is perceptually a good image, but over-segmentations are

appeared. Many hidden objects can be identified using edge detection which gives major points in actual truth behind the images.

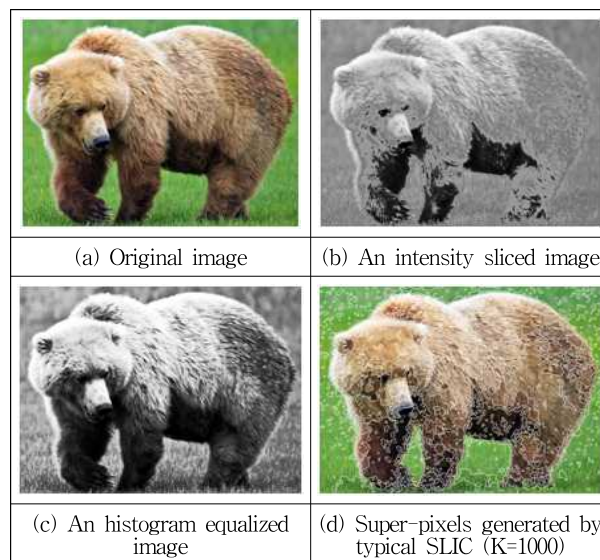


그림 3. 일반적인 SLIC의 슈퍼 픽셀  
 Fig. 3. Super-pixels by typical SLIC

Fig 3 (b) is an intensity sliced image, and (c) histogram equalized image. More meaningful image representation gives better recognition results, but it is difficult to demonstrate. Thus, we surveyed the principles that were used to build effective histogram based features and used them to produce mid-level features. Histogram with spatial bins is desirable. Changes of the details are to describe mid-level concepts. The image statistic improves the performance of conventional detectors. Fig 3 (d) shows super-pixels generated by simple SLIC through intensity slicing and histogram equalization. Here, equalized color map is corresponding to an image reshuffle. Here K is used with one thousand. With preprocessing related to gestalt aspects, smaller regions are merged with adjacent regions. Most structures are conserved and the reconstructed segmentation is very similar to the real image qualitatively. However, we may know that many contour details are unnecessary due to the process of over-segmentation. Furthermore, the uncertainties exist in various aspects of image



processing. As the image is always dynamic, fuzzy processing is very desirable.

These uncertainties include additive and non-additive noise in low level image processing, imprecision in the assumptions underlying the algorithms, and ambiguities in interpretation during high level image processing. We used 16 fuzzy edge shapes that represent the possible direction of the edges in the image and then calculate the divergence between the origin image and the 16 fuzzy shapes. When we compare this image to the human-mark segmentation, we find that some pixel details are lost in the process of under-segmentation. With reference to morphological thinning results, smaller regions are merged with adjacent regions. Most structures are conserved and the reconstructed segmentation is very similar to the real image qualitatively.

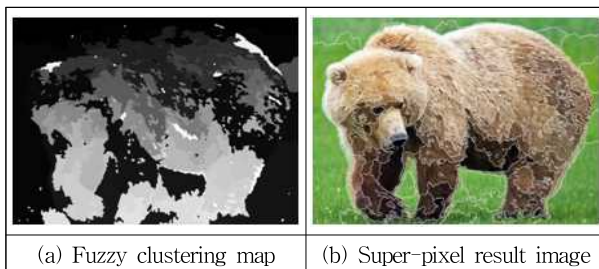


그림 4. 단순 SLIC와 퍼지 클러스터링의 결합  
 Fig. 4. Combining Simple SLIC with Fuzzy Clustering

Fig 4 is a result by the suggested fuzzy clustering of super-pixels. It shows a figure like as reconstruction of the human segmented image from the super-pixels. Fuzzy clustering assigns each super-pixel to a segment with the maximum overlapping area and extracts the super-pixel boundaries. According to this experiment, a super-pixel map generation reduces the complexity of images from many pixels. Also, pairwise contour restrictions allow for much wider range interactions between super-pixels, while only for neighbor pixels on the grid. In particular, fuzzy clustering of super-pixels is perceptually meaningful, for instance, all contours are most likely uniform in color and texture, etc. Under-segmentation error measures the error that

certain technique makes in segmenting an image with respect to human segmented images. This error is computed in terms of picking out the segment result in view of human segmentations. We used the boundary claim by Canny-fuzzy segmentation instead of the standard boundary claim measure<sup>[1]</sup>.

The internal boundaries of each super-pixel are used, and the boundary claim of Canny-fuzzy method is plotted against the number of super-pixels in Fig. 5. We compared our fuzzy approach with the improved SLIC distance measures such as adaptively normalized distance measure (ASLIC) and a geodesic distance measure (GSLIC)<sup>[21]</sup>. Our algorithm showed the lower under-segmentation error in Fig. 5 as well as high boundary recall. Note that the SLIC algorithm has the least under-segmentation error among the approaches computed in terms of the bleeding of the segments output by an algorithm when placed over ground truth segments. Further, cluster validity has been used to evaluate the fitness of partitions produced by clustering algorithms. Our algorithm is faster than density-based spatial clustering for any image size and it outputs the desired number of equal sized compact super pixels.

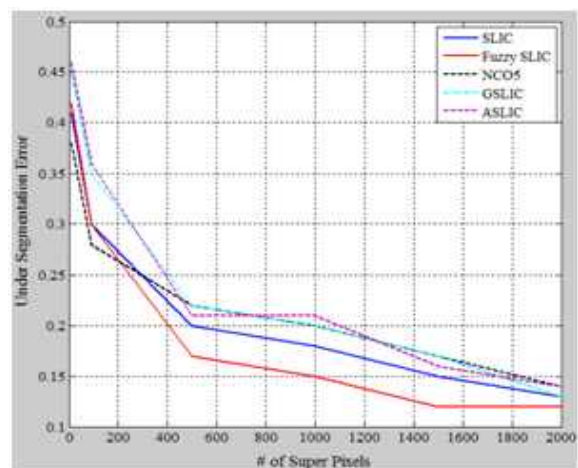


그림 5. 세분화속에서 오류 구성  
 Fig. 5. Under-segmentation Error Plot

## V. Conclusions

The major issue is what to do with the super-pixels.

The applications are numerous, ranging from image compression for perceptual grouping to segmentation in view of gestalt aspects. Currently, super-pixels are mainly used to avoid the complexity for labeling many more pixels. In the graph cuts segmentation, the affinity property can be obtained over super-pixels instead of pixels, and then the result produces a much smaller graph. Further, super-pixels can be considered as a compact image representation. For instance, each color of super-pixel can be approximated as three polynomials. Whereas the mean and the linear approximations are very poor, the quadratic approximation enhances the quality of the original image. Super-pixels belong to a restricted region segmentation that reduces image complexity by pixel grouping and under-segmentation error. Region segmentation algorithms have some problems such as local variation, mean-shift, or watershed. They can lead to under-segmentation error due to the absence of boundary aspects. Algorithms that cut down the compact restrictions such as N-Cuts are very important to reduce the under-segmentation error. SLIC algorithm performs a local clustering of pixels in the CIELAB color space and usual pixel coordinates. It is simple to implement and easily applied, and the only factor is the number of super-pixels. Berkeley benchmark dataset showed that SLIC is more efficient than the other methods. It shows equally or better segmentation results in under-segmentation errors. However, SLIC algorithm does not satisfy some features in view of the gestalt aspects such as non-sum, continuation, closure, perceptual constancy. For instance, human segmented images can be used as prototypes of good segmentations. In this paper, we combined SLIC algorithm with fuzzy clustering. Fuzzy clustering compensates the defects of SLIC algorithm in view of the gestalt aspects such as continuation, closure, and perceptual constancy. Also we applied Canny-fuzzy method to obtain a general human segmented image. As a result, super-pixels showed more natural and visual scenes than the simple SLIC

algorithm. Although we try to do thorough experiment, certain super pixel methods do not consider color information, while the other methods do. This may adversely impact their performance. The future work will try to resolve this problem. Also, it should have almost linear complexity in computational cost and memory usage.

## References

- [1] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua and S. Susstrunk, "SLIC Super pixels Compared to State-of-the-Art Super pixel Methods," *IEEE Transactions on pattern analysis and machine intelligence*, Vol. 34, No. 11. 2012, pp. 2274-2282.
- [2] A. Ayvaci and S. Soatto, "Motion segmentation with occlusions on the super pixel graph," *Proceedings of the Workshop on Dynamical Vision*, Kyoto, Japan, 2009.
- [3] A. V. Batera, "Image Edge Detection Using Ant Colony Optimization", *International Journal of Circuits, Systems and Signal Processing*, Issue 2, Vol. 4. No. 2. 2010, pp.58-67.
- [4] D. Comaniciu and P. Meer, "Mean shift: a robust approach toward feature space analysis." *Pattern Analysis and Machine*, Vol. 24, No. 5, 2002, pp. 603-619.
- [5] P. Felzenszwalb and D. Huttenlocher, "Efficient graph-based image segmentation." *IJCV*, 2004, pp. 167-181.
- [6] B. Fulkerson, A. Vedaldi, S. Soatto, "Class segmentation and object localization with super-pixel neighborhoods." *Computer Vision*, IEEE 12. 2009.
- [7] X. He, R. Zemel, D. Ray, "Learning and incorporating top-down cues in image segmentation." *ECCV*, 2006, pp. 338-351.
- [8] D. Hoiem, A. Efros, M. Hebert, "Automatic Photo Pop-up." *ACM Transactions on Graphics*, Vol. 24, 2005, pp. 577-584.

- [9] T. Kanungo, D. M. Mount, N. S. Netanyahu, C.D. Piatko, R. Silverman and A. Y. Wu, "A local search approximation algorithm for k-means clustering." Symposium on Computational geometry, 2002, pp.10-18.
- [10] M. Jafar, "Ant-based Clustering Algorithms: A Brief Survey", International Journal of Computer Theory and Engineering, Vol. 2, No. 5, 2010, pp. 1793-8201.
- [11] A. Levinstein, C. Sminchisescu and S. Dickinson, "Multi-scale symmetric part detection and grouping." International Journal of Computer Vision September, Vol. 104, No. 2, 2013, pp.117-134.
- [12] A. Levinstein, A. Stere, K. Kutulakos, D. Fleet, S. Dickinson and K. Siddiqi, "Turbo pixels: Fast super pixels using geometric flows." IEEE Transactions on Pattern Analysis & Machine Intelligence 2009, Vol.31, No.12. 2009, pp. 2290 - 2297.
- [13] Y. Li, J. Sun, C. K. Tang, H. Y. Shum, "Lazy snapping." ACM Transactions on Graphics, Vol. 23, 2004, pp. 303-308.
- [14] A. Moore, S. Prince, J. Warrell, U. Mohammed and G. Jones, "Super pixel Lattices." IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2008), 2008, pp. 24-26.
- [15] G. Mori, "Guiding model search using segmentation." Computer Vision, IEEE, 2005, pp.1417-1423.
- [16] X. Ren and J. Malik, "Learning a classification model for segmentation." International Conference on Computer Vision - ICCV, 2003, pp. 10-17.
- [17] J. Shi and J. Malik, "Normalized cuts and image segmentation." PAMI, 2000, pp. 888-905
- [18] A. Vedaldi and S. Soatto, "Quick shift and kernel methods for mode seeking." Computer Vision, ECCV, 2008, pp.705-718.
- [19] L. Vincent and P. Soille, "Watersheds in digital spaces: An efficient algorithm based on immersion simulations." PAMI, 1991, pp.583-598.
- [20] S.H. Lee, H. Jeong and K.I. Moon, "Image Segmentation Based on Fuzzy Clustering Super-Pixel," Proc. Advanced and Applied Convergence Letters, Yanbian, China, June 30-july 4, AAACL 05, 2015, pp. 177-180.
- [21] B. Sridevi and R. Nadarajan, "Fuzzy similarity measure for generalized fuzzy numbers," Int. J. Open Problems Compt. Math., Vol. 2, No. 2, June 2009, pp.240-253.
- [22] S. Gould, J. Rodgers, D. Cohen, G. Elidan, and D. Koller, "Multi-class segmentation with relative location prior," Int'l J. Computer Vision, vol. 80, no. 3, 2008, pp. 300-316.

#### 저자 소개

##### 김 용 길(정회원)



- 1990년 호남대학교 전산통계학과 졸업(이학사)
- 1992년 광주대학교 대학원 컴퓨터학과 졸업(공학석사)
- 2014년 ~ 현재 : 조선이공대학교 컴퓨터보안과 조교수

<주관심분야 : 네트워크보안, 통신시스템, 정보보호>

##### 문 경 일(정회원)



- 1982년 서울대학교 계산통계학과 졸업(이학사)
- 1988년 서울대학교 대학원 계산통계학과 졸업(이학석사)
- 1991년 서울대학교 대학원 계산통계학과 졸업(이학박사)
- 1987년~현재 호남대학교 컴퓨터공학과 교수

<주관심분야 : 지능 시스템, 복잡성 과학>