

---

## Perceptual image analysis

---

C. Henry and J.F. Peters\*

Department of Electrical and Computer Engineering,  
Rm E2-390 EITC Bldg., 75A Chancellor's Circle,  
University of Manitoba,  
Winnipeg, MB R3T 5V6 Manitoba, Canada  
E-mail: [chenry@ee.umanitoba.ca](mailto:chenry@ee.umanitoba.ca)  
E-mail: [jfpeters@ee.umanitoba.ca](mailto:jfpeters@ee.umanitoba.ca)  
\*Corresponding author

**Abstract:** The problem considered in this paper is one of extracting perceptually relevant information from groups of objects based on their descriptions. Object descriptions are qualitatively represented by feature-value vectors containing probe function values computed in a manner similar to feature extraction in pattern classification theory. The work presented here is a generalisation of a solution to extracting perceptual information from images using near sets theory which provides a framework for measuring the perceptual nearness of objects. Further, near set theory is used to define a perception-based approach to image analysis that is inspired by traditional mathematical morphology and an application of this methodology is given by way of segmentation evaluation. The contribution of this article is the introduction of a new method of unsupervised segmentation evaluation that is based on human perception rather than on properties of ideal segmentations as is normally the case.

**Keywords:** perception; near sets; morphology; unsupervised segmentation evaluation; probe functions; image analysis.

**Reference** to this paper should be made as follows: Henry, C. and Peters, J.F. (2010) 'Perceptual image analysis', *Int. J. Bio-Inspired Computation*, Vol. 2, Nos. 3/4, pp.271–281.

**Biographical notes:** Christopher Henry received his MSc in 2006 from the University of Manitoba. Currently, he is a full time PhD student in the Department of Electrical and Computer Engineering (ECE) at the University of Manitoba. He is also the President of the ECE Graduate Students' Association (GSA) as well as a Councillor at the University of Manitoba GSA. Since 2004, he has published 12 articles (five journal publications, six conference publications and one book chapter) on topics in the areas of reinforcement learning, rough sets, intelligent systems, near sets and image processing. His current research interests are in reinforcement learning, rough set theory, approximation spaces, near sets, image processing, pattern recognition and intelligent systems.

James F. Peters, PhD (1991), is Co-founder and Research Group Leader in the Computational Intelligence Laboratory (<http://wren.ece.umanitoba.ca/>) and Full Professor in the Department of Electrical and Computer Engineering (ECE) at the University of Manitoba. In 2002, he collaborated with Zdzislaw Pawlak on a descriptive view of the nearness of physical objects. In 2006, he introduced near sets, a generalisation of rough sets. This has led to feature-based solutions to the image correspondence problem. His current research interests are in tolerance spaces, image analysis, image correspondence, fuzzy sets, near sets, especially tolerance near sets and rough sets.

---

### 1 Introduction

This paper introduces a near set solution to the segmentation evaluation problem. Image segmentation is a partition of an image into non-overlapping regions where each region is associated with some perceptual meaning. Here, the term *perception* is understood in the context of psychophysics where perception is based on the relationship between visual stimuli and sensations (Bruce et al., 1996). Our senses can be considered a form of sensor or probe function providing information about the objects in our environment. Thus, we are interested in perception of objects based on object

descriptions and perception of objects based on classes created by object descriptions. Recently, it has been shown that near sets can be used in a perception-based approach to discovering correspondences between images based on object descriptions (see, e.g., Gupta and Patnaik, 2008; Peters, 2009a, 2008b; Peters and Wasilewski, 2009). Moreover, the perceptual approach to image analysis presented in this article is an outgrowth of research into applying near sets to the image processing problems of segmentation, segmentation evaluation (Henry and Peters, 2008), pattern recognition (Henry and Peters, 2007, 2009;

Peters, 2009a), face recognition systems (Gupta and Patnaik, 2008) and medical imaging (Hassanien et al., 2009). Work on a basis for near sets began in 2002, motivated by image analysis and inspired by a study of the perception of the nearness of perceptual objects carried out in cooperation with Z. Pawlak in Pawlak and Peters (2002). This initial work led to the introduction of near sets in Peters (2007a), which were further elaborated in Peters, (2007b, 2009c) and Peters and Wasilewski (2009).

There is a strong connection between near sets and rough sets (see, e.g., Peters and Wasilewski, 2009; Peters, 2009a). Briefly, any non-empty set  $X$  with non-empty approximation boundary between upper and lower approximations is a set that is approximately classified and such a set is called a *rough set* (Pawlak, 1981). Sets containing objects with similar descriptions are called near sets. That is, any non-empty set  $X$  containing perceptual objects with description similar to the description of perceptual objects in a set  $Y$  is an example of a near set; i.e., the set  $X$  is *perceptually near* the set  $Y$ . It can be shown that every rough set is a near set and yet, there are near sets that are not rough sets. Hence, near sets (Peters, 2007a, 2007b, 2007c, 2009a, 2009c; Peters et al., 2007) are considered a generalisation of rough sets (Peters, 2007b). It can also be shown that there is a connection between near sets and fuzzy sets (see, e.g., Peters, 2009a, 2009b).

The discovery of near sets begins with the selection of probe functions that provide a basis for describing and discerning affinities between sample objects (see, e.g., Peters, 2007a, 2008a; Skowron and Peters, 2008). A *probe function* is a real-valued function representing features of physical objects. The perceptual approach of near set theory is based on the idea that our mind identifies relationships between object features to form perceptions of sensed objects. Our senses gather the information of the objects we perceive and map sensations to values assimilated by the mind. Thus, our senses can be likened to perceptual probe functions in the form of a mapping of stimuli from objects in our environment to sensations (values used by the mind to perceive objects). It is this idea of probe functions that is at the heart of near sets.

In keeping with an insight by Serra (1982) concerning the perception of images, one first transforms an image in some fashion and then one measures the transformed image. In our case, thanks to the discovery by Pawlak (1981), a special form of equivalence relation called an indiscernibility relation makes a feature-based transformation of images possible. This is achieved by viewing an image as a set of points and then discovering affinities by comparing the descriptions of images regions, e.g., descriptions of sub-images or pixels contained in classes found in partitions of images. In this article, we give a practical application of near sets to the problem of evaluating the quality of image segmentations by way of extracting perceptually relevant information from a set of objects where each object has an associated vector describing object features (perceived object characteristics such as colour). It is the information contained in these

vectors that are used to extract perceptual information from classes of objects and to measure similarity among them. The contribution of this article is an introduction to a new method of unsupervised segmentation evaluation. This article is organised as follows: Section 2 presents a review of perception-based image processing using the framework established by near sets. Section 3 introduces perceptual morphology and Section 4 gives an application of the new morphology by way of unsupervised segmentation evaluation.

## 2 Perceptual image processing

This section introduces near sets and gives examples of their application to image processing. Formally, let  $O$  represent the set of all objects. Then, a description of an object  $x \in O$  is given by:

$$\phi(x) = (\phi_1(x), \phi_2(x), \dots, \phi_l(x), \dots, \phi_l(x)),$$

where  $l$  is the length of the description and each  $\phi_l(x)$  is a probe function that describes the object  $x$ . Furthermore, we can define a set  $\mathbb{F}$  that represents all the probe functions used to describe an object  $x$ . Next, a perceptual information system  $S$  can be defined as  $S = \langle O, \mathbb{F} \{Val_{\phi_l}\}_{\phi_l \in \mathbb{F}} \rangle$ , where

$\mathbb{F}$  is the set of all possible probe functions that take as the domain objects in  $O$  and  $\{Val_{\phi_l}\}_{\phi_l \in \mathbb{F}}$  is the value range of a function  $\phi_l \in \mathbb{F}$ . For simplicity, a perceptual system is abbreviated as  $\langle O, \mathbb{F} \rangle$  when the range of the probe functions is understood.

*Definition 1. Indiscernibility Relation:* Let  $\langle O, \mathbb{F} \rangle$  be a perceptual system. For every  $\mathcal{B} \subseteq \mathbb{F}$  the indiscernibility relation  $\sim_{\mathcal{B}}$  is defined as follows:

$$\sim_{\mathcal{B}} = \{(x, y) \in O \times O : \|\phi(x) - \phi(y)\| = 0\},$$

where  $\|\cdot\|$  represents the  $l^2$  norm. If  $\mathcal{B} = \{\phi\}$  for some  $\phi \in \mathbb{F}$ , instead of  $\sim_{\{\phi\}}$  we write  $\sim_{\phi}$ .

Definition 1 is a refinement of the original indiscernibility relation given by Pawlak (1981). Using the indiscernibility relation, objects with matching descriptions can be grouped together forming granules of highest object resolution determined by the probe functions in  $\mathcal{B}$ . This gives rise to an elementary set:

$$x / \sim_{\mathcal{B}} = \{y \in O \mid y \sim_{\mathcal{B}} x\},$$

defined as a set where all objects have the same description. Similarly, a quotient set is the set of all elementary sets defined as:

$$O / \sim_{\mathcal{B}} = \{x / \sim_{\mathcal{B}} \mid x \in O\}.$$

Near set theory can be easily applied to images. For example, let us define a RGB image as  $f = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_T\}$ , where  $\mathbf{p}_i = (c, r, R, G, B)^T$ ,  $c \in [1, M]$ ,  $r \in [1, N]$ ,  $R, G, B \in [0, 255]$  and  $M, N$  respectively denote the width and height of the image and  $M \times N = T$ . Further, define a square sub-image as  $f_i \subset f$  with the following conditions:

$$\begin{aligned} f_1 \cap f_2 \dots \cap f_s &= \emptyset, \\ f_1 \cup f_2 \dots \cup f_s &= f, \end{aligned} \quad (1)$$

where  $s$  is the number of sub-images in  $f$ . We restrict all sub-images to be square except when doing so violates equation (1). For example, the images in the Berkeley segmentation dataset (Martin et al., 2001) often have the dimension  $321 \times 481$ . Consequently, a square sub-image size of 25 will produce 6,240 square sub-images, 96 sub-images of size  $1 \times 5$ , 64 sub-images of size  $5 \times 1$  and one sub-image consisting of a single pixel. Next, we can define  $O$  as the set of all sub-images, i.e.,  $O = \{f_1, \dots, f_s\}$ , and  $\mathbb{F}$  is a set of functions that operate on images (see, e.g., Table 1 for examples of probe functions used in this paper or Marti et al. (2001) for other examples). Once the set  $\mathcal{B}$  has been selected, the elementary sets are simply created by grouping all objects with the same description and the quotient set is made up of all the elementary sets. Finally, a simple example of these concepts is given in Figure 1 where the left image contains an octagon with a radius of 100 pixels located at the centre of the  $400 \times 400$  image and the right image contains the elementary sets obtained using  $\mathcal{B} = \{\phi_{\text{avg}}(f_s)\}$  and a sub-image size of  $10 \times 10$ .

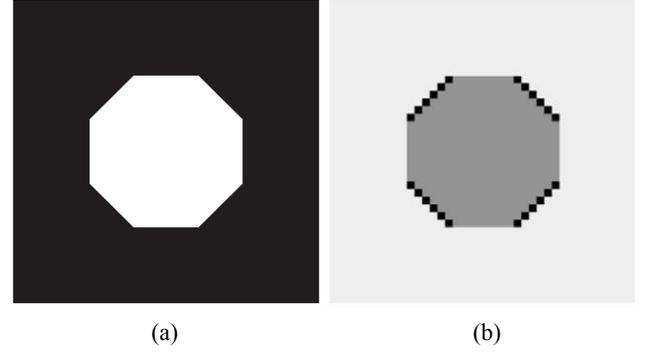
**Table 1** Example probe functions

$\phi$	Description
$\phi_{\text{avg}}(f_s)$	Average greyscale value of sub-image
$\phi_{\text{C}}(f_s)$	Information content of sub-image (Henry and Peters, 2008, 2009)
$\phi_{\text{NormR}}(f_s)$	Average normalised R value of sub-image (Marti et al., 2001)

Observe that three elementary sets are obtained in Figure 1(b), namely, the light grey background, the dark grey octagon interior and the black squares along the diagonals. The black squares are created by sub-images that contain both black and white pixels (in the original image) and are located only on the diagonals due to the sub-image size and shape, and the position and radius of the hexagon. All other sub-images are uniformly white or black. Thus, we are presented with perceptual information in the form of three equivalence classes when restricted to only being able to describe the original image with the probe function  $\mathcal{B} = \{\phi_{\text{avg}}(f_s)\}$  and a sub-image size of  $10 \times 10$ . This example clearly demonstrates that perceptual information obtained from the application of near set theory is represented by the elementary sets (formed by the grouping

of objects with similar descriptions), and the information gained is always presented with respect to the probe functions contained in  $\mathcal{B}$ .

**Figure 1** Example of near set theory in the context of image processing. (a) original image (b) elementary sets obtained from (a) using  $\phi_{\text{avg}}(f_s)$



### 3 Morphological image processing

This section introduces a new form of morphological image processing based on perception rather than geometry. Gonzalez and Woods (2002) define morphology as the study of form and structure in complex biological organisms. Similarly, they define mathematical morphology in terms of set theory where the sets consist of tuples containing pixel coordinates in a binary image. We propose a new form of morphology, inspired by traditional morphology, in which the operations are based on objects defined in near set sense. The next two subsections first introduce mathematical morphology as it is commonly presented followed by our perceptual-based approach.

#### 3.1 Mathematical morphology

The following review uses the same notation given in Gonzalez and Woods (2002). Mathematical morphology starts with the assumption that an image is a set of points. Then image points (pixels) are represented by two dimensional vectors. With this view of an image, it is then possible to consider set theory-based operations on images, namely, dilation  $\oplus$  and erosion  $\ominus$ . Let the  $\hat{\cdot}$  operator represent the reflection of a set  $B$ , defined as:

$$\hat{B} = \{w \mid w = -b, \forall b \in B\},$$

and let the operator  $(\cdot)_z$  denote the translation of a set by point  $z = (z_1, z_2)^T$ , that is,

$$(A)_z = \{c \mid c = a + z, \forall a \in A\}.$$

Then we can respectively define the dilation and erosion of an image  $A$  by a structuring element (SE)  $B$  as:

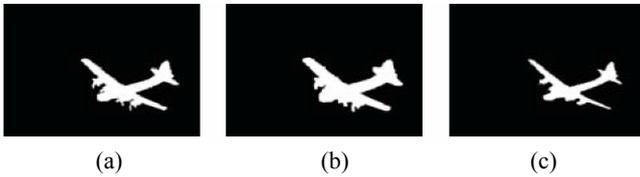
$$A \oplus B = \left\{ z \mid \left( (\hat{B})_z \cap A \neq \emptyset \right) \right\},$$

and an erosion operation  $A \ominus B$  is:

$$A \ominus B = \{z \mid (B)_z \subseteq A\},$$

where  $A$  usually indicates the image to be transformed, and the SE  $B$  consists of a geometric arrangement of pixels in which the shape and binary value of the pixels play an important role in the transformation of  $A$ . A simple example of dilation and erosion is given in Figure 2 [see both Dougherty and Lotufo (2003) and Gonzalez and Woods (2002) for more examples]. Notice that dilation and erosion operators respectively cause the set  $A$  to grow and reduce in size, where the terms grow and reduce are defined with respect to area since the objects in the sets correspond to coordinates of pixels.

**Figure 2** Example of mathematical morphology, (a) segment obtained from Berkeley segmentation dataset (Martin et al., 2001), (b) dilation of (a) using square a SE of width 5 and (c) erosion of (a) using the same SE



### 3.2 Perceptual morphology

If set theory is the language of mathematical morphology, then near set theory is language of perceptual morphology. The central idea between the two approaches is the same, i.e., both methods increase or decrease set membership based on comparison with an SE. The difference lies in the definition of the objects, in particular, the perceptual approach defines objects in the near set sense. Thus, objects can be anything as long as it is possible to define probe functions that operate on the objects under consideration. In terms of image processing, the goals are the same. Mathematical morphology is used for identifying structure (Gonzalez and Woods, 2002), whereas the near set approach is used to identify perceptual concepts in images. The latter is similar to the former in that identification of structure facilitates human perception of images (Wang et al., 2004b).

We propose to use the same approach as mathematical morphology to introduce a new morphology that the perceptual information inherent to near set theory. Again, let the set of objects be represented by  $O$  and the quotient set, as given above, is  $O / \sim_B$ , where  $B$  is a set of probe functions on objects in  $O$  selected from  $\mathbb{F}$ . Now, define a set  $A \subseteq O$  such that it has some *a priori* perceptual meaning associated with it, i.e., this set has definite meaning in a perceptual sense outside of the probe functions in  $B$ . Next, let the quotient set represent the SE from traditional mathematical morphology, in other words let  $B = O / \sim_B^{-1}$ .

As will be seen shortly, the quotient set is used as the SE in perceptual morphology, since it contains the perceptual information necessary to augment the set  $A$  in a

perceptually meaningful way. This perceptual information is in the form of elementary sets (collections of objects with the same descriptions) since, as was mentioned, we perceive objects by the features that describe them and that people tend to grasp not single objects, but classes of them (Orłowska, 1982).

Keeping the above in mind, we define perception-based dilation as:

$$A \oplus B = \{x / \sim_B \in B \mid x / \sim_B \cap A \neq \emptyset\}, \quad (2)$$

and the perception-based erosion is defined as:

$$A \ominus B = \bigcup_{x / \sim_B \in B} \{x / \sim_B \cap A\}. \quad (3)$$

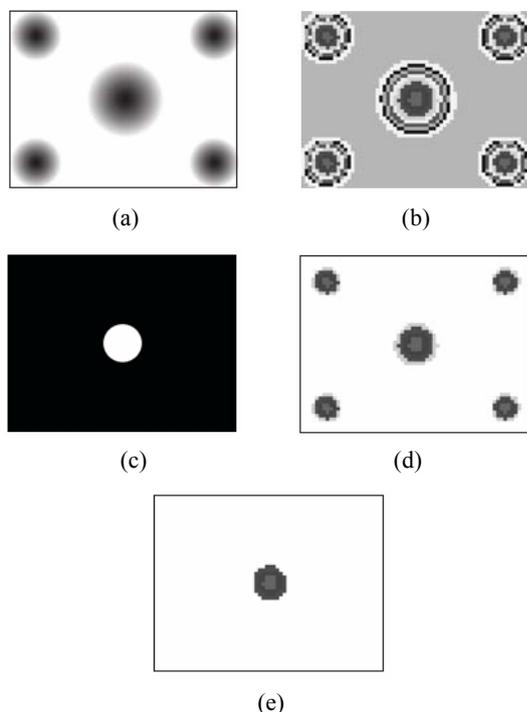
Notice that set  $A$  is grown perceptually by the SE  $B$  (and consequently by the probe functions in  $B$ ) using the dilation operator by including objects in the result that have similar descriptions to those contained in  $A$ . In other words, the dilation operation perceptually enhances the set  $A$  by including the full membership of the elementary sets that have at least one object in  $A$ . Conversely, the erosion operation essentially masks the set  $B$  (using  $A$ ) by including in the result only the portions of the elementary sets contained in  $A$  already. Thus, perceptual information can be reduced if the entire elementary set is not contained in the result.

By way of example, the above concepts are illustrated with images. Using the definition of an image and sub-image given in Section 2, let  $O$  contain sub-images as objects, i.e.,  $O = \{f_1, f_2, \dots, f_s\}$  and let  $A$  be a subset of the sub-images<sup>2</sup>. Then, the SE,  $B = O / \sim_B$ , can be viewed as an image where each class is assigned a unique colour (or grey value). Similarly, the results of dilation and erosion can be viewed as images as well where the objects in the result are assigned the same colour as the objects in  $B$  and the rest of the image can be coloured white representing an absence of objects (since not all of the sub-images in the original image will be included in the result).

Figure 3 is an example of these techniques applied to a simple greyscale image containing five circles in which a gradient operation from white to black was applied as shown in Figure 3(a). The gradient was used so that there is no crisp boundary to indicate the start of the circle. Next, Figure 3(b) contains the elementary sets created using a sub-image size of  $10 \times 10$  and  $B = \{\phi_{\text{avg}}(f_s)\}$  where each colour represents a different set. Notice that each circle has similar elementary sets since they are identical (except for the centre circle which is slightly larger). Next, Figure 3(c) contains the set  $A$  representing *a priori* perceptual information in this case it is a segment representing the point at which the centre circle is predominately black. The result of perceptually growing the segment using the dilation operator defined in equation (2) is given in Figure 3(d)<sup>3</sup>. Notice that the results show that we were able to grow the segment  $A$  to include the other four circles that

are perceptually similar with respect to the probe functions in  $\mathcal{B}$ . Thus, we were able to gain more perceptual information about the segment represented by  $A$ . Similarly, the result of perceptually reducing  $A$  using equation (3) is given in Figure 3(e). Perceptual information is reduced in that we are not taking full advantage of all the information contained in  $B = O_{/\sim_B}$  because there is no representation of the other four circles in the result which are perceptually similar to the centre circle. However, the operation is still useful in that it gives us perceptual information (in the form of elementary sets) about  $A$  with respect to the probe functions in  $\mathcal{B}$ .

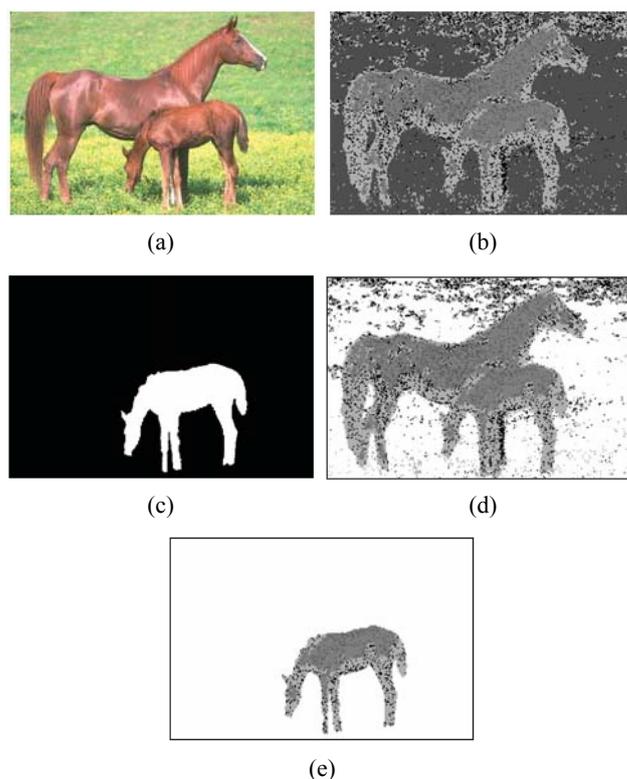
**Figure 3** Example of perceptual morphology on simple greyscale image, (a) original image, (b) quotient set of (a) created using method described in Section 2, (c) perceptual segmentation of the centre circle in (a), (d) perceptual dilation of (b) and (e) perceptual erosion of (b)



Another example of perceptual morphology is given in Figure 4. In this case, Figure 4(a) contains an image from the Berkeley segmentation dataset (Martin et al., 2001) and Figure 4(c) is a single ground truth segment from the same dataset. As before, Figure 4(b) contains the elementary sets, this time created using a sub-image of size  $2 \times 2$  and  $\mathcal{B} = \{\phi_C(f_s), \phi_{\text{NormG}}(f_s)\}$ . Notice, for the most part, that the elementary sets represent perceptual concepts of the image, e.g., the horses tend to share the same elementary sets which differ from those of the background. Next, the set  $A$  represented by the segment of the younger horse is perceptually dilated in Figure 4(d). Again, the result now includes other areas of the image that are perceptually similar (with respect to probe functions in  $\mathcal{B}$ ) including the

other horse and parts of the background. Recall that we may not find the background perceptually similar to the horse but the similarity occurs using only the probe functions in  $\mathcal{B}$ , i.e., information content and the normalised green value from the RGB colour model. Similar to the last example, the perceptual reduction caused by the erosion operator occurs due to the lack of inclusion of the full elementary sets in the result. Although, as was mentioned before, this result still contains perceptually valuable information of the original set  $A$  as will be seen in Section 4.

**Figure 4** Example of perceptual morphology on image from Berkeley segmentation dataset (Martin et al., 2001), (a) original image, (b) quotient set of (a) created using method described in Section 2, (c) perceptual segmentation of younger horse in (a), (d) perceptual dilation of (b) and (e) perceptual erosion of (b) (see online version for colours)



#### 4 Application

This section presents an application of perceptual morphology in the form of a method for segmentation evaluation called the near set index (NSI) first introduced in Henry and Peters (2008). The NSI index was created out of a need for an unsupervised segmentation evaluation method that could be used in real world systems yet still based on human perception rather than characteristics of ideal segmentations that is popular in most unsupervised segmentation evaluation methods. This section begins with a discussion of segmentation evaluation followed by an introduction of a popular segmentation technique called the mean shift algorithm (Comaniciu, 2002) which is the method used to generate segmentations in this article. Next, a supervised segmentation evaluation measure called the

normalised probabilistic rand (NPR) index (Unnikrishnan et al., 2007) is presented as a benchmark for the NSI index. Finally, a description of the NSI index is given with some examples.

#### 4.1 Segmentation evaluation

There are generally three classes of segmentation evaluation techniques, namely, analytic, empirical discrepancy and empirical goodness methods (Zhang, 1996; Zhang et al., 2005a). Analytic methods generally perform evaluation solely by examining the algorithm without considering the resultant segmentations. These methods are concerned with processing complexity and strategy and are not useful for evaluating the perceptual relevance of a given segmentation. The next category is empirical discrepancy (also called relative or supervised evaluation) and is characterised by the comparison of the test segmentation with ground truth images (segmentations performed by people or experts). These methods are popular because they evaluate segmentations based on the perceptual groupings created by people which are the end result of any segmentation algorithm. Unfortunately, it is not realistic to assume that systems incorporating image segmentation will have access to ground truth images and so these methods are generally used for the comparison of segmentation algorithms during the design phase. The last type of segmentation category is called empirical goodness (also known as standalone or unsupervised). These methods are based on some properties that ideal segmentations should contain. Generally, it is more difficult for these evaluation methods to be based on human perception of objects due to the lack of ground truth input or a formal framework for the quantification of perception and similarities of objects.

There are many examples of empirical goodness methods for segmentation evaluation. For example, Zhang et al. (2005a) implements a co-evaluation framework in which multiple unsupervised methods are combined with learning algorithms to take advantage of different measures. Examples of the measures they consider are ones based on squared colour error ratio to segment area (Borsotti et al., 1998; Liu and Yang, 1994), on entropy (information content) of an image and the minimum description length principle (MDL) (Zhang et al., 2005b), on the geometric shape of a segment (e.g., compactness, circularity and elongation), and a contrast measure between the inside and outside of a segment (Correia and Pereira, 2003). Similarly, Chabrier et al. (2004) present a review of six unsupervised methods based on image features such as segmentation contrast, standard deviation and colour error. Likewise, Zhang (1996) is another often cited survey of segmentation evaluation techniques also describing unsupervised measures where again the unsupervised methods are based solely on image characteristics. Notice that all these methods suffer because of the lack of the perceptual information contained in a ground truth image or the lack of a formal framework for the quantification of perception and similarity of objects introduced by near set theory. This is a

problem we attempt to rectify with the introduction of the NSI.

#### 4.2 Mean shift segmentation algorithm

The mean shift algorithm, introduced in Comaniciu (2002), is an image segmentation algorithm that creates segments based on the assumption that the image can be represented by a mixture model of multivariate density functions. The basic idea is that for each pixel the mean shift algorithm iteratively searches for a mode (peak) in the local density. Then, a pixel is assigned to the region for which all pixels have the same mode (peak) (Wang et al., 2004a). The process of finding the modes for an image is based on theory from the field of kernel density estimation. This is a non-parametric technique for estimating the probability density function of a random variable based on observations. Specifically, both the number of observations within a volume in  $d$ -dimensional space centred on  $\mathbf{x}$  and a kernel that weights the importance of the observations determines the estimate of the distribution (Duda et al., 2001). The segmentations used in this article were created using the implementation of the mean shift algorithm called EDISON (Christoudias et al., 2002), a system for which both the source code and binaries are freely available online.

#### 4.3 Normalised probabilistic rand index

The NPR index [introduced in Unnikrishnan et al. (2007) and summarised here] is a non-parametric technique for evaluating the performance of an image segmentation algorithm. This index is a supervised technique in that evaluation is performed with respect to ground truth images. A supervised measure was selected as the benchmark for the NSI index since human perceptual grouping is inherent to the evaluation of the segmentation due to the use of ground truth images. Furthermore, the NPR index was selected due to its use of multiple ground truth images when evaluating a proposed segmentation thus taking into account multiple perceptual sources.

The NPR index has its roots in the rand index, a measure developed based on the idea of counting pixel pairs that have the same segmentation labels. First, define an image  $X = \{x_1, \dots, x_N\}$  of  $N$  pixels and two segmentations of  $X$ ,  $S$  and  $S'$  where each segmentation respectively assigns labels  $l_i$  and  $l'_i$  to the pixels in  $X$ . Then the rand index is given as:

$$R(S, S') = \frac{1}{\binom{N}{2}} \sum_{\substack{i,j \\ i \neq j}} [\Pi(l_i = l_j \wedge l'_i = l'_j) + \Pi(l_i \neq l_j \wedge l'_i \neq l'_j)],$$

where  $\Pi$  is the identity function and  $\binom{N}{2}$  is the number of unique pixel pairs in  $X$ . Next, the rand index was extended to allow the use of more ground truth images in the evaluation of a proposed segmentation. The idea is that observing the same pixel pair in each ground truth image is considered a Bernoulli trial with the two outcomes being

either they have the same label or they do not. Then, the set of observations across all ground truth segmentations forms a Bernoulli distribution with expected value  $p_{ij}$ . Thus, given a set of manual segmentations  $\{S_1, \dots, S_K\}$ , a segment for evaluation  $S_{\text{test}}$  and a label of  $x_i$  denoted as  $l_i^S$  where  $S$  denotes the segment used to label  $x_i$ , the probabilistic rand (PR) index is defined as:

$$PR(S_{\text{test}}, \{S_k\}) = \frac{1}{\binom{N}{2}} \sum_{\substack{i,j \\ i < j}} [c_{ij} p_{ij} + (1 - c_{ij})(1 - p_{ij})],$$

where

$$c_{ij} = \Pi(l_i^{S_{\text{test}}} = l_j^{S_{\text{test}}}).$$

Finally, the NPR index extends the PR index by normalising with respect to its baseline. The selected baseline is the expected value of PR index. Consequently, the NPR index is defined as:

$$\text{NPR} = \frac{\text{PR} - \mathbb{E}[\text{PR}]}{\max(\text{PR}) - \mathbb{E}[\text{PR}]},$$

where the maximum value is taken to be 1 and the expected value of the PR index is given as:

$$\begin{aligned} \mathbb{E}[PR(S_{\text{test}}, \{S_k\})] &= \frac{1}{\binom{N}{2}} \sum_{\substack{i,j \\ i < j}} \left\{ \mathbb{E}[\Pi(l_i^{S_{\text{test}}} = l_j^{S_{\text{test}}})] p_{ij} \right. \\ &\quad \left. + \mathbb{E}[\Pi(l_i^{S_{\text{test}}} \neq l_j^{S_{\text{test}}})] (1 - p_{ij}) \right\}, \\ &= \frac{1}{\binom{N}{2}} \sum_{\substack{i,j \\ i < j}} [p'_{ij} p_{ij} + (1 - p'_{ij})(1 - p_{ij})], \end{aligned}$$

where  $p'_{ij} = \mathbb{E}[\Pi(l_i^{S_{\text{test}}} = l_j^{S_{\text{test}}})]$ . To make the baseline representative of perceptually consistent groupings,  $p'_{ij}$  is estimated from segmentations of all images for all unordered pairs. In other words, given  $\Phi$  as the number of images in the database used for testing,  $p'_{ij}$  is defined as:

$$p'_{ij} = \frac{1}{\Phi} \sum_{\phi} \frac{1}{K_{\phi}} \sum_{k=1}^{K_{\phi}} \Pi(l_i^{S_k^{\phi}} = l_j^{S_k^{\phi}}).$$

#### 4.4 Near set index

This section introduces a method for segmentation evaluation using perceptual morphology presented above. The goal of image segmentation is to partition an image into disjoint regions such that each one reflects some perceptual component of the image<sup>4</sup>. Since it has been observed that the quotient set captures perceptual information of objects (Orłowska, 1985), it makes sense to use the quotient set to measure the quality of a segmentation, i.e., the degree to which a segmentation represents an image component.

Let  $f$  represent an RGB image and let  $A$  represent *a priori* information in the form of an image segment. Then, the result of perceptual erosion can be used to evaluate the quality of the segment since it only contains perceptual information about the set  $A$ . Further, since this set should represent a perceptual component within  $f$  and the quotient set represents perceptual information about the sub-images in  $O$ , it should be possible to select probe functions in  $\mathcal{B}$  such that the elementary sets begin to represent these components. As such, we propose that a good measure of segmentation quality is the variability of the classes contained in the perceptual erosion of  $A$ . In general, the measure of variability of objects that take on labels from a discrete set is called the information content and it takes on values in the interval  $[0, \log_2 L]$  where  $L$  is the number of different labels the objects can assume (MacKay, 2003; Seemann, 2002). A value of 0 is produced when the objects contain all the same labels and the highest value occurs when each label occurs with equal frequency. Thus, for this application, low value of information content of the erosion of  $A$  corresponds to good segmentations and vice versa. This leads to the following definition:

*Definition 2. NSI:* Let  $A$  represent a single image segment for evaluation, and let  $B = O_{/\mathcal{B}}$  represent the quotient set obtained using the probe functions in  $\mathcal{B}$ . Then, the NSI is the information content of the perception-based erosion of  $A$ .

#### 4.5 Shannon and Pal entropy

Shannon's entropy (also called information content) is a measure of the information gained in a system by receiving a message from a finite set of messages, and is very useful in image processing. Work in Pal and Pal (1991, 1992) shows that Shannon's definition of entropy has some limitations. The following reviews Shannon's definition of entropy and presents exponential entropy introduced in Pal and Pal (1991).

Let the probability of receiving a message  $i$  of  $n$  messages be  $p_i$ , then the information gain of a message can be written as:

$$\Delta I = \log(1/p_i) = -\log(p_i), \quad (4)$$

and the entropy of the system is the expected value of the gain and is calculated as:

$$H = -\sum_{i=1}^n p_i \log(p_i).$$

Shannon's definition of entropy suffers from the following problems: it is undefined when  $p_i = 0$ ; in practise the information gain tends to lie at the limits of the interval  $[0, 1]$ ; and statistically speaking, a better measure of ignorance is  $1 - p_i$  rather than  $1/p_i$  (Pal and Pal, 1991). As

a result, a new definition of entropy can be defined with the following desirable properties:

- P1  $\Delta I(p_i)$  is defined at all points in  $[0, 1]$
- P2  $\lim_{p_i \rightarrow 0} \Delta I(p_i) = \Delta I(p_i = 0) = k_1, k_1 > 0$  and finite
- P3  $\lim_{p_i \rightarrow 1} \Delta I(p_i) = \Delta I(p_i = 1) = k_2, k_2 > 0$  and finite
- P4  $k_2 < k_1$
- P5 With increase in  $p_i, \Delta I(p_i)$  decreases exponentially
- P6  $\Delta I(p)$  and  $H$ , the entropy, are continuous for  $0 \leq p \leq 1$
- P7  $H$  is maximum when all  $p_i$ 's are equal, i.e.,  $H(p_1, \dots, p_n) \leq H(1/n, \dots, 1/n)$ .

With these in mind, Pal and Pal (1991) define the gain in information from an event as:

$$\Delta I(p_i) = e^{(1-p_i)},$$

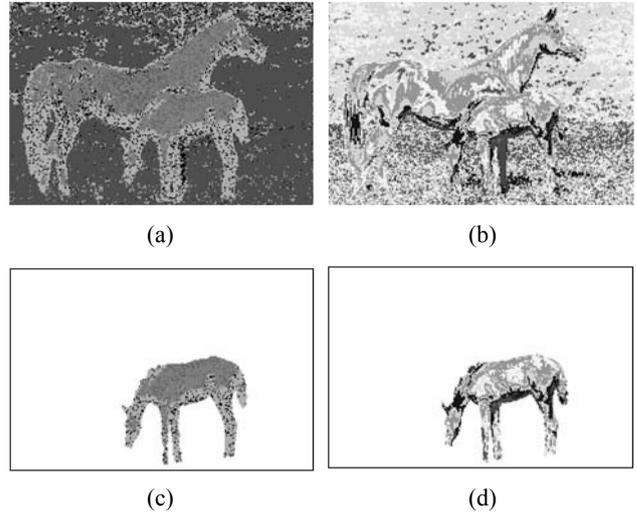
which gives a new measure of entropy as:

$$H = \sum_{i=1}^n p_i e^{(1-p_i)}.$$

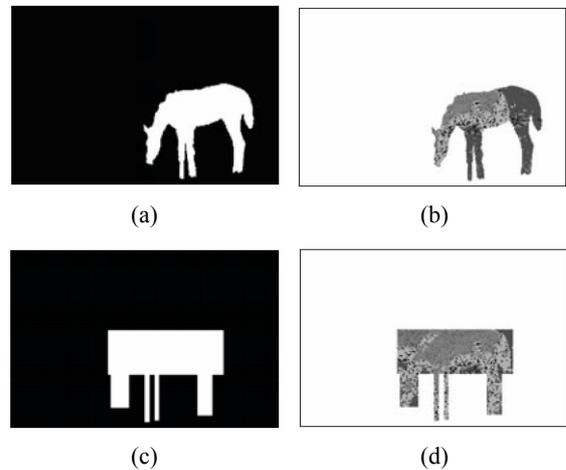
#### 4.6 Segmentation evaluation examples

To demonstrate these concepts two examples are given in Figures 5 and 6 and the resultant information content is given in Table 2. The first example shows that poor choices of  $\mathcal{B}$  lead to poor segmentation evaluations. This demonstrates that the method is dependent on the selection of probe functions in the same manner that pattern recognition is dependent on feature selection. Notice that the elementary sets of Figure 5(a) tend to capture the perceptual components of Figure 4(a) better than those in Figure 5(b). Consequently, the information content of the erosion given in Figure 5(d) is higher (and so worse) than that of Figure 5(c). Thus, the same segment can have different NSI values depending on the choice of  $\mathcal{B}$ . This example was given to highlight the need for careful probe function selection for a given application. The next example (given in Figure 6) demonstrates the ability of the NSI to evaluate different segmentations. The segmentation given in Figure 6(a) is the same as Figure 4(c) shifted to the right, and the segmentation given in Figure 6(c) was created by placing boxes of approximately the right size over the horse. Both of these segmentations are bad in the sense that they do not capture the perceptual component representing the smaller horse in the image as well as the segmentation in Figure 4(c). This is reflected by the information content values given in Table 2 which are higher than that of Figure 5(c).

**Figure 5** Example of quotient set and perception-based erosion of Figure 4(a) using different probe functions, (a) Figure 4(b) created using probe functions  $\mathcal{B} = \{\phi_C(f_s), \phi_{\text{NormG}}(f_s)\}$  and repeated here for comparison, (b) quotient set obtained using  $\mathcal{B} = \{\phi_{\text{avg}}(f_s)\}$ , (c) perceptual-based erosion using (a) as the SE and (d) perceptual-based erosion using (b) as the SE



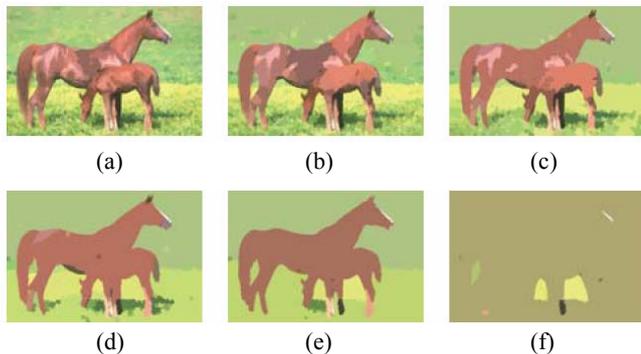
**Figure 6** Segmentations of the smaller horse in Figure 4(a) and their perception-based erosions, (a) sample segmentation, (b) erosion of (a), (c) sample segmentation and (d) erosion of (b)



**Table 2** Information content of perceptual-based erosion

Erosion image	Information content
Figure 5(c)	1.6619
Figure 5(d)	2.5682
Figure 6(b)	2.1954
Figure 6(d)	2.0797

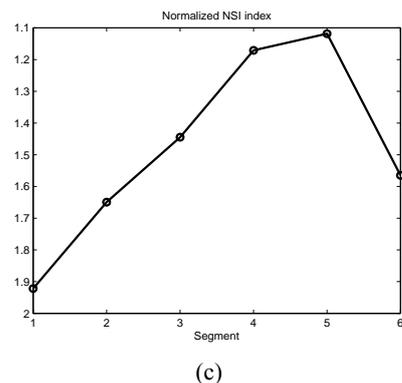
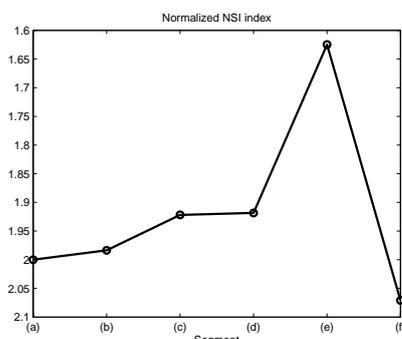
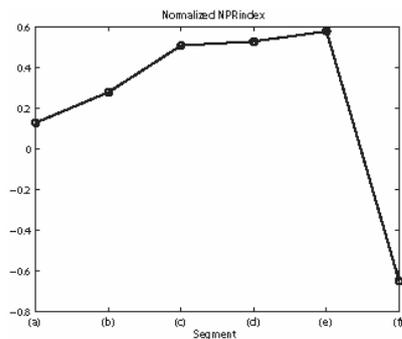
**Figure 7** Mean shift segmentation of Figure 4(a) with  $h_s(\text{spatial}) = 7$  and  $h_r(\text{range}) = 3, 7, 11, 15, 19, 23$  for (a), (b), (c), (d), (e) and (f) (see online version for colours)



**Figure 8** Results of evaluating the segmentations given in Figure 7, (a) NPR index and (b), (c) NSI on a window size of  $2 \times 2$  using respectively

$$\mathcal{B} = \left\{ \phi_{H_{\text{Shannon}}}(f_s), \phi_{\text{NormG}}(f_s) \right\} \text{ and}$$

$$\mathcal{B} = \left\{ \phi_{H_{\text{Pal}}}(f_s), \phi_{\text{NormG}}(f_s) \right\}$$



Next, further demonstration of the ability of the NSI index to evaluate an image segmentation,  $S_{\text{test}}$ , is given by way of comparison with the NPR index. But first, the NSI must be extended to handle more than one segment. This is easily accomplished by letting the NSI of a proposed image segmentation be the information content of the worst segment in the image, i.e., let the NSI be the value of the highest information content resulting from the erosions of all the segments in  $S_{\text{test}}$ . The idea being that a proposed segmentation is only as good as its worst region. Using this approach, the NSI was used to evaluate the segmentations from an example given in Unnikrishnan et al. (2007) and repeated in Figure 7. The results of evaluating the segmentations in Figure 7 using both the NPR and the NSI index are reported in Figure 8. Notice that all three measures give similar evaluations of the segmentations of Figure 7, i.e., Figures 7(c)–7(e) are all ranked the highest [with Figure 7(e) ranked the best]. Additionally, the NSI index gives similar (poor) ratings for both the over and under segmentations shown in Figures 7(a) and 7(f), rather than treating one case much worse than the other (a rather nice result since both are usually equally bad outcomes).

### 5 Conclusions

This article presents an introduction to a form of an unsupervised segmentation evaluation measure called the NSI. This index is based on a new form of morphology inspired by traditional morphology using near set theory to incorporate perceptual information. Initial results obtained by comparison with the NPR index demonstrate that the NSI may be a reasonable measure for unsupervised segmentation evaluation. Further, this new method has an advantage over other existing unsupervised methods because it takes into account perceptual information in the form of elementary sets. It is important to note that this method is dependent on probe function selection and that poor selection results in less perceptually relevant information. This is to be expected and is similar to the feature extraction problem in pattern recognition. Also of note, perceptual information is always presented with respect to the probe functions contained in  $\mathcal{B}$  just as our senses define our perception of the world. For example, our ability to view light in the ‘visible spectrum’ rather than infra red or microwaves spectrum defines our perception of the world just as the selection of probe functions constrains the amount of perceptual information available for extraction from a set of objects. Future work will consist of in-depth comparison testing of the NSI with the NPR to further demonstrate the validity of the NSI as an unsupervised segmentation evaluation technique.

### Acknowledgements

The authors would like to thank Sankar K. Pal, Piotr Wasilewski, Andrzej Skowron, Jarislaw Stepanski, Sheela Ramanna, Homa Fashandi, Amir H. Meghdadi,

Dan Lockery and Leszek Puzio for their insights and suggestions concerning topics in this paper. This research has been supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) grant 185986, Canadian Arthritis Network grant SRI-BIO-05, Manitoba Hydro grant T277 and an NSERC Postgraduate Fellowship.

## References

- Borsotti, M., Campadelli, P. and Schettini, R. (1998) 'Quantitative evaluation of color image segmentation results', *Pattern Recognition Letters*, Vol. 19, pp.741–747.
- Bruce, V., Green, P.R. and Georgeson, M.A. (1996) *Visual Perception: Physiology, Psychology and Ecology*, Psychology Press, Hove, East Sussex, UK.
- Chabrier, S., Emile, B., Laurent, H., Rosenberger, C. and Marche, P. (2004) 'Unsupervised evaluation of image segmentation application to multi-spectral images', in *Proceedings of the 17th International Conference on Pattern Recognition (ICPR 2004)*, Vol. 1, pp.576–579.
- Christoudias, C., Georgescu, B. and Meer, P. (2002) 'Synergism in low level vision', in *Proceedings of the 16th International Conference on Pattern Recognition*, Quebec City, Vol. 4, pp.150–156.
- Comaniciu, D. (2002) 'Mean shift: a robust approach toward feature space analysis', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 5, pp.603–619.
- Correia, P.L. and Pereira, F. (2003) 'Objective evaluation of video segmentation quality', *IEEE Transactions on Image Processing*, Vol. 12, No. 2, pp.156–200.
- Dougherty, E.R. and Lotufo, R.A. (2003) *Hands-on Morphological Image Processing*, Vol. TT59, SPIE Press.
- Duda, R., Hart, P. and Stork, D. (2001) *Pattern Classification*, 2nd ed., Wiley.
- Gonzalez, R.C. and Woods, R.E. (2002) *Digital Image Processing*, 2nd ed., Prentice-Hall, Toronto.
- Gupta, S. and Patnaik, K. (2008) 'Enhancing performance of face recognition systems by using near set approach for selecting facial features', *Journal of Theoretical and Applied Information Technology*, Vol. 4, No. 5, pp.433–441.
- Hassanien, A.E., Abraham, A., Peters, J.F., Schaefer, G. and Henry, C. (2009) 'Rough sets and near sets in medical imaging: a review', *IEEE Transactions on Information Technology in Biomedicine*, digital object identifier: 10.1109/TITB.2009.2017017, in press.
- Henry, C. and Peters, J.F. (2007) 'Image pattern recognition using approximation spaces and near sets', in *Proceedings of the Eleventh International Conference on Rough Sets, Fuzzy Sets, Data Mining and Granular Computer (RSFDGrC 2007), Joint Rough Set Symposium (JRS07)*, Lecture notes in artificial intelligence, Vol. 4482, pp.475–482.
- Henry, C. and Peters, J.F. (2008) 'Near set index in an objective image segmentation evaluation framework', in *Proceedings of the GEOgraphic Object Based Image Analysis: Pixels, Objects, Intelligence*, University of Calgary, Alberta, pp.1–6.
- Henry, C. and Peters, J.F. (2009) *International Journal of Intelligent Computing and Cybernetics*, submitted.
- Liu, J. and Yang, Y.H. (1994) 'Multi-resolution color image segmentation', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 16, No. 7, pp.689–700.
- MacKay, D.J.C. (2003) *Information Theory, Inference, and Learning Algorithms*, Cambridge University Press, UK.
- Marti, J., Freixenet, J., Batlle, J. and Casals, A. (2001) 'A new approach to outdoor scene description based on learning and top-down segmentation', *Image and Vision Computing*, Vol. 19, pp.1041–1055.
- Martin, D., Fowlkes, C., Tal, D. and Malik, J. (2001) 'A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics', in *Proceedings of the 8th International Conference on Computer Vision*, Vol. 2, pp.416–423.
- Orłowska, E. (1982) *Semantics of Vague Concepts, Applications of Rough Sets*, Tech. Rep. 469, Institute for Computer Science, Polish Academy of Sciences.
- Orłowska, E. (1985) 'Semantics of vague concepts', in G. Dorn and P. Weingartner (Eds.): *Foundations of Logic and Linguistics. Problems and Solutions*, pp.465–482, Plenum Press, London/NY.
- Pal, N.R. and Pal, S.K. (1991) 'Entropy: a new definition and its applications', *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 21, No. 5, pp.1260–1270.
- Pal, N.R. and Pal, S.K. (1992) 'Some properties of the exponential entropy', *Information Sciences*, Vol. 66, pp.119–137.
- Pawlak, Z. (1981) *Classification of Objects by Means of Attributes*, Tech. Rep. PAS 429, Institute for Computer Science, Polish Academy of Sciences.
- Pawlak, Z. and Peters, J.F. (2002) 'Jak blisko (how near)', *Systemy Wspomagania Decyzji*, Vol. 1, pp.57, 109.
- Peters, J.F. (2007a) 'Classification of objects by means of features', in *Proceedings of the IEEE Symposium Series on Foundations of Computational Intelligence (IEEE SSCI 2007)*, Honolulu, Hawaii, pp.1–8.
- Peters, J.F. (2007b) 'Near sets: general theory about nearness of objects', *Applied Mathematical Sciences*, Vol. 1, No. 53, pp.2609–2629.
- Peters, J. (2007c) 'Near sets: special theory about nearness of objects', *Fundamenta Informaticae*, Vol. 75, Nos. 1–4, pp.407–433.
- Peters, J.F. (2008a) 'Classification of perceptual objects by means of features', *International Journal of Information Technology & Intelligent Computing*, Vol. 3, No. 2, pp.1–35.
- Peters, J.F. (2008b) 'Discovery of perceptually near information granules', in J.T. Yao (Ed.): *Novel Developments in Granular Computing: Applications of Advanced Human Reasoning and Soft Computation*, Information Science Reference, Hersey, N.Y., USA, to appear.
- Peters, J.F. (2009a) 'Affinities between perceptual granules: foundations and perspectives', in A. Bargiela and W. Redrycz (Eds.): *Human-Centric Information Processing Through Granular Modelling*, pp.49–66, Springer-Verlag, Berlin.
- Peters, J.F. (2009b) 'Fuzzy sets, near sets, and rough sets for your computational intelligence toolbox', in A.E. Hassanien, A. Abraham, F. Herrera, W. Pedrycz, A. de Carvalho, P. Siarry and A.P. Engelbrecht (Eds.): *Studies in Computational Intelligence*, Springer, Heidelberg, Vol. 202, Springer, Heidelberg, ISBN: 978-3-642-01532-8.
- Peters, J. (2009c) 'Tolerance near sets and image correspondence', *Int. J. Bio-Inspired Computation*, Vol. 1, No. 4, pp.239–245.
- Peters, J.F. and Wasilewski, P. (2009) *Foundations of Near Sets. An International Journal*, Vol. 179, pp.3091–3109, digital object identifier: doi:10.1016/j.ins.2009.04.018, in press.

- Peters, J.F., Skowron, A. and Stepaniuk, J. (2007) 'Nearness of objects: extension of approximation space model', *Fundamenta Informaticae*, Vol. 79, Nos. 3–4, pp.497–512.
- Seemann, T. (2002) 'Digital image processing using local segmentation', Phd dissertation, School of Computer Science and Software Engineering, Monash University.
- Serra, J. (1982) *Image Analysis and Mathematical Morphology*, Academic Pres., London.
- Skowron, A. and Peters, J.F. (2008) 'Rough-granular computing', in V.K.W. Pedrycz and A. Skowron (Eds.): *Handbook on Granular Computing*, pp.285–328, John Wiley & Sons, Inc., NY, USA.
- Unnikrishnan, R., Pantofaru, C. and Hebert, M. (2007) 'Toward objective evaluation of image segmentation algorithms', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 29, No. 6, pp.929–944.
- Wang, J., Thiesson, B., Xu, Y. and Cohen, M. (2004a) 'Image and video segmentation by anisotropic kernel mean shift', in *In. Proc. European Conference on Computer Vision (ECCV'04)*, Vol. 2, pp.238–249.
- Wang, Z., Bovik, A.C., Sheikh, H.R. and Simoncelli, E.P. (2004b) 'Image quality assesment: from error visibility to structural similarity', *IEEE Transactions on Image Processing*, Vol. 13, No. 4, pp.600–612.
- Zhang, H., Fritts, J.E. and Goldman, S.A. (2005a) 'A co-evaluation framework for improving segmentation evaluation', in *SPIE Defense and Security Symposium – Signal Processing, Sensor Fusion, and Target Recognition XIV*, pp.420–430.
- Zhang, H., Fritts, J.E. and Goldman, S.A. (2005b) 'A fast texture feature extraction method for region-based image segmentation', in *Proceedings of IS&T/SPIE's 16th Annual Symposium on Image and Video Communication and Processing*, Vol. 5685.
- Zhang, Y.J. (1996) 'A survey on evaluation methods for image segmentation', *Pattern Recognition*, Vol. 29, No. 8, pp.1335–1346.

## Notes

- 1 The quotient set is being relabelled only to be notationally consistent with traditional mathematical morphology defined in Subsection 3.1.
- 2 Sometimes a set  $A$  will be given in terms of a segment of the original image [as in Figure 3(c)]. In this case it is necessary to rasterise the set by converting from a set with pixel granulation to a set with sub-image granulation.
- 3 Note that the white area of Figures 3(d) and 3(e) and Figures 4(d) and 4(e) do not represent an elementary set. This area is meant to be a background used to indicate a lack of elementary sets that are found in the same location in the image of the quotient set.
- 4 Here we refer to the *objects* contained in images as components to avoid confusion with *objects* in the near set sense.