Information Sciences 248 (2013) 214–238

Contents lists available at SciVerse ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Incorporating local image structure in normalized cut based graph partitioning for grouping of pixels

Debashis Sen [a,*], Niloy Gupta [b], Sankar K. Pal [a]

[a] Center for Soft Computing Research, Indian Statistical Institute, 203 B.T. Road, Kolkata, West Bengal 700 108, India
[b] Department of Computer Engineering, National Institute of Technology Karnataka Surathkal, Mangalore 575 025, India

ARTICLE INFO

Article history:
Received 17 January 2012
Received in revised form 31 December 2012
Accepted 2 June 2013
Available online 22 June 2013

Keywords:
Perceptual grouping
Early human vision
Image pixel grouping
Local image structure
Graph partitioning
Normalized cut

ABSTRACT

Graph partitioning for grouping of image pixels has been explored a lot, with normalized cut based graph partitioning being one of the popular ones. In order to have a credible allegiance to the perceptual grouping taking place in early human vision, we propose and study in this paper the incorporation of local image structure/context in normalized cut based graph partitioning for grouping of image pixels. Similarity and proximity, which have been studied earlier for grouping of image pixels, are only two among many perceptual cues that act during grouping in early human vision. In addition to the said two cues, we study three other such cues, namely, common fate, common region and continuity, and find indications of local image structure utilization during grouping of image pixels. Appropriate incorporation of local image structure/context is achieved by representing it using neighborhood in the form of histogram and fuzzy set. We demonstrate both qualitatively and quantitatively through experimental results that the incorporation of local image structure improves performance of grouping of image pixels.

1. Introduction

Image segmentation and the allied issue of grouping of image pixels have received a lot of attention of researchers over the last half a century. Although both image segmentation and grouping of image pixels refer to appropriate partitioning of images, we, in this paper, would like to have a distinction between them based on the cues driving the processes. As mentioned in [37], we consider image segmentation as a process that partitions images based on lower-level cues such as coherence of brightness, color and texture, and spatial proximity along with higher-level cues such as shape, model and knowledge of objects/regions. On the other hand, we consider grouping of pixels in images as a process that partitions images based only on lower-level cues, that is, image segmentation based on lower-level cues alone. Therefore, we would expect image segmentation to partition images into meaningful regions and grouping of pixels to partition images into homogeneous and spatially compact regions.

We consider grouping of image pixels in the spirit of perceptual grouping in early human vision. Perceptual grouping in early human vision refers to the human visual ability to group image pixels by inferring relations among them from lower-level primitive image features without any knowledge of the image content [22,26]. Proximity, similarity, common fate, continuity, closure, common region and element connectedness are some cues that are known to drive such perceptual grouping [22,32,44]. Lately, evidences have been found which suggest that perceptual grouping occurs at multiple stages in human visual processing and not only at early stages [32]. The said evidences further corroborate the rationale behind our

* Corresponding author.
E-mail addresses: dsen.cscr.isi@gmail.com (D. Sen), niloygupta@gmail.com (N. Gupta), sankar@isical.ac.in (S.K. Pal).

0020-0255/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ins.2013.06.036

distinction of image segmentation and grouping of image pixels. Image segmentation can be seen as the computational analogue to the entire multi-stage perceptual grouping process of human vision, as opposed to grouping of image pixels, which can be seen as the computational analogue to the perceptual grouping occurring only during early human vision.

Graph theoretic formulation of grouping of image pixels has been extensively investigated in literature. Grouping of image pixels based on graph theoretic clustering usually represents an image as a weighted undirected graph where each vertex corresponds to a pixel of the image. A brief survey of some work related to graph theoretic clustering based grouping of image pixels is given in Section 2. As one can infer from the survey, most graph theoretic methods of grouping of image pixels concentrate on separation (cut) criterion, speed and memory, features of a graph representing clusters/groups, better approximation and constraint incorporation. The computation of edge weights, which represent the affinity among image pixels, is also an important factor that effects the outcome of many approaches of grouping of image pixels. The extensive study given in [24,25] on computation of edge weights corresponding to the grouping of image pixels through normalized cut based graph partitioning of [36,37], emphasizes the importance. The appropriateness of computation of edge weights with respect to the cues which drive perceptual grouping of image pixels has seldom been studied.

As mentioned earlier, proximity, similarity, common fate, common region and continuity are a few perceptual cues that drive grouping of image pixels. In grouping of image pixels, proximity and similarity refer to the spatial farness and the coherence in gray/color/texture value, respectively. These two cues have been considered in many cut based graph partitioning approaches for grouping of image pixels (see Section 2). However, to the best knowledge of the authors, cues such as common fate and region of a pixel, and continuity between pixels have not been considered in any approach of grouping of image pixels, graph-theoretic or other. Cues such as common fate, common region and continuity point to the utilization of local image structure or in other words local image context in grouping of image pixels.

In this paper, we propose a technique to compute edge weights of the graph representing an image for graph theoretic clustering based grouping of image pixels such that local image structure is incorporated. We study the proposed computation of the edge weights for grouping of image pixels using normalized cut based graph partitioning. Distributions computed from neighborhoods and fuzzy set theoretic formulation of neighborhoods are used to represent local image structure/context. The said neighborhood based representation is justified with respect to the aforesaid perceptual cues, namely, common fate, common region and continuity. Being a global optimization technique, normalized cut based graph partitioning augers well with the idea of interpreting a global gist of a scene, which is central to perceptual grouping [37] and hence we consider it. Moreover, it considers local interactions between pixels that are represented by affinity values (edge weights) between them. We demonstrate both qualitatively and quantitatively through experimental results that the aforesaid incorporation of local image structure improves performance of grouping of image pixels.

The organization of the paper is as follows. Section 2 briefly reviews some work on graph theoretic clustering for grouping of image pixels. In Section 3, normalized cut based graph partitioning for grouping of image pixels is presented. The study of some important perceptual cues acting in early human vision, the proposed representation of local image structure/context using neighborhoods along with its analysis, and the proposed incorporation of local image structure into grouping of image pixels are given in Section 4. In Section 5, experimental results are shown in order to demonstrate the effectiveness of incorporation of local image structure into grouping of image pixels. A brief summary of the paper is given and conclusions are drawn in Section 6.

2. A brief survey of some work on graph theoretic clustering based grouping of image pixels

Here, we briefly review some work on grouping of image pixels using graph theoretic clustering. A few classical graph theoretic techniques for data clustering are given in [17,40,47], which are based on minimal spanning tree, the concept of limited neighborhood sets and cluster representation by directed trees, respectively. A recently developed graph theoretic data clustering technique is given in [19] that avoids approximated iterative approaches by presenting an analytical solution. The said clustering methods can be used to perform grouping of image pixels.

Grouping of image pixels through clustering based on cut values corresponding to a weighted graph representing the underlying image was introduced in [45], where the minimization of the maximum of cut values between subgraphs representing image regions is considered. The method used similarity between image pixels, which is calculated using a local derivative operator, to get the weights of the edges in the graph. In [6], similarity based weights of edges and faces in a planar graph representing the underlying image are considered for grouping of image pixels. The said approach is claimed to be an improvement over that of [45], as both interior region and boundary information are considered instead of the latter alone. As mentioned in [37], the authors of [45] noticed that the minimum cut criterion in their approach favors results having small and isolated image regions. This drawback is addressed in [36,37], where a normalization of cut value is carried out to provide a measure referred as normalized cut and it is used instead of cut. Grouping of image pixels is carried out by performing recursive two-way or simultaneous multi-way partitioning, and by considering similarity and proximity among image pixels for computation of the edge weights. In [43], it is shown how the graph partitioning approach developed in [36] can be thought in the light of a general spectral partitioning framework.

In the approach of [37], the minimization of normalized cut is NP-complete and hence an approximate solution can only be found. In [39], approximate multi-way graph partitioning for grouping of image pixels is performed, where the edges in the graph are iteratively re-weighted in a manner that it eventually disconnects into a prescribed number of components. A few eigenvectors with smaller eigenvalues obtained at each iteration are used to determine the new edge weights. The approximation in [39] is referred to as spectral rounding, which is shown to have some empirical advantages among those concerning the normalized cut criterion. In [5], energy minimization based graph partitioning is considered and efficient approximation algorithms are provided.

Several other cut based graph partitioning approaches for grouping of image pixels have been reported in literature. In [42], minimization of a cost function referred to as ratio cut is performed for partitioning, which guarantees that the regions produced are connected and allows boundaries of regions to align with image edges. In [23], images are represented in tensor space and user initialized regions are used to compute edge weights for graph partitioning based grouping of image pixels in order to achieve superior performance. A similar interactive algorithm is presented in [41], where an explicit connectivity prior is considered through user input after graph partitioning. A computationally efficient algorithm using graph partitioning based thresholding for grouping of image pixels to perform object extraction is presented in [38], where gray levels are considered as graph vertices instead of pixels. In [15], grouping of pixels in color textured images is considered based on a new texture descriptor and cut based graph partitioning. In [18], a new energy function without regularizing parameter is considered in energy minimization based graph partitioning for grouping of image pixels. A reformulation of the normalized cut based graph partitioning of [37] is presented in [8] such that it can handle linear equality constraints exactly, allowing the incorporation of priori information in grouping of image pixels. In [35], a kernel function is used in graph partitioning based grouping of image pixels so that the image data can be mapped to higher dimensional feature in order to have better separability.

Attempts have been made to improve speed and memory requirements of cut based graph partitioning algorithms, which are computationally intensive, for grouping of image pixels. Isoperimetric graph partitioning based grouping of image pixels is presented in [11], which is carried out by solving a linear system. The said approach is claimed to provide improvement in speed and stability compared to eigenvector based approach of graph partitioning. In [7], the mathematical equivalence between weighted kernel k-means objective and weighted graph clustering objective is exploited to eliminate the requirement of eigenvector computation in cut based graph partitioning for grouping of image pixels achieving higher speed and lower memory usage.

Various other graph theoretic approaches of grouping of image pixels are present in literature. A nearly linear-time graph based algorithm for grouping of image pixels is presented in [10], which is adapted based on similarities between pixels within and between regions enabling the capture of perceptually important non-local properties of images. In [33], the analogy between a dominant set of vertices, which generalizes the notion of maximal clique in the context of edge weighted graphs, and the concept of a cluster is exploited to perform grouping of image pixels. In [13], an algorithm is developed that searches for embedded optimal cycles in a graph representing the underlying image to perform grouping. A ratio of a measure of flow of quantity such as intensity (intensity gradient) into or out of a region to a measure of length of the region boundary is considered for the said purpose.

3. Normalized cut based graph partitioning for grouping of image pixels

Normalized cut based graph partitioning for grouping of image pixels has been developed in [37]. In the approach, an image is modeled as a weighted undirected graph $G = (V, E)$, where the vertices of the graph are the pixels of the image. Each pair of vertices are connected by edges, which are associated with weights ω. A weight ω_{ij} gives the affinity between ith and jth vertices. Grouping of image pixels is done by partitioning the set of vertices into disjoint sets V_k, $k = 1, \ldots, N$ such that affinity among the vertices within every V_k is high and affinity across the vertices in different V_ks is low.

In [37], both recursive two-way partitioning and simultaneous N-way partitioning have been considered to divide the graph into multiple disjoint sets. We consider recursive two-way partitioning, which does not require predefining of the number of groups. Bi-partitioning/two-way partitioning of a set of vertices S is carried out by removing the edges connecting two parts A, B of it, where $A \cup B = S$ and $A \cap B = \emptyset$. Optimal partitioning is obtained by choosing A and B such that a measure called normalized cut is minimized. The normalized cut value is defined as follows:

$$Ncut(A,B) = \frac{\sum_{x\in A, y\in B}\omega_{xy}}{\sum_{x\in A, z\in S}\omega_{xz}} + \frac{\sum_{x\in A, y\in B}\omega_{xy}}{\sum_{y\in B, z\in S}\omega_{yz}} \quad (1)$$

From (1), it is evident that normalized cut is an overall affinity measure between all the vertices in A and B with respect to the overall affinity of the vertices in A and B with all the vertices in S. As shown in [37], minimization of normalized cut is same as minimization of association between the groups, which is equivalent to maximization of association within the groups.

As mentioned in [37], minimization of normalized cut is NP-complete. However, an approximate and efficient solution is available through spectral graph partitioning. The requirement is to solve the following generalized eigenvalue problem:

$$(\mathbf{D} - \mathbf{W})\mathbf{x} = \lambda \mathbf{D}\mathbf{x} \quad (2)$$

where

$$\mathbf{D}(i,j) = \begin{cases} \sum_l \omega_{il} & i = j \\ 0 & i \neq j \end{cases}$$
$$\mathbf{W}(i,j) = \omega_{ij}$$

We use the ‘eigs’ function of MATLAB® to solve the generalized eigenvalue problem given above. The eigenvector corresponding to the second smallest eigenvalue of the generalized eigenvalue system in (2) is the solution to the problem of minimization of normalized cut. A splitting point is required to be chosen in the said eigenvector to partition the graph into two parts. As suggested in [37], we consider the splitting point such that the partitioning results in the best (smallest) normalized cut value (*Ncut*(*A*, *B*)). We also consider the partition stability criterion given in [37], where it is checked whether the eigenvector has a shape of a continuous function and partitioning is not performed if it is found so. An eigenvector taking a shape of a continuous function is an indication of the fact that there is no sure way of splitting the eigenvector and hence a forced division might lead to an inappropriate result.

Therefore, the algorithm of [37] for grouping of image pixels is as follows:

(1) Given an image, model it as a weighted undirected graph $G = (V, E)$, where the weights (ω) of the edges (E) give the affinity between image pixels represented by the vertices (V).
(2) Solve $(\mathbf{D} - \mathbf{W})\mathbf{x} = \lambda\mathbf{D}\mathbf{x}$ to get the eigenvector corresponding to the second smallest eigenvalue.
(3) Find a splitting point in the eigenvector corresponding to the second smallest eigenvalue that minimizes the normalized cut value in order to bipartition the graph G.
(4) Perform the bi-partitioning of the graph G only if the associated partition stability criterion is met and the associated normalized cut value is below a pre-defined threshold.
(5) Recursively bipartition the parts obtained after previous bi-partitioning in a manner similar to the previous four steps until no further bi-partition- ing can take place.

Now, an important part of the aforesaid algorithm for grouping of image pixels is the determination of the edge weights (ω_{ij}) which give the affinities between image pixels. In [37], the following weight has been suggested:

$$\omega_{ij} = e^{\frac{-\|\mathbf{F}(i)-\mathbf{F}(j)\|_2^2}{\sigma_g^2}} \times \begin{cases} e^{\frac{-\|\mathbf{X}(i)-\mathbf{X}(j)\|_2^2}{\sigma_s^2}} & \|\mathbf{X}(i) - \mathbf{X}(j)\|_2 < r \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

where **F** represents feature vector, which will be only a gray value in the case of grayscale images and **X** represents spatial location. The quantities σ_g and σ_s determine the extent of existence of affinity between image pixels based on their feature values and spatial locations, respectively, and r represents a threshold in terms of spatial proximity beyond which the underlying image pixels do not have any affinity.

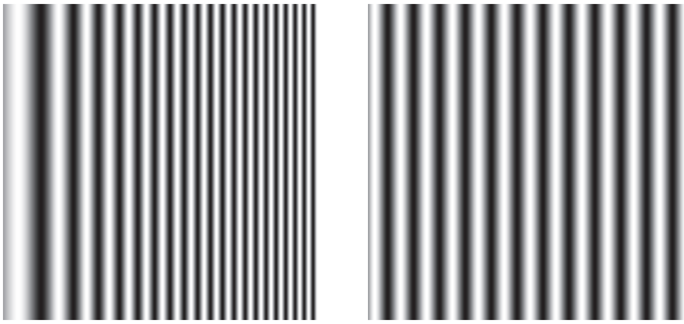
4. Perceptual cues, representation of local image structure and its incorporation into grouping of image pixels

As described in Section 1, proximity, similarity, common fate, common region and continuity are a few perceptual cues that drive grouping of image pixels in early human vision. Proximity and similarity, which refer to spatial farness and coherence in feature value, are the two cues that are inherently considered in the algorithm for grouping of image pixels explained in Section 3. The other three cues, namely, common fate, common region and continuity are also very important cues which might play a crucial role in an algorithm for grouping of image pixels, if they are suitably considered. Let us now elaborately consider the three aforementioned cues from the point of view of grouping.

4.1. The three perceptual cues from the point of view of grouping in images

4.1.1. Common fate

Common fate is a perceptual cue by which there is a tendency to group together the elements that have undergone similar phenomenon [44]. In images, common fate of pixels would refer to the coherence in the spatial phenomenon/condition to which they are subjected to. For an illustration, consider the images shown in Fig. 1. Fig. 1a shows an image containing sinusoidal gratings whose frequency increases from left to right. On the other hand, Fig. 1b shows an image containing sinusoidal gratings whose frequency does not change. Now, consider a pixel each from the left and right halves of both the images. It is clear that in the image of Fig. 1a, the two pixels from the two halves are subjected to (or a part of) different spatial phenomena. Whereas, in the image of Fig. 1b, the two pixels from the two halves are subjected to the same spatial phenomenon. Therefore, in accordance to only the common fate cue (ignoring other cues), the two pixels in the image of Fig. 1b would have more affinity than the two pixels in the image of Fig. 1a. It is evident from the aforesaid illustration that the common fate cue indicates towards the utilization of local image structure/context in grouping of image pixels, such that the local structure considered around a pixel involves some impression of the spatial phenomenon to which that pixel is subjected to.



(a) Sinusoidal gratings with frequency increasing from left to right

(b) Sinusoidal gratings with no change in frequency from left to right

Fig. 1. Illustration of the common fate cue in images.

(a) Two non-intersecting bounded areas

(b) Two intersecting bounded areas

Fig. 2. Illustration of the common region cue in images.

4.1.2. Common region

Common region is a perceptual cue by which there is a tendency to group together the elements that lie within the same bounded area [32]. In images, common region of pixels would refer to the commonness between the bounded areas within which they are contained. For an illustration, consider the images shown in Fig. 2. Fig. 2a shows an image containing two non-intersecting bounded areas. On the other hand, Fig. 2b shows an image containing two intersecting bounded areas. Now consider a pixel each from the two bounded areas in both the images. It is clear that the commonness between the two bounded areas of the two pixels from the image of Fig. 2a is less compared to the commonness between the two bounded areas of the two pixels from the image of Fig. 2b. Therefore, in accordance to only the common region cue (ignoring other cues), the two pixels in the image of Fig. 2b would have more affinity than the two pixels in the image of Fig. 2a. It is evident from the aforesaid illustration that the common region cue indicates towards the utilization of local image structure/context in grouping of image pixels, such that the local structure considered around a pixel involves some impression of the bounded area to which that pixel is contained.

4.1.3. Continuity

Continuity is a perceptual cue by which there is a tendency to group together the elements that naturally seem to be successive parts of a whole following one another [44]. In images, continuity between pixels would refer to the homogeneity of the paths connecting them. For an illustration, consider the images shown in Fig. 3. Fig. 3a shows an image containing a white patch on a black surface. On the other hand, Fig. 3b shows an image containing a completely black surface with no patches. Now consider one pixel each from the black areas on the either side of the white patch in the image of Fig. 3a. Consider two pixels in the image of Fig. 3b such that they would have lied on different sides of the white patch, if they were taken from the image of Fig. 3a. It is clear that the homogeneity of paths between the two pixels from the image of Fig. 3a is less compared to the homogeneity of paths between the two pixels from the image of Fig. 3b. Therefore, in accordance to only the continuity cue (ignoring other cues), the two pixels in the image of Fig. 3b would have more affinity than the two pixels in the image of Fig. 3a. It is evident from the aforesaid illustration that the continuity cue indicates towards

(a) A white patch on a black surface (b) A completely black surface with no patches

Fig. 3. Illustration of the continuity cue in images.

the utilization of local image structure/context in grouping of image pixels, such that the local structure considered around a pixel involves some impression of the paths between that pixel and others.

Research activities, which raise question on the use of only spatial proximity and feature similarity in grouping algorithms, have been reported in literature [2,9,12,34]. In [34], the authors discuss the proximity cue considering proximity between phenomena represented by elements rather than proximity between elements themselves. The higher involvement of surfaces compared to features in tasks such as visual search is pointed out in [12]. In [2], a model of perceptual grouping based on spatial correlation computed in a limited time-dependent range is putforth. In [9], the authors mention about the presence of articulated substructures and complex spatial relation among elements of a grouping unit.

Findings in [2,9,12,34] such as involvement of spatial relation and correlation, surfaces and phenomenal proximity in visual/perceptual grouping point towards the utilization of local image structure/context. Such research outcomes provide motivation to explore the use of cues such as common fate, common region and continuity along with spatial proximity and feature similarity in grouping of image pixels, because as mentioned in the elaboration of the common fate, common region and continuity cues, they also indicate towards utilization of local image structure/context in grouping of image pixels.

Local image structure/context corresponding to an image pixel can be represented by considering a neighborhood around the pixel. For a given purpose, a neighborhood can be used in different forms. For example, [31,48] consider a neighborhood in the form of a binary and m-ary number for texture analysis, respectively. Considering the distribution of values within a neighborhood or a neighborhood as a whole in the form of an array might be suitable for grouping of image pixels. Note that, in such cases a bunch of multiple values would represent a neighborhood, unlike the case of grouping considering feature values calculated from neighborhoods.

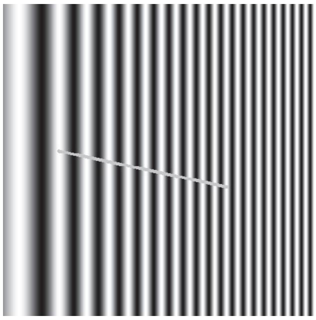
Some work on grouping of image pixels considering neighborhoods of pixels has been reported in literature [21,28,30]. In [21], local histograms computed from multiple filtered outputs of an image are considered. Histograms of image regions obtained from multiple initial estimates are used in [28] to perform a fusion to get the final result. In [30], homogeneity among local histograms computed around pixels is checked to determine those lying within the same region. It is evident that the aforesaid uses of neighborhoods of pixels consider a neighborhood in the form of local histogram/distribution. We shall consider a neighborhood both in the form of local histogram and as a whole in the form of an array.

We use neighborhoods to represent local image structure and alikeness between local image structures around pixels is related to the affinity among pixels due to the common fate, common region and continuity cues. Therefore, irrespective of the form in which we consider a neighborhood of a pixel, we are required to measure the alikeness between two given neighborhoods. From (3) we see that in the algorithm for grouping of image pixels described in Section 3, only proximity and similarity cues are considered and the affinity among pixels due to these cues are measured in terms of Euclidean distance. In order to incorporate affinity among pixels due to the common fate, common region and continuity cues, when we consider neighborhoods in the form of local histograms, we shall use measure of alikeness between probability distributions. It is well known that a probability distribution can be computed from a histogram through a normalization operation. When we consider neighborhoods in the form of an array, we shall consider that all the elements in the array form a universe and define a fuzzy set in that universe. The affinity among pixels due to the common fate, common region and continuity cues shall be measured in terms of alikeness between fuzzy sets (like the one presented in [1]). Let us now elaborately consider alikeness measures between neighborhoods and appropriateness of the aforementioned representation of local image structure.

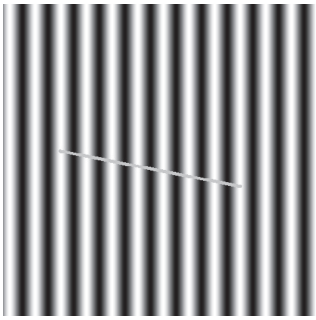
4.2. The measures of alikeness between neighborhoods and appropriateness of the local image structure representation

As mentioned earlier, we represent local image structure using neighborhoods in various forms around pixels. Here, we shall present three measures of alikeness between neighborhoods and justify the representation of local image structure by neighborhoods using the measures of alikeness. The justification would be based on the ability of the measures of alikeness to represent the affinity among pixels due to the common fate, common region and continuity cues.

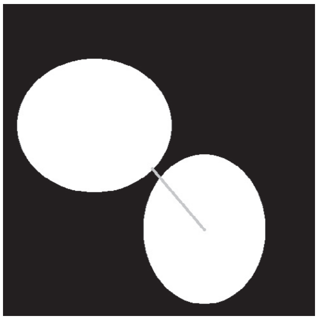
The three measures of alikeness considered are based on Bhattacharyya coefficient [4,14] when neighborhoods are considered in the form of local histograms, and based on fuzzy divergence [3] and fuzzy correlation [29] when neighborhoods are considered in the form of a fuzzy set. Consider the images shown in Fig. 4. The images are those given in Figs. 1–3. In each image, a gray line joins two pixels and we wish to measure the affinity between them. In the images of Fig. 4a and b, the two pixels are respectively subjected to different and same spatial phenomenon. Therefore, the measures of alikeness representing affinity due to only the common fate cue should suggest that the two pixels in the image of Fig. 4b have more affinity than the two pixels in the image of Fig. 4a. In the images of Fig. 4c and d, the commonness between the bounded areas within which the two pixels belong is respectively lower and higher. Therefore, the measures of alikeness representing affinity due



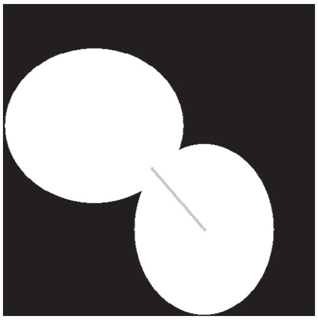
(a) Two pixels subjected to different spatial phenomenon



(b) Two pixels subjected to same spatial phenomenon



(c) Two pixels with lower commonness between their bounded areas



(d) Two pixels with higher commonness between their bounded areas

(e) Two pixels with less homogeneous paths connecting them

(f) Two pixels with more homogeneous paths connecting them

Fig. 4. Demonstration of affinity among pixels due to common fate, common region and continuity cues in terms of measures of alikeness between neighborhoods.

Table 1
Ability of feature similarity and spatial proximity to represent the affinity among pixels due to the common fate, common region and continuity cues.

	Feature similarity	Spatial proximity	Cue
Fig. 4a	0	281.258	Common fate
Fig. 4b	0	298.881	
Fig. 4c	0	131.894	Common region
Fig. 4d	0	131.894	
Fig. 4e	0	54	Continuity
Fig. 4f	0	54	

to only the common region cue should suggest that the two pixels in the image of Fig. 4d have more affinity than the two pixels in the image of Fig. 4c. In the images of Fig. 4e and f, the paths connecting the two pixels are respectively of lower and higher homogeneity. Therefore, the measures of alikeness representing affinity due to only the continuity cue should suggest that the two pixels in the image of Fig. 4f have more affinity than the two pixels in the image of Fig. 4e.

Before we elaborate the said measures of alikeness and study their ability to represent the affinity among pixels due to the common fate, common region and continuity cues, let us discuss the ability of spatial proximity and feature similarity, which have been popularly used for grouping of image pixels, to do the same. Table 1 gives the similarity (gray value similarity) and proximity (spatial proximity) values between the chosen two pixels in each of the images in Fig. 4 measured in terms of Euclidean distance. Therefore, a higher (lower) value would represent less (more) affinity.

As can be seen from the table, the feature similarity values do not suggest any difference between the affinity among the two pixels in Fig. 4a and the affinity among the two pixels in Fig. 4b. This observation is contrary to the affinity created by the common fate cue. Similar observations can be made from the table about affinity in terms of similarity values among the pixels in Fig. 4c and d, and Fig. 4e and f, which are contrary to the affinity created by the common region and continuity cues, respectively. It is also evident from the table that affinity in terms of spatial proximity values among the pixels in Fig. 4a and b, Fig. 4c and d, and Fig. 4e and f do not correspond to the affinity created by the common fate, common region and continuity cues. Therefore, the said gray value similarity and spatial proximity measures, which are respectively intended to represent affinity due to similarity and proximity cues, lack the ability to represent the affinity among pixels due to the common fate, common region and continuity cues.

Let us now consider the elaboration of the said measures of alikeness between neighborhoods, study their ability to represent the affinity among pixels due to the common fate, common region and continuity cues and hence justify neighborhood based representation of local image structure.

4.2.1. Bhattacharyya coefficient based

Here we present a measure of alikeness between neighborhoods based on Bhattacharyya coefficient. Let N be a neighborhood around a pixel in a L-level grayscale image. Let n be a pixel in the neighborhood N ($n \in N$) and $g(n)$ be the gray value at n, where $g \in G = \{0, \ldots, L-1\}$. Let H be the local gray-level histogram $H(l)$, $l = 0, \ldots, L-1$ calculated from the neighborhood N by performing $H(l) = H(l) + 1$, when $l = g(n)\ \forall n \in N$ (with all $H(l)$s initialized as zero). Therefore, the probabilities of occurrence of gray values in the neighborhood N is given as

$$p(l) = \frac{H(l)}{\sum_{k=0}^{L-1} H(k)}, \quad l = 0, \ldots, L-1 \tag{4}$$

Note that, p is referred to as the probability density of gray value occurrence in the neighborhood N. Now consider two neighborhoods N_1 and N_2 in a grayscale image, and let p_1 and p_2 be the corresponding probability densities. The Bhattacharyya coefficient [4,14] for the two densities measuring the discrepancy between them is given as

$$B = \sum_{l=0}^{L-1} \sqrt{p_1(l)p_2(l)} \tag{5}$$

A lower (higher) value of B represents more (less) discrepancy. Hellinger distance, which is a type of f-divergence and obeys the triangular inequality, can then be computed from the Bhattacharyya coefficient as follows

$$D = \sqrt{1 - B} \tag{6}$$

We then compute the measure of alikeness between the neighborhoods N_1 and N_2 from D as

$$\Omega = 1 - D \tag{7}$$

A higher (lower) value of Ω represents more (less) alikeness and Ω lies in the range [0 1]. Note that during the computation of the Ω in (7) neighborhoods are considered in the form of local histograms.

Table 2
Ability of Bhattacharyya coefficient based neighborhood alikeness to represent the affinity among pixels due to the common fate, common region and continuity cues.

	Neighborhood alikeness	Cue
Fig. 4a	0.0665	Common fate
Fig. 4b	0.0665	
Fig. 4c	0.5445	Common region
Fig. 4d	1	
Fig. 4e	1	Continuity
Fig. 4f	1	

Table 2 gives the measured Bhattacharyya coefficient based neighborhood alikeness values between the chosen two pixels in each of the images in Fig. 4. A higher (lower) value would represent more (less) affinity. The neighborhood around a pixel is chosen by taking a 11×11 window of pixels around it. As can be seen from the table, the Bhattacharyya coefficient based neighborhood alikeness values do not suggest any difference between the affinity among the two pixels in Fig. 4a and the affinity among the two pixels in Fig. 4b. Similarly they do not suggest any difference between the affinity among the two pixels in Fig. 4e and the affinity among the two pixels in Fig. 4f. These observations are respectively contrary to the affinity created by the common fate and continuity cues. However, the Bhattacharyya coefficient based neighborhood alikeness is in agreement with the common region cue, as it suggests that the affinity among the two pixels in Fig. 4c is less than the affinity among the two pixels in Fig. 4d.

The failure of Bhattacharyya coefficient based neighborhood alikeness values to represent the affinity created by the common fate and continuity cues is due to the fact that neighborhoods in the form of local histograms do not posses the information about the location of pixels. Note that the grouping methods of [21,28,30] discussed earlier use neighborhoods in the form of local histograms, and hence they would also suffer from the aforesaid drawback.

As mentioned earlier in Section 4.1, the neighborhood representing the local structure considered around a pixel should contain some impression of the circumstances that cause affinity due to the three mentioned cues. The 11×11 window of pixels taken as a neighborhood around a pixel satisfies the aforesaid condition in case of the images in Fig. 4.

With the aforesaid condition satisfied, it is found from Table 2 that Bhattacharyya coefficient based neighborhood alikeness values lack the ability to represent the affinity among pixels due to the common fate and continuity cues. However, it has the ability to represent the affinity among pixels due to the common region cue.

4.2.2. Fuzzy divergence based

Here we present a measure of alikeness between neighborhoods based on fuzzy divergence. Let N be a neighborhood around a pixel in a L-level grayscale image. Let n be a pixel in the neighborhood N ($n \in N$) and $g(n)$ be the gray value at n, where $g \in G = \{0, \ldots, L-1\}$. Consider N as a universal set of pixels. A fuzzy set [16] is then defined in N as follows

$$F = \{(n, \mu(n))\}, \quad n \in N \tag{8}$$

where μ is the corresponding membership function, which given as

$$\mu(n) = \frac{g(n)}{L-1} \tag{9}$$

The fuzzy set F is defined such that it answers the question - how much bright is a pixel in N? A pixel is considered to be more (less) bright if its gray value is nearer to (farther from) $L-1$. The membership value $\mu(n)$ (ranging between 0 and 1) gives the extent to which the pixel n is bright.

Now consider two neighborhoods N_1 and N_2 in a grayscale image, and let F_1 and F_2 be the two fuzzy sets defined respectively in them with the corresponding membership functions μ_1 and μ_2. Let the number of pixels in $N1$ be equal to that in N_2 and hence $|N_1| = |N_2| = |N|$. The fuzzy divergence [3] between the two fuzzy sets measuring the discrepancy between them is given as

$$D = \frac{1}{|N|}\sum_{k=1}^{|N|}(D_k(F1, F2) + D_k(F2, F1)) \tag{10}$$

where

$$D_k(F1, F2) = \mu_1(k)\log\left(\frac{\mu_1(k)}{\mu_2(k)}\right) + (1-\mu_1(k))\log\left(\frac{1-\mu_1(k)}{1-\mu_2(k)}\right) \tag{11}$$

$$D_k(F2, F1) = \mu_2(k)\log\left(\frac{\mu_2(k)}{\mu_1(k)}\right) + (1-\mu_2(k))\log\left(\frac{1-\mu_2(k)}{1-\mu_1(k)}\right) \tag{12}$$

Table 3
Ability of fuzzy divergence based neighborhood alikeness to represent the affinity among pixels due to the common fate, common region and continuity cues.

	Neighborhood alikeness	Cue
Fig. 4a	0	Common fate
Fig. 4b	1	
Fig. 4c	0	Common region
Fig. 4d	1	
Fig. 4e	0	Continuity
Fig. 4f	1	

Note that $\sum_{k=1}^{|N|}$ in (10) essentially means that the summation considers all $n \in N_1$ or N_2. A lower (higher) value of D represents less (more) discrepancy. We then compute the measure of alikeness between the neighborhoods N_1 and N_2 from D as

$$\Omega = \exp(-D) \tag{13}$$

A higher (lower) value of Ω represents more (less) alikeness and Ω lies in the range [0 1]. Note that during the computation of the Ω in (13) neighborhoods are considered in the form of fuzzy sets.

Table 3 gives the measured fuzzy divergence based neighborhood alikeness values between the chosen two pixels in each of the images in Fig. 4. A higher (lower) value would represent more (less) affinity. Similar to Section 4.2.1, the neighborhood around a pixel is chosen by taking a 11×11 window of pixels around it. As can be seen from the table, the fuzzy divergence based neighborhood alikeness values suggest that the affinities among the two pixels in Fig. 4 are respectively less than the affinities among the two pixels in Fig. 4. These observations are respectively in agreement to the affinity created by the common fate, common region and continuity cues.

Fuzzy divergence based neighborhood alikeness values are able to represent the affinity created by the common fate and continuity cues unlike Bhattacharyya coefficient based neighborhood alikeness values due to the fact that neighborhoods in the form of fuzzy sets posses the information about the location of pixels.

As mentioned in Section 4.2.1, the 11×11 window of pixels taken as a neighborhood around a pixel captures an impression of the circumstances that cause affinity due to the three mentioned cues in case of the images in Fig. 4. With the aforesaid statement being true, it is found from Table 3 that fuzzy divergence based neighborhood alikeness values posses the ability to represent the affinity among pixels due to the common fate, common region and continuity cues.

4.2.3. Fuzzy correlation based

Here we present a measure of alikeness between neighborhoods based on fuzzy correlation. Consider two neighborhoods N_1 and N_2 in a grayscale image, and let F_1 and F_2 be two fuzzy sets defined respectively in them with the corresponding membership functions μ_1 and μ_2 as explained in Section 4.2.2. Let the number of pixels in $N1$ be equal to that in N_2 and hence $|N_1| = |N_2| = |N|$. The fuzzy correlation [29] between the two fuzzy sets measuring the discrepancy between them is given as

$$C = \begin{cases} 1 - \frac{4\sum_{k=1}^{|N|}(\mu_1(k)-\mu_2(k))^2}{X_1+X_2} & X_1 + X_2 \neq 0 \\ 1 & X_1 + X_2 = 0 \end{cases} \tag{14}$$

where

$$X_1 = \sum_{k=1}^{|N|}(2\mu_1(k) - 1)^2 \tag{15}$$

$$X_2 = \sum_{k=1}^{|N|}(2\mu_2(k) - 1)^2 \tag{16}$$

As in Section 4.2.2, $\sum_{k=1}^{|N|}$ in (14) essentially means that the summation considers all $n \in N_1$ or N_2. A lower (higher) value of C represents more (less) discrepancy. We then compute the measure of alikeness between the neighborhoods N_1 and N_2 from C as

$$\Omega = \frac{C + 1}{2} \tag{17}$$

A higher (lower) value of Ω represents more (less) alikeness and Ω lies in the range [0 1]. Note that, similar to Section 4.2.2, during the computation of the Ω in (17) neighborhoods are considered in the form of fuzzy sets.

Table 4 gives the measured fuzzy correlation based neighborhood alikeness values between the chosen two pixels in each of the images in Fig. 4. A higher (lower) value would represent more (less) affinity. Similar to Sections 4.2.1 and 4.2.2, the neighborhood around a pixel is chosen by taking a 11×11 window of pixels around it. As can be seen from the table, the fuzzy correlation based neighborhood alikeness values suggest that the affinities among the two pixels in Fig. 4 are respec-

Table 4
Ability of fuzzy correlation based neighborhood alikeness to represent the affinity among pixels due to the common fate, common region and continuity cues.

	Neighborhood alikeness	Cue
Fig. 4a	0.8182	Common fate
Fig. 4b	1	
Fig. 4c	0.6281	Common region
Fig. 4d	1	
Fig. 4e	0.4545	Continuity
Fig. 4f	1	

tively less than the affinities among the two pixels in Fig. 4. These observations are respectively in agreement to the affinity created by the common fate, common region and continuity cues.

Similar to fuzzy divergence based neighborhood alikeness values, fuzzy correlation based neighborhood alikeness values are able to represent the affinity created by the common fate and continuity cues unlike Bhattacharyya coefficient based neighborhood alikeness values due to the fact that neighborhoods in the form of fuzzy sets posses the information about the location of pixels. Note that fuzzy divergence based neighborhood alikeness values are more severe than fuzzy correlation based ones in discriminating the affinities among the two pixels in Fig. 4 respectively from the affinities among the two pixels in Fig. 4.

As mentioned in Sections 4.2.1 and 4.2.2, the 11×11 window of pixels taken as a neighborhood around a pixel captures an impression of the circumstances that cause affinity due to the three mentioned cues in case of the images in Fig. 4. With the aforesaid statement being true, it is found from Table 4 that fuzzy correlation based neighborhood alikeness values posses the ability to represent the affinity among pixels due to the common fate, common region and continuity cues.

Once the neighborhood alikenesses (Ω_{ij}) between image pixels have been computed, we redefine the edge weights $\left(\omega_{ij} \to \omega'_{ij}\right)$ of Section 3 that are used for normalized cut based graph partitioning for grouping of image pixels by modifying the formulation given in (3) as

$$\omega'_{ij} = (1 + \Omega_{ij}) \times \omega_{ij} \tag{18}$$

$$= (1 + \Omega_{ij}) \times e^{\frac{-\|\mathbf{F}(i) - \mathbf{F}(j)\|_2^2}{\sigma_g^2}} \times \begin{cases} e^{\frac{-\|\mathbf{X}(i) - \mathbf{X}(j)\|_2^2}{\sigma_s^2}} & \|\mathbf{X}(i) - \mathbf{X}(j)\|_2 < r \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

The formulation of edge weights ω'_{ij} is done such that:

(1) The contribution of similarity and proximity cues is not affected by the contribution of common fate, common region and continuity cues. The contribution due to similarity and proximity cues is given by $1 \times \omega_{ij}$.
(2) The contribution of common fate, common region and continuity cues is modulated by the contribution of similarity and proximity cues. This is because assigning high affinity due to common fate, common region and continuity cues between pixels which are not nearby or have incoherent feature would not be useful for grouping of image pixels as the assignment does not hold much meaning. The term containing the contribution due to common fate, common region and continuity cues is $\Omega_{ij} \times \omega_{ij}$.

The edge weights $\left(\omega'_{ij}\right)$ in (19) give the affinities between image pixels due to similarity, proximity, common fate, common region and continuity cues as opposed to the edge weights (ω_{ij}) in (3) which give the affinities due to similarity and proximity cues alone. In our approach of grouping of image pixels, we use the algorithm of [37] that employs the proposed edge weights $\left(\omega'_{ij}\right)$ in (19).

5. Experimental results

In this section, we present experimental results in order to demonstrate the effectiveness of incorporation of local image structure in normalized cut based graph partitioning for grouping of image pixels. In order to do so, we consider both qualitative and quantitative evaluation of performance of grouping of image pixels with and without the incorporation of local image structure. We also study the time consumed by normalized cut based graph partitioning for grouping of image pixels with and without the incorporation of local image structure.

Recall that we have considered local image structure using neighborhoods around pixels and we have studied three different measures of alikeness between neighborhoods for use in the normalized cut based graph partitioning for grouping of image pixels. The incorporation of local image structure is achieved by modifying the edge weights used in the normalized cut based grouping of image pixels such that neighborhood alikeness is involved. In the experimental results given in this section, the use of all the three measures of alikeness between neighborhoods are considered while studying the case when local image structure is incorporated in grouping of image pixels.

As evident from Section 3, the normalized cut based graph partitioning for grouping of image pixels without the incorporation of local image structure depends on six pre-defined parameters. They are as follows:

P1. $\sigma_g \rightarrow$ a parameter that defines the extent of decrease (increase) in similarity between feature values with increase (decrease) in the Euclidean distance between them.
P2. $\sigma_s \rightarrow$ a parameter that defines the extent of decrease (increase) in proximity between pixels with increase (decrease) in the Euclidean distance (spatial) between them.
P3. $r \rightarrow$ a parameter that defines the spatial extent in terms of Euclidean distance beyond which pixels have no affinity.
P4. A stability criterion threshold value → if a characteristic value calculated from an eigenvector based on which a bi-partitioning has been performed is greater than the said threshold value then the resulting partitions would be unstable.
P5. Minimum number of pixels in a valid group → if the number of pixels in a group obtained after a bi-partitioning operation is less the said minimum number, it is declared invalid.
P6. A normalized cut value based stopping criterion threshold value → if the associated normalized cut value corresponding to a bi-partitioning is greater than the said threshold value, the bi-partitioning is considered invalid.

When the normalized cut based graph partitioning for grouping of image pixels is considered with the incorporation of local image structure, another parameter is added to the aforesaid list, which is:

P7. Size of a neighborhood → The size of the square window of pixels taken around a pixel in order to define the neighborhood around it.

Throughout this section, we fix the values of P1 and P4, and compute P2 and P5 in the same manner. We take the value of P1 as 5 feature (gray) values and the value of P4 is taken as 0.06 as suggested in [37]. P2 is computed as $\sqrt{R^2 + C^2}/20$ pixel width and P5 is computed as $\lfloor\sqrt{R^2 + C^2}/50\rceil$ pixels, where R and C respectively are the number of rows and columns in the image under consideration. Note that the said values of P1, P2 and P5 have been considered based on empirical observations. We take the value of P7 as 11×11, same as that in Section 4.2, and we also provide some results with other values of P7 to justify the said choice. The value of P7 is kept same for computations at the border areas of images by applying zero-padding.

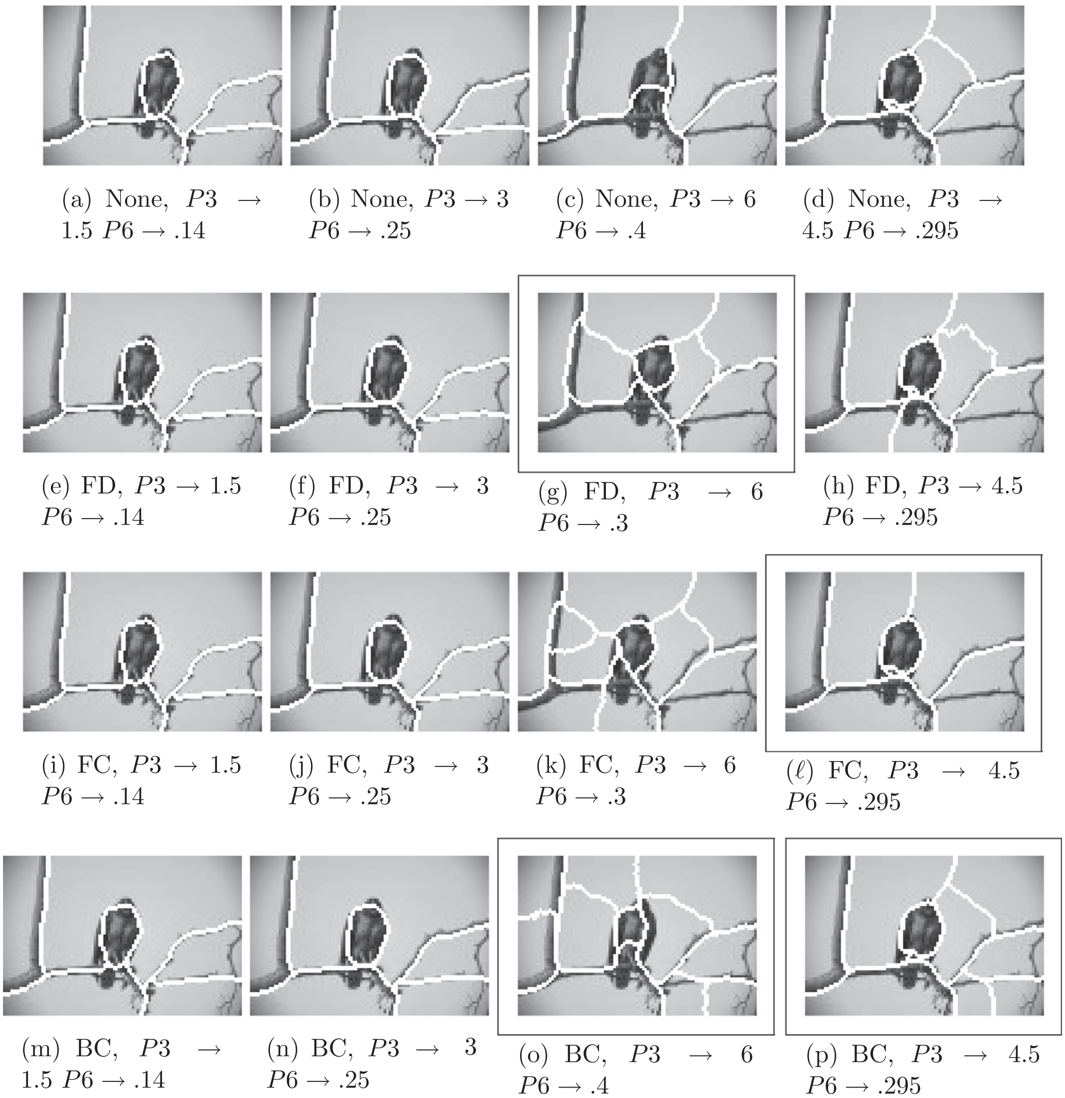
5.1. Qualitative analysis

Here we shall consider the qualitative evaluation of performance of normalized cut based graph partitioning for grouping of image pixels with and without the incorporation of local image structure. Note that as mentioned in Section 1, we consider grouping of image pixels in the spirit of perceptual grouping only in early human vision. Perceptual grouping occurs at multiple stages in human visual processing and hence only a technique (segmentation) trying to mimic perceptual grouping based on the entire human visual processing should be expected to segment the image into meaningful regions. However, a technique (grouping of image pixels), like the ones considered here, trying to mimic perceptual grouping based only on the human visual processing at the early stages should be expected to produce crude results in the form of homogeneous and spatially compact groups. These groups obtained through grouping of image pixels can be conceived as an initial estimate of the final segmentation result one wishes to obtain. Hence, one would like a technique of grouping of image pixels to produce homogeneous and spatially compact groups that are close to meaningful regions of the underlying image.

In the qualitative analysis, the values of the parameters P1, P2, P4, P5 and P7 are considered as mentioned earlier and the values of the parameters P3 and P6 are mentioned at the concerned places. In this analysis, for a single image, four different values of P3 are taken and that value of P6 is considered which gives the best (by visual observation) result when the other parameters are fixed as mentioned. When the results of grouping of image pixels performed with the incorporation of local image structure (proposed) are shown, we employ 'FD', 'FC' and 'BC' respectively to indicate the use of neighborhood alikeness based on fuzzy divergence, fuzzy correlation and Bhattacharyya coefficient. On the other hand, when the results of grouping of image pixels performed without the incorporation of local image structure (existing) are shown, we use 'None'.

It is known from [37] that for fixed values of other parameters, the normalized cut based graph partitioning technique for grouping of image pixels with a larger value of P6 would yield more number of groups. On the otherhand, for fixed values of other parameters, a larger value of P3 would yield less number of groups. The value of P3 is a matter of user's choice and the value P6 needs to be chosen with all other parameters fixed such that the best result is obtained. Note that a larger value of P6 is usually associated with a larger value of P3. As P3 is a matter of user's choice, it is desirable that acceptable results are yielded across a range of values of P3. Hence, we test the repeatability of acceptable results for different values of P3 while comparing results of image pixel grouping with and without the incorporation of local image structure.

Now, let us consider the results of grouping of image pixels shown in Figs. 5–8. In the mentioned figures, the type of neighborhood alikeness measure used is the same row-wise and the value of P3 considered is the same column-wise. For a given value of P3, we compare the performance of grouping of image pixels obtained with (all three – FD, FC and BC) and without the incorporation of local image structure and make our observations. Note that, as our goal is only to determine



(a) None, $P3 \rightarrow 1.5$ $P6 \rightarrow .14$
(b) None, $P3 \rightarrow 3$ $P6 \rightarrow .25$
(c) None, $P3 \rightarrow 6$ $P6 \rightarrow .4$
(d) None, $P3 \rightarrow 4.5$ $P6 \rightarrow .295$
(e) FD, $P3 \rightarrow 1.5$ $P6 \rightarrow .14$
(f) FD, $P3 \rightarrow 3$ $P6 \rightarrow .25$
(g) FD, $P3 \rightarrow 6$ $P6 \rightarrow .3$
(h) FD, $P3 \rightarrow 4.5$ $P6 \rightarrow .295$
(i) FC, $P3 \rightarrow 1.5$ $P6 \rightarrow .14$
(j) FC, $P3 \rightarrow 3$ $P6 \rightarrow .25$
(k) FC, $P3 \rightarrow 6$ $P6 \rightarrow .3$
(ℓ) FC, $P3 \rightarrow 4.5$ $P6 \rightarrow .295$
(m) BC, $P3 \rightarrow 1.5$ $P6 \rightarrow .14$
(n) BC, $P3 \rightarrow 3$ $P6 \rightarrow .25$
(o) BC, $P3 \rightarrow 6$ $P6 \rightarrow .4$
(p) BC, $P3 \rightarrow 4.5$ $P6 \rightarrow .295$

Fig. 5. Results of grouping of image pixels by normalized cut based graph partitioning with and without the incorporation of local image structure in an image containing tree branches and a bird.

the utility of the incorporation of local image structure, the results shown here may not contain the best result possible ones that may be obtained choosing the best values for the parameters P1–P7 varying all of them across their possible ranges.

As can be seen in Fig. 5, when the value of P3 is considered as 1.5 and 3, all the results of grouping of image pixels obtained with and without the incorporation of local image structure are the same and sufficiently satisfactory. However, when the value of P3 is taken as 6 and 4.5, we observe that a few results are better than the others. When P3 is taken as 6, the incorporation of local image structure using fuzzy divergence and Bhattacharyya coefficient based neighborhood alikeness measurement gives better result compared to the other two. The use of fuzzy divergence produces a result that locates the bird and contains spatially compact groups with acceptable homogeneity. Whereas the use of Bhattacharyya coefficient produces a result that roughly locates the bird, and contains homogeneous and compact groups that include those below the lower branch at the right-bottom of the image. In the other two, either the bird is not located or some homogeneous areas are unnecessarily dissected without any gains elsewhere. When P3 is taken as 4.5, the incorporation of local image structure using fuzzy correlation and Bhattacharyya coefficient based neighborhood alikeness measurement gives better result compared to the other two. The use of fuzzy correlation produces a result that locates the bird and does not dissect any fairly homogeneous group. Whereas the use of Bhattacharyya coefficient produces a result that locates the bird, and contains homogeneous and compact groups that include those below the lower branch at the right-bottom of the image. In the other two, some homogeneous areas are unnecessarily dissected without any gains elsewhere.

As can be seen in Fig. 6, when P3 is taken as 1.5, the non-incorporation of local image structure and incorporation of local image structure using fuzzy correlation based neighborhood alikeness measurement give better result compared to the other

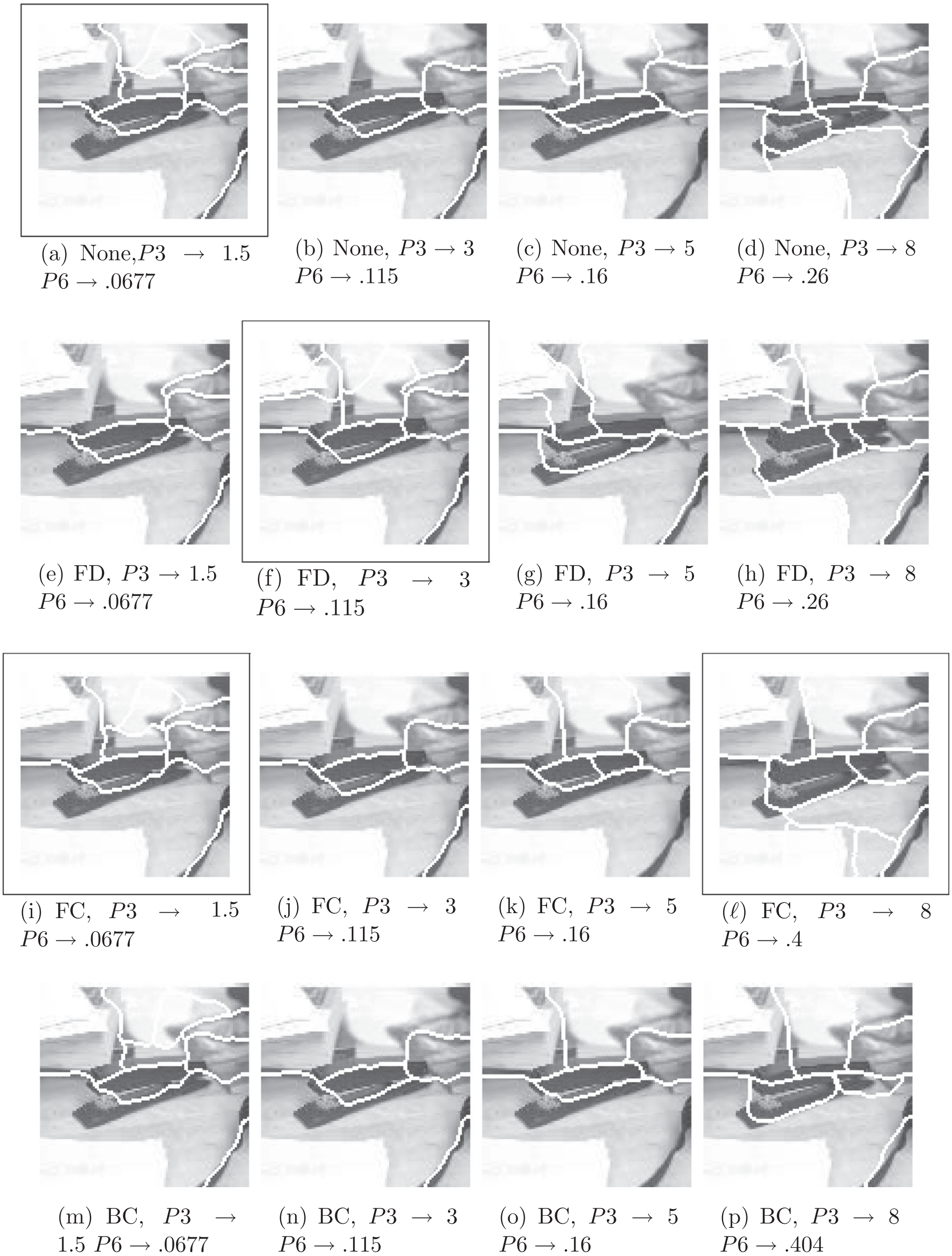
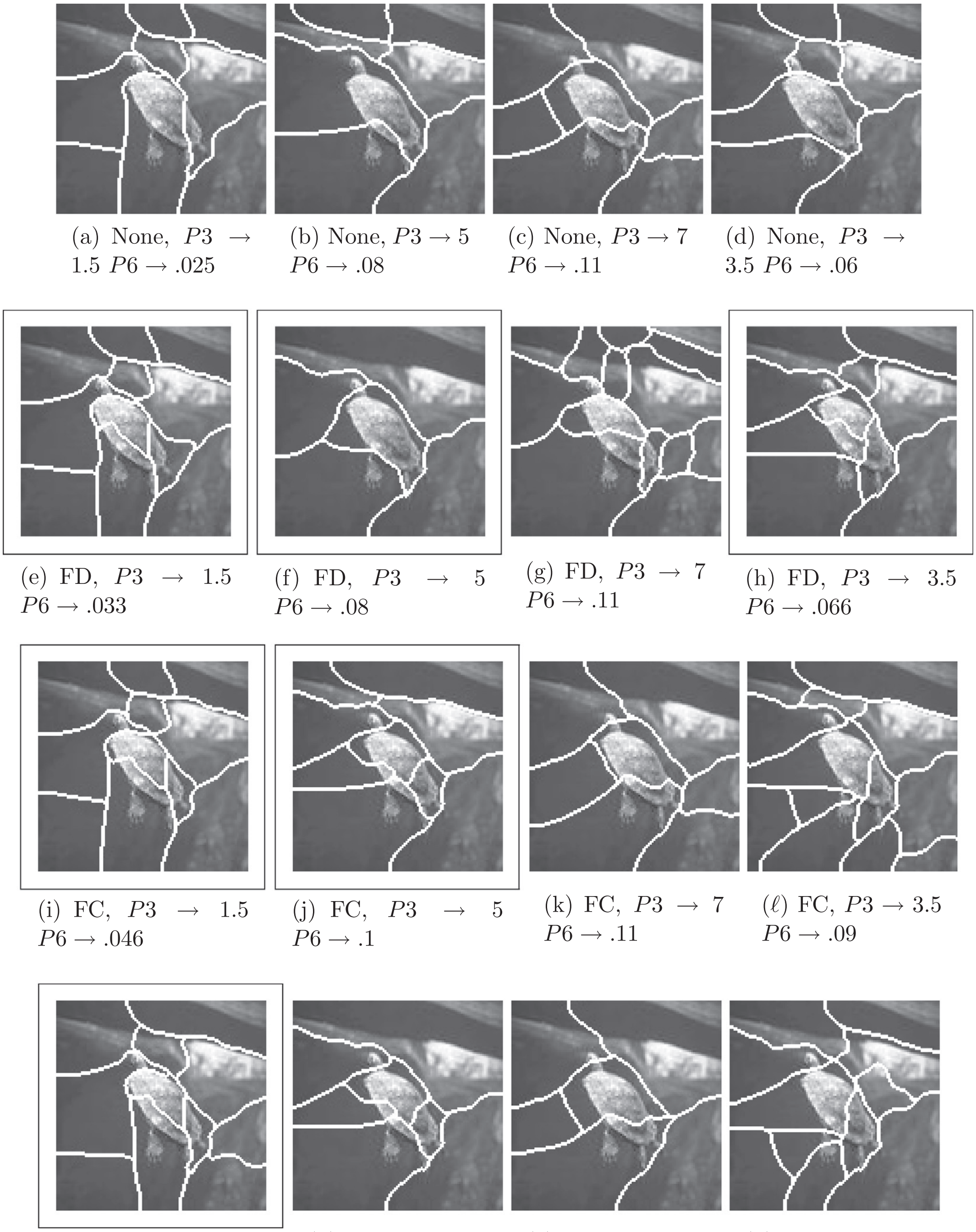


Fig. 6. Results of grouping of image pixels by normalized cut based graph partitioning with and without the incorporation of local image structure in an image containing papers, a book and a stapler on a table.

two. Both the non-incorporation and the fuzzy correlation based incorporation produce results that roughly locate the stapler, the book and most of the papers, while only a few homogeneous groups are dissected. In the other two, either a few sets of papers are not located or some homogeneous areas are unnecessarily dissected without any gains elsewhere. When P3 is taken as 3, the incorporation of local image structure using fuzzy divergence based neighborhood alikeness measurement gives better result compared to the other three. The use of fuzzy divergence produces a result that roughly locates most of the papers, the book and the stapler, although a few homogeneous groups are dissected. In all the other three, the papers



(a) None, $P3 \rightarrow 1.5$ $P6 \rightarrow .025$ (b) None, $P3 \rightarrow 5$ $P6 \rightarrow .08$ (c) None, $P3 \rightarrow 7$ $P6 \rightarrow .11$ (d) None, $P3 \rightarrow 3.5$ $P6 \rightarrow .06$

(e) FD, $P3 \rightarrow 1.5$ $P6 \rightarrow .033$ (f) FD, $P3 \rightarrow 5$ $P6 \rightarrow .08$ (g) FD, $P3 \rightarrow 7$ $P6 \rightarrow .11$ (h) FD, $P3 \rightarrow 3.5$ $P6 \rightarrow .066$

(i) FC, $P3 \rightarrow 1.5$ $P6 \rightarrow .046$ (j) FC, $P3 \rightarrow 5$ $P6 \rightarrow .1$ (k) FC, $P3 \rightarrow 7$ $P6 \rightarrow .11$ (ℓ) FC, $P3 \rightarrow 3.5$ $P6 \rightarrow .09$

(m) BC, $P3 \rightarrow 1.5$ $P6 \rightarrow .032$ (n) BC, $P3 \rightarrow 5$ $P6 \rightarrow .1$ (o) BC, $P3 \rightarrow 7$ $P6 \rightarrow .11$ (p) BC, $P3 \rightarrow 3.5$ $P6 \rightarrow .1$

Fig. 7. Results of grouping of image pixels by normalized cut based graph partitioning with and without the incorporation of local image structure in an image containing a turtle and a rocky structure.

and the book are properly not located. When P3 is taken as 5, all the results of grouping of image pixels obtained with and without the incorporation of local image structure are of similar quality. In them, either a few of the mentioned objects and the group at the right-bottom of the image are not located or some homogeneous and spatially compact groups are unnecessarily dissected. However, in all of them, a majority of suitable objects/groups are satisfactorily located. When P3 is taken as 8, the incorporation of local image structure using fuzzy correlation based neighborhood alikeness measurement gives better result compared to the other three. The use of fuzzy correlation produces a result that roughly locates the stapler,

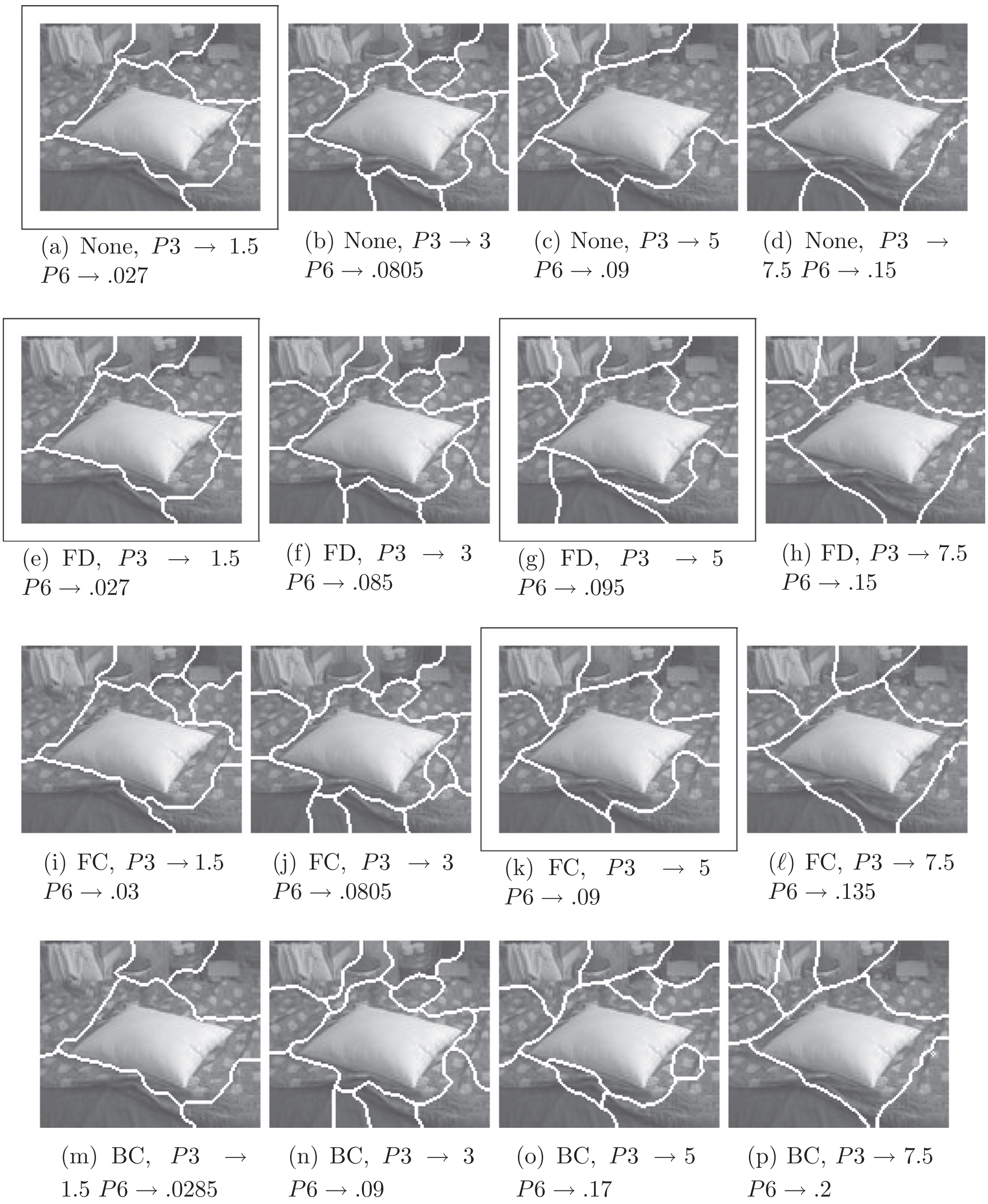


Fig. 8. Results of grouping of image pixels by normalized cut based graph partitioning with and without the incorporation of local image structure in an image containing a pillow on a bed covered with a bed sheet along with some small objects in the background.

the book, the papers including the one at the left-bottom of the image and the group at the right-bottom of the image, although a few homogeneous groups are dissected. In the other three, either homogeneous groups are unnecessarily dissected or a few objects/groups are not located.

As can be seen in Fig. 7, when P3 is taken as 1.5, the incorporation of local image structure by all the three means gives better result compared to the non-incorporation of the same. The said incorporation produces results that roughly locate the turtle unlike the result produced when the non-incorporation is considered. When P3 is taken as 5, the incorporation of local image structure using fuzzy divergence and fuzzy correlation based neighborhood alikeness measurement gives better result compared to the other two. The non-incorporation of local image structure produces a result that fails to locate the turtle; however, it contains grouping in the rocky structure. The use of Bhattacharyya coefficient produces a result that roughly locates the turtle and contains no grouping in the rocky structure unlike the use of fuzzy correlation, which produces a result that not only roughly locates the turtle but also contains grouping in the rocky structure. The use of fuzzy divergence produces a result that locates the turtle better than the other three, although it does not contain grouping in the rocky structure. When P3 is taken as 7, all the results of grouping of image pixels obtained with and without the incorporation of local image structure are of similar quality. In them, either a few of the objects present are not located or some homogeneous and spatially compact groups are unnecessarily dissected. However, in all of them, the turtle is roughly located. When P3 is taken as 3.5, the incorporation of local image structure using fuzzy divergence based neighborhood alikeness measurement gives better result compared to the other three. The use of fuzzy divergence, unlike the other three, produces a result that roughly locates the turtle as well as only a few homogeneous and compact groups are dissected.

As can be seen in Fig. 8, all the results of grouping of image pixels are not satisfactory in terms of locating the small objects in the background. When P3 is taken as 1.5, the non-incorporation of local image structure and incorporation of local image structure using fuzzy divergence based neighborhood alikeness measurement give better result compared to the other two. Both the non-incorporation and the fuzzy divergence based incorporation produce results that locates the pillow better than the other two. When P3 is taken as 5, the incorporation of local image structure using fuzzy divergence and fuzzy correlation based neighborhood alikeness measurement gives better result compared to the other two. The use of fuzzy divergence and fuzzy correlation produce results that roughly locates the pillow and contains only a few unwanted dissections. In the other two, either the pillow is not properly located or there are too many unwanted dissections. When the value of P3 is considered as 3 and 7.5, all the results of grouping of image pixels obtained with and without the incorporation of local image structure are almost the same and roughly locate the pillow.

Note that in Figs. 5–8, the better results among the relevant lot have been marked using a black colored box. It can be noticed that the box appears 7 times around a result produced considering local image structure incorporation using fuzzy divergence, 6 times considering incorporation using fuzzy correlation, 3 times considering incorporation using Bhattacharyya coefficient and 2 times considering non-incorporation of local image structure. In relation to the discussion in Section 4, the said observation suggests the following:

(1) The characteristics of local structure/context around a pixel can be captured using a neighborhood around the pixel and appropriate use of local image structure in grouping of image pixels improves performance.
(2) Representing a neighborhood using a fuzzy set defined over an universe (array of pixels) is more appropriate than representing a neighborhood using local histogram.
(3) Fuzzy divergence and fuzzy correlation based neighborhood alikeness values indeed posses the ability to represent the affinity among pixels due to the common fate, common region and continuity cues, which apart from the similarity and proximity cues are the ones that might play a crucial role in grouping of image pixels. The said ability of fuzzy divergence and fuzzy correlation makes the use of local image structure based on them in grouping of image pixels effective in terms of performance.

In order to obtain the results shown in Figs. 5–8, we have taken the value of parameter P7 as 11×11. Let us consider the results in Figs. 9 and 10 in order to justify the said choice. In order to arrive at the value 11×11 of P7, we considered values from 5×5 to 17×17. Figs. 9 and 10 show the results of grouping of image pixels with the value of $P7$ as 5×5 and 17×17, respectively. These results corresponds to only the better results (marked using a black colored box, same value of P3 and P6 used) obtained in Figs. 5–8 when local image structure is incorporated, as these results would demonstrate the reason for choosing the value of P7 as 11×11. As can be seen from Fig. 10, there is no difference in the results obtained using P7 as 11×11 and 17×17. However, it is evident from Fig. 9, that some results obtained using P7 as 5×5 is different from that when P7 is 11×11 and is similar to those obtained without the incorporation of local image structure. This suggests that when P7 is 5×5, in some cases, the local image structure is not appropriately captured. We made such observations proceeding from P7 as 5×5, until we considered P7 as 11×11, after which the results remained the same till P7 as 17×17. As computation time increases with increase in the size represented by P7, we consider 11×11 as the appropriate value of P7. Note that, although proceeding further with the value of P7 might yield results which are different to when P7 is 11×11, we ignore them as use of such values of P7 would be computationally inefficient.

We shall now consider quantitative analysis using a large number of images in order to find out whether there is statistical evidence of the statements enumerated above.

5.2. Quantitative analysis

Here, we consider human labeled segmentation ground truth based quantitative evaluation of performance in order to carry out rigorous analysis and establish the effectiveness of incorporating local image structure in normalized cut based graph partitioning for grouping of image pixels. Note that techniques for grouping of image pixels are expected to produce only crude results in the form of homogeneous and spatially compact groups and not human-like results in the form of meaningful regions, which can be expected from segmentation techniques. In spite of the aforesaid fact, we consider human labeled segmentation ground truth for evaluating performance of grouping of image pixels because result of grouping of image pixels can be conceived as an initial estimate of the final segmentation result desired, and hence, one would like grouping of image pixels to produce homogeneous and spatially compact groups that are close to meaningful regions of the underlying image. Comparisons are carried out between the results of grouping of image pixels obtained using fuzzy

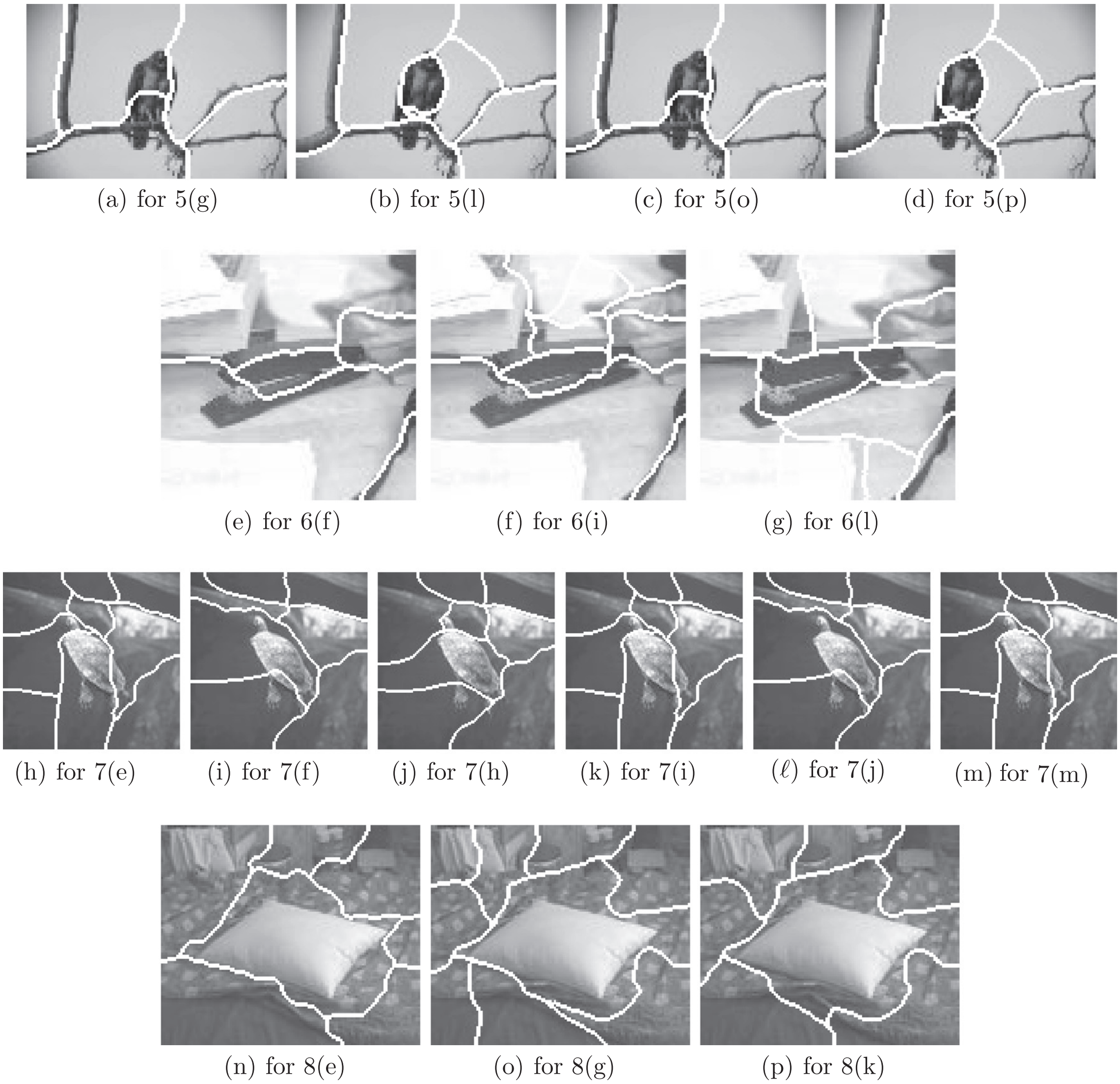


Fig. 9. Results of grouping of image pixels by normalized cut based graph partitioning with the value of *P*7 as 5×5.

divergence, fuzzy correlation and Bhattacharyya coefficient based local image structure incorporation and the results of grouping of image pixels obtained without the incorporation of local image structure.

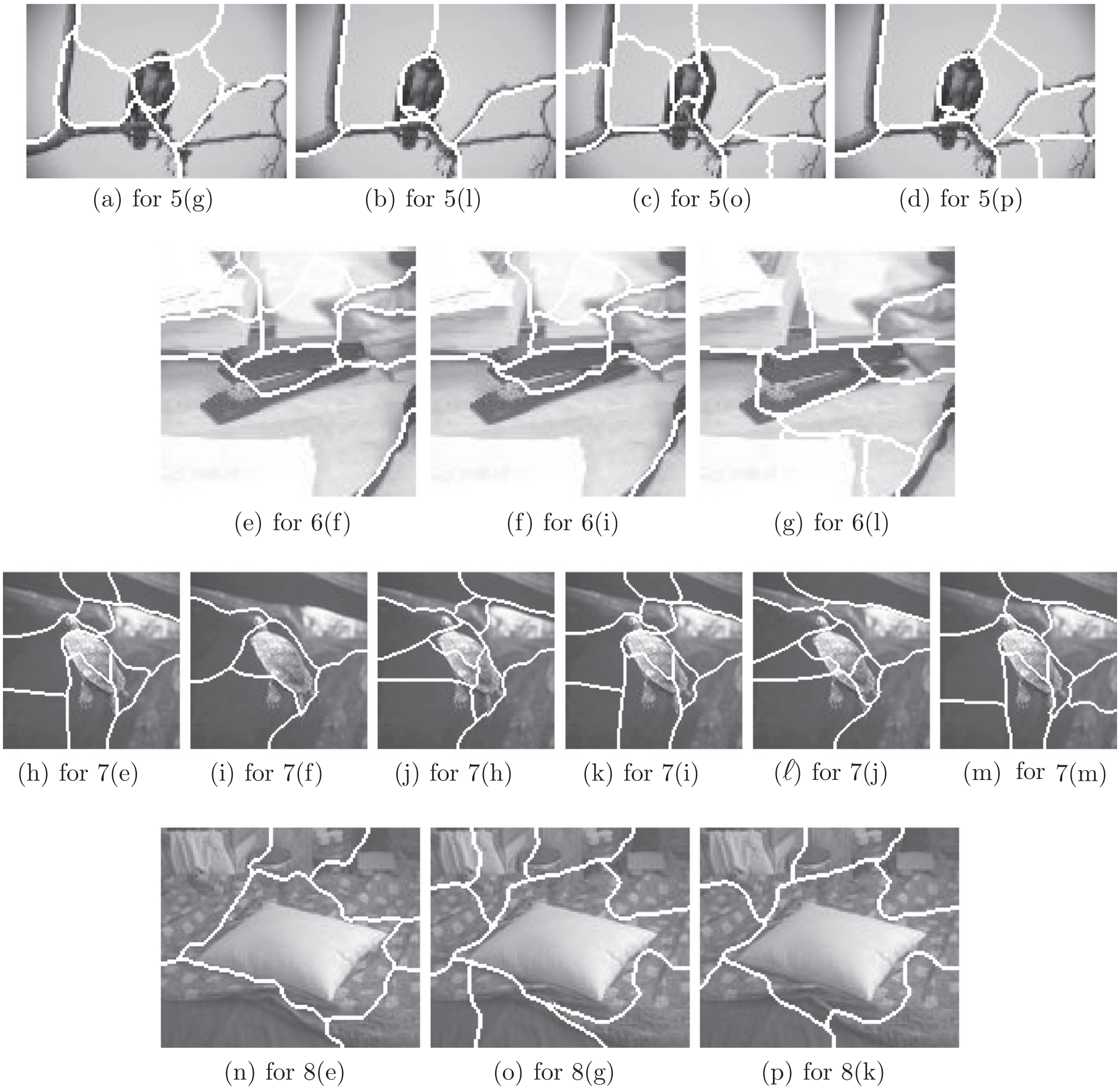
In the quantitative analysis, the value of the parameter P7 is considered as 11×11, the values of the parameters P1, P2, P4, and P5 are considered as mentioned earlier, and the values of the parameters P3 and P6 are fixed at 2.5 and 0.05, respectively.

5.2.1. The image dataset considered

We consider 100 grayscale images from the 'Berkeley Segmentation Dataset and Benchmark' [27] (http://www.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/segbench/, the test set of images in BSDS300). Each one of the 100 images considered are associated with multiple segmentation results hand labeled by multiple human subjects and hence we have multiple segmentation ground truths for every single image.

5.2.2. The evaluation measures considered

We use the global consistency error (GCE) and the local consistency error (LCE) measures defined in [27] in order to judge the appropriateness of results of grouping of image pixels. Consider S_H as a segmentation result hand labeled by a human subject and S_A as a result of grouping of image pixels obtained applying an algorithm. The GCE and LCE measures representing the appropriateness of S_A with reference to the ground truth S_H are given as



(a) for 5(g) (b) for 5(l) (c) for 5(o) (d) for 5(p)

(e) for 6(f) (f) for 6(i) (g) for 6(l)

(h) for 7(e) (i) for 7(f) (j) for 7(h) (k) for 7(i) (ℓ) for 7(j) (m) for 7(m)

(n) for 8(e) (o) for 8(g) (p) for 8(k)

Fig. 10. Results of grouping of image pixels by normalized cut based graph partitioning with the value of *P*7 as 17 × 17.

$$\mathrm{GCE}(S_H, S_A) = \frac{1}{n} \min \left\{ \sum_{i=1}^{n} E(S_H, S_A, p_i), \sum_{i=1}^{n} E(S_A, S_H, p_i) \right\} \tag{20}$$

$$\mathrm{LCE}(S_H, S_A) = \frac{1}{n} \sum_{i=1}^{n} \min\{E(S_H, S_A, p_i), E(S_A, S_H, p_i)\} \tag{21}$$

where

$$E(S_1, S_2, p) = \frac{|R(S_1, p) \setminus R(S_2, p)|}{|R(S_1, p)|} \tag{22}$$

In the above, $\setminus$ represents set difference, $|x|$ represents the cardinality of a set x, $R(S, p)$ represents the set of pixels corresponding to the region/group in segmentation result/result of grouping of image pixels S that contains the pixel p and n represents the number of pixels in the image under consideration. Both GCE and LCE take values in the range [0, 1] and GCE is a tougher measure than LCE, that is, LCE $(S_H, S_A) \leqslant$ GCE (S_H, S_A). In case of both these error measures, a smaller value indicates more appropriateness of the result of image pixel grouping S_A (with reference to the ground truth S_H). One may refer to [27] for an elaborate description of the aforesaid error measures.

5.2.3. Statistical analysis of performance

We calculate the GCE and LCE measures corresponding to all the results of grouping of image pixels obtained using fuzzy divergence, fuzzy correlation and Bhattacharyya coefficient based local image structure incorporation, and without the incorporation of local image structure, with reference to all segmentation ground truths available for every image among the 100 images considered.

When one algorithm for grouping of image pixels is considered for one image, a set of GCE values and a set of LCE values are obtained. Every element in a set of GCE/LCE values corresponds to one of the multiple segmentation ground truths available for the image considered and hence the number of elements in a set of GCE/LCE values obtained for an image equals the number of segmentation ground truths associated with that image.

While considering each one of the 100 images, we take the help of statistical hypothesis testing in order to compare the results of grouping of image pixels obtained incorporating local image structure (aforementioned all three ways) and not incorporating the same. The comparisons are carried out considering the error measures, that is, the sets of GCE and LCE values, obtained when local image structure is incorporated and not incorporated. We use the statistical *t*-test [20] assuming that the GCE values and the LCE values corresponding to those being compared have come from normally distributed populations with unknown and possibly unequal variances.

We perform one-sided *t*-tests [20] considering the alternative hypothesis (H_1) that 'the average error occurred when local image structure is not incorporated is greater than the average error occurred when local image structure is incorporated'. The *p*-value [20] obtained from such a *t*-test gives the probability that a superior performance obtained, when local image structure is incorporated, would have been due to chance alone. Such one-sided *t*-tests are performed for all the 100 images. The *p*-values obtained from these tests are shown using blue colored bars in Figs. 11 and 12.

We also perform one-sided *t*-tests considering the alternative hypothesis (H_2) that 'the average segmentation error occurred when local image structure is incorporated is greater than the average segmentation error occurred when local image structure is not incorporated'. The *p*-value obtained from such a *t*-test gives the probability that a superior performance obtained, when local image structure is not incorporated, would have been due to chance alone. Such one-sided *t*-tests are performed for all the 100 images. The *p*-values obtained from these tests are shown using red colored bars in Figs. 11 and 12.

The segmentation error considered in the experiments represented by Fig. 11 is GCE and the segmentation error considered in the experiments represented by Fig. 12 is LCE. In Figs. 11 and 12, Figs. 11 and 12, and Fig. 12a, Fig. 12a, and Fig. 12b, and Fig. 12b, and Fig. 12c correspond to the comparisons between grouping of image pixels with fuzzy divergence (FD), fuzzy correlation (FC) and Bhattacharyya coefficient (BC) based local image structure incorporation and grouping of image pixels without ('None') local image structure incorporation, respectively.

Recall that we have considered the value of the parameter P6 as 0.05. The said value of P6 is fixed based on the empirical observation that neither too many or too few groups are formed when all the techniques for grouping of image pixels considered here for comparison are applied on the 100 images. This observation is important as the usage of GCE and LCE measures requires such a condition in order to appropriately calculate segmentation error value with respect to a ground truth [27].

It is evident from the illustrations in Fig. 11a, Fig. 12a, Fig. 12a, and Fig. 12b that the blue colored bars are in general shorter than the red colored bars. This signifies that, for most of the 100 images considered, it is less likely that a better performance of grouping of image pixels obtained when local image structure is incorporated using fuzzy divergence and fuzzy correlation is merely due to chance alone and not due to a real effect, compared to the case when local image structure is not incorporated. However, from the illustrations in Fig. 12b and Fig. 12c, we see that the blue colored bars are shorter than the red colored bars only in a slight majority of the 100 images. This signifies that the likeliness of a better performance of grouping of image pixels obtained when local image structure is incorporated using Bhattacharyya coefficient is merely due to chance alone and not due to a real effect is only slightly less than the case when local image structure is not incorporated.

The aforesaid observations point to the clear superiority of results of grouping of image pixels obtained incorporating local image structure using fuzzy divergence and fuzzy correlation over those obtained not incorporating local image structure. They also point to the slight superiority of results of grouping of image pixels obtained incorporating local image structure using Bhattacharyya coefficient over those obtained not incorporating local image structure.

Now, using all the *p*-values (shown in Fig. 11 and Fig. 12) obtained corresponding to all the 100 images, we perform another statistical analysis in order to compare the *p*-values obtained performing the aforementioned *t*-tests with alternative hypotheses H_1 and H_2. It is assumed that the *p*-values corresponding to the *t*-tests with alternative hypotheses H_1 and H_2 have come from normally distributed populations with unknown and possibly unequal variances. We perform one-sided *t*-tests considering the following two alternative hypotheses:

H_I: The average *p*-value obtained from the *t*-test with alternative hypothesis H_2 is greater than the average *p*-value obtained from the *t*-test with alternative hypothesis H_1.

H_{II}: The average *p*-value obtained from the *t*-test with alternative hypothesis H_1 is greater than the average *p*-value obtained from the *t*-test with alternative hypothesis H_2.

The *p*-values obtained from the aforesaid *t*-tests are given in Table 5. The *p*-values listed in the table correspond to GCE and LCE based comparison between the results of grouping of image pixels obtained using fuzzy divergence (FD), fuzzy correlation (FC) and Bhattacharyya coefficient (BC) based local image structure incorporation and the results of grouping of image

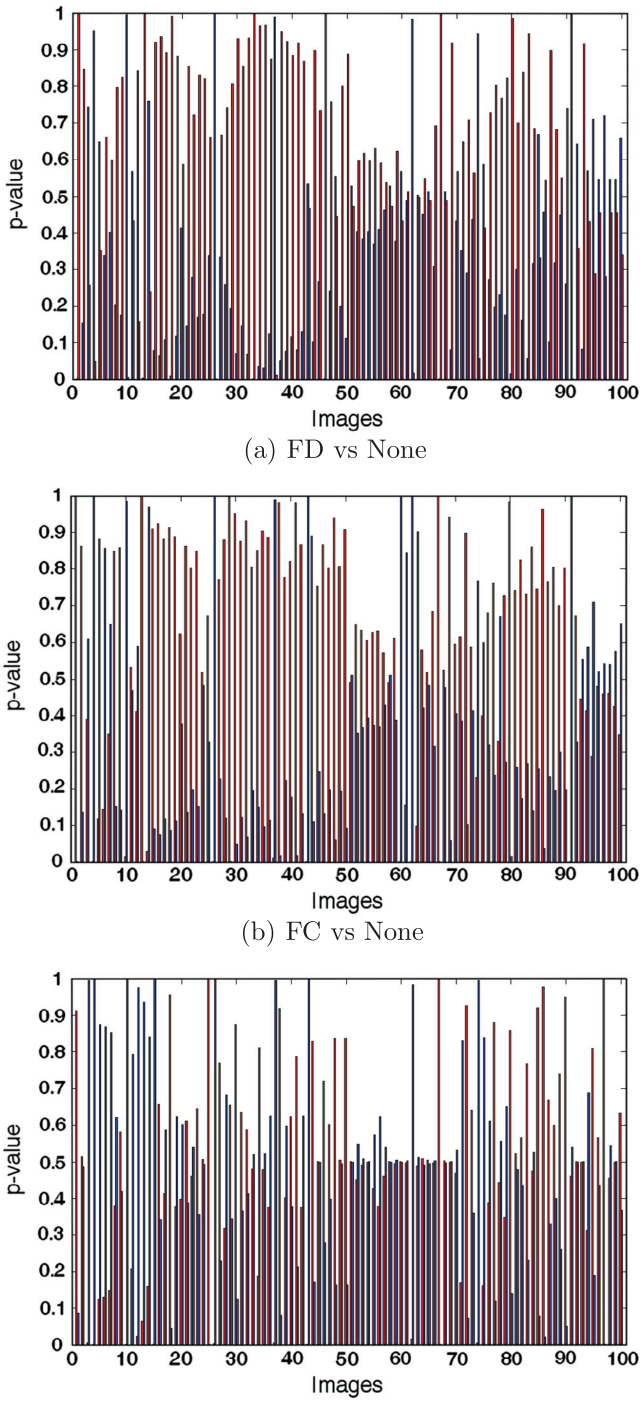
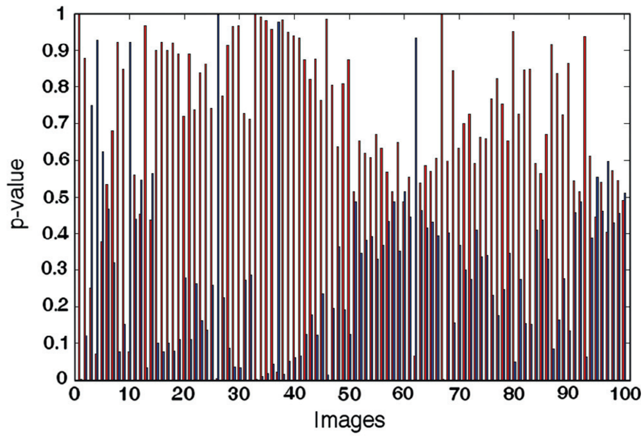
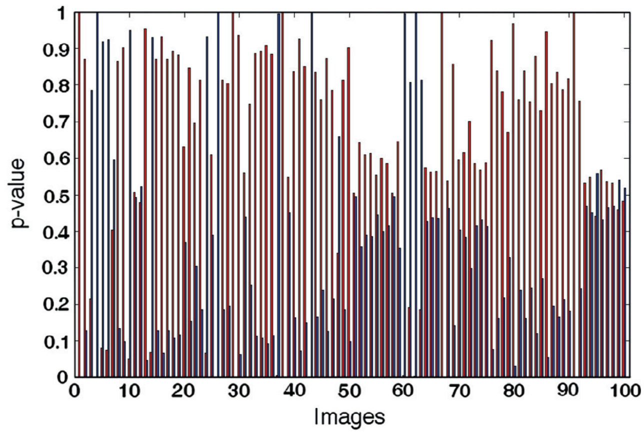


Fig. 11. The *p*-values obtained corresponding to the one sided *t*-tests with alternative hypotheses H_1 and H_2 considering GCE.

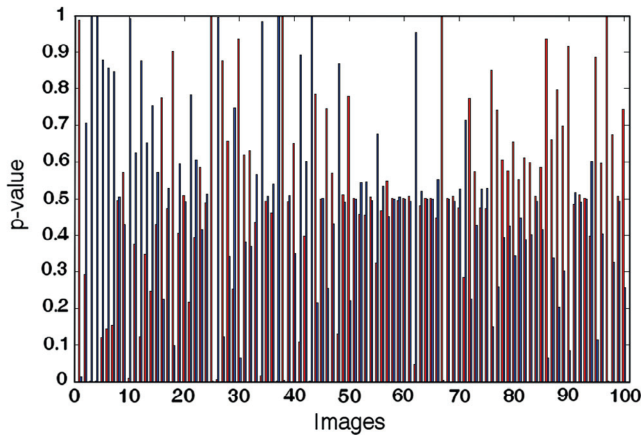
pixels obtained without ('None') the incorporation of local image structure. We observe from Table 5 that in the cases when fuzzy divergence and fuzzy correlation based local image structure incorporation is compared to the non-incorporation of local image structure, the *p*-values obtained for the *t*-tests performed with alternative hypothesis H_{I} are considerably smaller than the corresponding *p*-values obtained for the *t*-tests performed with alternative hypothesis H_{II}. This observation suggests that it is almost unlikely that a smaller *p*-value obtained when the *t*-tests with alternative hypothesis H_1 is performed is merely due to chance alone and not due to a real effect. On the other hand, it also suggests that it is almost certain that a



(a) FD vs None



(b) FC vs None



(c) BC vs None

Fig. 12. The *p*-values obtained corresponding to the one sided *t*-tests with alternative hypotheses H_1 and H_2 considering LCE.

smaller *p*-value obtained when the *t*-tests with alternative hypothesis H_2 is performed is merely due to chance alone and not due to a real effect. However, when the case of comparison of Bhattacharyya coefficient based local image structure incorporation with the non-incorporation of local image structure is considered, it is observed that the *p*-values obtained for the *t*-tests performed with alternative hypothesis H_{I} are only slightly smaller than the corresponding *p*-values obtained for the *t*-tests performed with alternative hypothesis H_{II}. Therefore, it is found that there is absolute superiority of results of grouping of image pixels obtained incorporating local image structure using fuzzy divergence and fuzzy correlation over those

Table 5
The *p*-values obtained corresponding to the one sided *t*-tests with alternative hypotheses H_I and H_{II}.

Comparison	Alternative hypothesis	*p*-values	
		GCE	LCE
FD vs.	H_I	1.5476×10^{-12}	0
None	H_{II}	≈ 1	1
FC vs.	H_I	8.1719×10^{-10}	1.3243×10^{-11}
None	H_{II}	≈ 1	≈ 1
BC vs.	H_I	0.4974	0.2938
None	H_{II}	0.5026	0.7062

Table 6
Time consumed (in seconds) by grouping of image pixels with and without the incorporation of local image structure when applied on four different images

Local image structure using ⇒	None (s)	Fuzzy divergence (s)	Fuzzy orrelation (s)	Bhattacharyya coefficient (s)
Image in Fig. 5	12	30	15.5	19.5
Image in Fig. 6	29	57	33	40
Image in Fig. 7	38	72	44	52
Image in Fig. 8	40.5	74	46	54

obtained not incorporating local image structure and only a slight superiority of results of grouping of image pixels obtained incorporating local image structure using Bhattacharyya coefficient over those obtained not incorporating local image structure.

Based on the above given quantitative analysis, we infer that there is statistical evidence of the concluding statements enumerated at the end of qualitative analysis.

5.3. Time consumption analysis

The time consumed in performing the grouping of image pixels in the images of Figs. 5–8 using normalized cut based graph partitioning with fuzzy divergence (FD), fuzzy correlation (FC) and Bhattacharyya coefficient (BC) based local image structure incorporation and without ('None') local image structure incorporation is given in Table 6. The programming is performed using MATLAB® in a machine having Intel® core 2 duo processor with processing speed of 2.53 GHz each, and 2 gigabytes of RAM. It is evident from the table that as expected the time consumed when incorporation of local image structure is not considered is less than when the incorporation is done. Grouping of image pixels using fuzzy divergence based local image structure incorporation, which as evident from the qualitative and quantitative analyses is expected to produce the best results, consumes more time than that using the other two approaches of local image structure incorporation. The time consumed by grouping of image pixels using Bhattacharyya coefficient based local image structure incorporation is slightly more than the time consumed by that using fuzzy correlation based local image structure incorporation. It is empirically observed that the time taken by the algorithms for grouping of image pixels is dependent on the number pixels in the image under consideration and their dependency on the parameters P1–P7 is fairly insignificant.

6. Conclusion

Incorporation of local image structure/context in normalized cut based graph partitioning for grouping of image pixels have been proposed and studied in this paper. The said incorporation of local image structure have been proposed based on findings from the study of some perceptual cues acting in early human vision. Local image structure/context has been represented using neighborhood in the form of histogram and fuzzy set, and the said representation has been justified with respect to the perceptual cues. Experimental results have been given in order to demonstrate the effectiveness of the incorporation of local image structure/context into grouping of image pixels.

In addition to the similarity and proximity perceptual cues that are usually considered, common fate, common region and continuity perceptual cues have been considered in this paper for grouping of image pixels. It has been found that the incorporation of local image structure in normalized cut based graph partitioning for grouping of image pixels helps in adding the effect of the common fate, common region and continuity cues to that of the similarity and proximity cues. It has also been found that a neighborhood around a pixel can be appropriately used to represent the local image structure corresponding to that pixel. The proposed way of considering neighborhoods in the form of fuzzy sets has been found to be better than considering neighborhoods in the form of histogram. It has been found qualitatively and quantitatively that the incorporation of local image structure improves performance of grouping of image pixels. Note that, the approach of using the five perceptual cues depicted in this paper might be employed to fruitfully modify the existing algorithms (such as the one in [46]) that consider only the two standard cues, similarity and proximity.

References

[1] R. Belohlavek, M. Krupka, Grouping fuzzy sets by similarity, Information Sciences 179 (15) (2009) 2656–2661.
[2] M.B. Ben-Av, D. Sagi, Perceptual grouping by similarity and proximity: experimental results can be predicted by intensity autocorrelations, Vision Research 35 (6) (1995) 853–866.
[3] D. Bhandari, N.R. Pal, D.D. Majumdar, Fuzzy divergence, probability measure of fuzzy events and image thresholding, Pattern Recognition Letters 13 (12) (1992) 857–867.
[4] A. Bhattacharyya, On a measure of divergence between two statistical populations defined by their probability distributions, Bulletin of the Calcutta Mathematical Society 35 (1943) 99–109.
[5] Y. Boykov, O. Veksler, R. Zabih, Fast approximate energy minimization via graph cuts, IEEE Transactions on Pattern Analysis and Machine Intelligence 23 (11) (2001) 1222–1239.
[6] I.J. Cox, S.B. Rao, Y. Zhong, Ratio regions: a technique for image segmentation, in: Proceedings of the 13th International Conference on Pattern Recognition, vol. 2, 1996.
[7] I.S. Dhillon, Y. Guan, B. Kulis, Weighted graph cuts without eigenvectors a multilevel approach, IEEE Transactions on Pattern Analysis and Machine Intelligence 29 (11) (2007) 1944–1957.
[8] A. Eriksson, C. Olsson, F. Kahl, Normalized cuts revisited: a reformulation for segmentation with linear grouping constraints, Journal of Mathematical Imaging and Vision 39 (1) (2011) 45–61.
[9] J. Feldman, The role of objects in perceptual grouping, Acta Psychologica 102 (2–3) (1999) 137–163.
[10] P.F. Felzenszwalb, D.P. Huttenlocher, Efficient graph-based image segmentation, International Journal of Computer Vision 59 (2) (2004) 167–181.
[11] L. Grady, E.L. Schwartz, Isoperimetric graph partitioning for image segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence 28 (3) (2006) 469–475.
[12] Z.J. He, K. Nakayama, Surfaces versus features in visual search, Nature 359 (1992) 231–233.
[13] I.H. Jermyn, H. Ishikawa, Globally optimal regions and boundaries as minimum ratio weight cycles, IEEE Transactions on Pattern Analysis and Machine Intelligence 23 (10) (2001) 1075–1088.
[14] T. Kailath, The divergence and Bhattacharyya distance measures in signal selection, IEEE Transactions on Communication Technology 15 (1) (1967) 52–60.
[15] J.S. Kim, K.S. Hong, Color-texture segmentation using unsupervised graph cuts, Pattern Recognition 42 (5) (2009) 735–750.
[16] G. Klir, B. Yuan, Fuzzy Sets and Fuzzy Logic: Theory and Applications, Prentice Hall, New Delhi, India, 2005.
[17] W.L.G. Koontz, P.M. Narendra, K. Fukunaga, A graph-theoretic approach to nonparametric cluster analysis, IEEE Transactions on Computers 25 (9) (1976) 936–944.
[18] T.H. Le, S.W. Jung, K.S. Choi, S.J. Ko, Image segmentation based on modified graph-cut algorithm, Electronic Letters 46 (16) (2010) 1121–1123.
[19] C.H. Lee, O.R. Zaïane, H.H. Park, J. Huang, R. Greiner, Clustering high dimensional data: a graph-based relaxed optimization approach, Information Sciences 178 (23) (2008) 4501–4511.
[20] E.L. Lehmann, J.P. Romano, Testing Statistical Hypothesis, third ed., Springer, New York, USA, 2005.
[21] X. Liu, D. Wang, Image and texture segmentation using local spectral histograms, IEEE Transactions on Image Processing 15 (10) (2006) 3066–3077.
[22] D. Lowe, Perceptual Organization and Visual Recognition, Kluwer Academic Publishers, Norwell, MA, USA, 1985.
[23] J. Malcolm, Y. Rathi, A. Tannenbaum, A graph cut approach to image segmentation in tensor space, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2007.
[24] J. Malik, S. Belongie, T. Leung, J. Shi, Contour and texture analysis for image segmentation, International Journal of Computer Vision 43 (1) (2001) 7–27.
[25] J. Malik, S. Belongie, J. Shi, T. Leung, Textons, contours and regions: cue integration in image segmentation, in: Proceedings of the International Conference on Computer Vision, 1999.
[26] D. Marr, Vision: A Computational Investigation into the Human Representation and Processing of Visual Information, W.H. Freeman, San Francisco, CA, USA, 1982.
[27] D. Martin, C. Fowlkes, D. Tal, J. Malik, A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics, in: Proceedings of 8th International Conference on Computer Vision, vol. 2, 2001.
[28] M. Mignotte, Segmentation by fusion of histogram-based K-means clusters in different color spaces, IEEE Transactions on Image Processing 17 (5) (2008) 780–787.
[29] C.A. Murthy, S.K. Pal, D.D. Majumder, Correlation between two fuzzy membership functions, Fuzzy Sets and Systems 17 (1) (1985) 23–38.
[30] K. Ni, X. Bresson, T. Chan, S. Esedoglu, Local histogram based segmentation using the wasserstein distance, International Journal of Computer Vision 84 (1) (2009) 97–111.
[31] T. Ojala, M. Pietikainen, D. Harwood, A comparative study of texture measures with classification based on featured distributions, Pattern Recognition 29 (1) (1996) 51–59.
[32] S.E. Palmer, J.L. Brooks, R. Nelson, When does grouping happen?, Acta Psychologica 114 (3) (2003) 311–330
[33] M. Pavan, M. Pelillo, A new graph-theoretic approach to clustering and segmentation, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, 2003.
[34] I. Rock, L. Brosgole, Grouping based on phenomenal proximity, Journal of Experimental Psychology 67 (6) (1964) 531–538.
[35] M.B. Salah, A. Mitiche, I.B. Ayed, Multiregion image segmentation by parametric kernel graph cuts, IEEE Transactions on Image Processing 20 (2) (2011) 545–557.
[36] J. Shi, J. Malik, Normalized cuts and image segmentation, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 1997.
[37] J. Shi, J. Malik, Normalized cuts and image segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (8) (2000) 888–905.
[38] W. Tao, H. Jin, Y. Zhang, L. Liu, D. Wang, Image thresholding using graph cuts, IEEE Transactions on Systems, Man, and Cybernetics, A 38 (5) (2008) 1181–1195.
[39] D.A. Tolliver, G.L. Miller, Graph partitioning by spectral rounding: applications in image segmentation and clustering, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2006.
[40] R. Urquhart, Graph theoretical clustering based on limited neighbourhood sets, Pattern Recognition 15 (3) (1982) 173–187.
[41] S. Vicente, V. Kolmogorov, C. Rother, Graph cut based image segmentation with connectivity priors, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2008.
[42] S. Wang, J.M. Siskind, Image segmentation with ratio cut, IEEE Transactions on Pattern Analysis and Machine Intelligence 25 (6) (2003) 675–690.
[43] Y. Weiss, Segmentation using eigenvectors: a unifying view, in: Proceedings of the 7th IEEE International Conference on Computer Vision, vol. 2, 1999.
[44] M. Wertheimer, Laws of organization in perceptual forms, in: W.D. Ellis (Ed.), A Source Book of Gestalt Psychology, Kegan Paul, Trench, Trubner & Company, London, England, 1938, pp. 71–88.

[45] Z. Wu, R. Leahy, An optimal graph theoretic approach to data clustering: theory and its application to image segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence 15 (11) (1993) 1101–1113.
[46] Y. Xia, D. Feng, T. Wang, R. Zhao, Y. Zhang, Image segmentation by clustering of spatial patterns, Pattern Recognition Letters 28 (12) (2007) 1548–1555.
[47] C.T. Zahn, Graph-theoretical methods for detecting and describing gestalt clusters, IEEE Transactions on Computing 20 (1) (1971) 68–86.
[48] C. Zhu, R. Wang, Local multiple patterns based multiresolution gray-scale and rotation invariant texture classification, Information Sciences 187 (2012) 93–108.