# ILL-CONFIGURED OBJECT REPRESENTATION BY NEIGHBOUR SET WITH APPLICATIONS TO AERIAL IMAGE ANALYSES

T. Watanabe<sup>a, \*</sup>, K. Sugawara<sup>a</sup>

<sup>a</sup> Graduate School of Information Systems, University of Electro-Communications, 1-5-1 Chofu-ga-oka, Chofu-shi, Tokyo, 182-8585 Japan – (watanabe, sugawara)@sd.is.uec.ac.jp

KEY WORDS: Object, Configuration, Model, Theory, Algorithm, Recognition, Registration, Image

## **ABSTRACT:**

We introduce herein the concept of the ill-configured object (ICO). An ICO is a geometrical object having a stable (unique) name but varying configurations (shape, size, components, and component layout). In addition, we introduce the concept of the neighbour set representation (NSR) of an object, and show that the NSR is well-fitted to the ICO. Moreover, we show that any object, either non-ICO or ICO, can be characterized as a solution of a set theoretic equation defined on its NSR. An algorithm is thus designed to detect ICOs in images. Two applications of the proposed theory are then presented. The first is ICO recognition in aerial images, and the second is automatic matching of highly deviated landmark-less images. The latter provides a foundation for automatic land cover change analysis using satellite/aerial images obtained under different conditions (time, height, and direction).

### 1. INTRODUCTION

Although number of objects of concern to us may have specific names, their configurations may vary. The shape, the components, and component layout may change depending on the case. For example, an aerial image of school is usually composed of a number of components, such as a school building, a playground, and a pool. However, the overall size, shape, components, and their layout vary, as shown in Figure 1. In this paper, this type of object is referred to as an ill-configured object (ICO). In contrast, an object that has a very stable configuration, such as a human face, is an example of a non-ICO. The problem of ICO recognition in segmented images is examined herein. We will skip discussions on aerial image segmentation and refer the reader to our recent paper (Watanabe, 2002). The solution of this problem will contribute to various image analysis tasks that require object recognition, especially tasks that require inexact matching (Shapiro, 1981).

In order to solve this problem, we should prepare an appropriate computational representation of ICO. A typical representation is a graph model (equivalently, a relational model) in which nodes represent component regions and arcs represent region adjacency relations (Shapiro, 1981; Vosselman, 1992; Kim, 1991). Object recognition then becomes the problem of finding a sub-graph (of the larger graph representing the whole image) that exactly matches the model graph. In order to cope with ICOs in this setting, we must prepare various graph models corresponding to the topological variants of the ICOs. Using the inexact matching method, which finds a sub-graph that is similar to a given model, we can reduce the model set size at the expense of high computation cost (Shapiro, 1981; Shapiro, 1982; Shapiro, 1985; Vosselman, 1992).

A powerful algorithm for inexact matching of trees has recently been developed, however, this algorithm is as of yet inapplicable for general graphs (Oommen, 2001). In addition, non-graphical representations of objects, such as MRF (Markov random field) (Geman, 1984), attribute grammar (Young, 1986), logical rules (Ohta, 1985), and 2D string (Chang, 1987), also exist. However, in order to deal with ICOs, these representations So, to solve the ICO recognition problem, we are required to go back to the basics and investigate the possibility of a new representation scheme for ICO. We introduce herein the concept of neighbor set representation (NSR) of objects as a possible solution and investigate its properties. We show that the number of models required for ICO representation is far fewer than for the graph models. We show that NSR is a unified representation for both non-ICOs and ICOs by proving that both objects can be characterized as a solution, i.e., a fixed point, of a set theoretic formula defined on the NSR of the object. Fortunately, this formula permits an iterative solution on which we can build an ICO search algorithm.

We present two applications of aerial image analysis in order to demonstrate the usefulness of the proposed concepts: ICO recognition in artificial and real images (Suto, 2000), and

also have high cost, with respect to either model preparation and/or in computation. Therefore, a new representation and a new matching method are required in order to solve the ICO recognition problem. Recently, a new method ACC (adaptive combination of classifiers) is proposed to deal with a kind of ICO having a non-fixed but stable component layout (Mohan, 2001). It's typical target is the articulated human body composed of components including, head, right/left arums, body, and foots. ACC is composed of several low level component classifiers and a high level combination classifier. Using the fact that both the place and the extent of each component is stable, each component classifier monitors the existence of a relevant component in a relevant window and the combination classifier decides the existence of a human body using the outputs of these component classifiers. In both layers, classifiers are realized using SVM (support vector machine) (Vapnik, 1998). Although superior performance than the traditional non component-based complete person detector is reported, very high SVM training cost of nearly  $O(10^3)$  positive and  $O(10^4)$  negative examples are required for each classifier. So, ACC looses its power for ICOs that have unstable components configuration and/or permits only small training examples as seen in aerial image.

<sup>\*</sup> Corresponding author.

automatic matching of highly deviated landmark-less aerial images (Nishikawa, 2001), which provides a fundamental function in realizing automated land cover change monitoring. In the present paper, basic concepts are discussed in section 2 and applications are presented in section 3. For the sake of simplicity, primary discussions are limited to two-dimensional and two-level hierarchical objects, although extensions to more general objects are possible.





(a) School; (b) segmented image; (c) and (d) layout change; (e) layout and component change; and (f) size change. An example representation of b is,

O.attrib = school, O.area = (x, y, w, h), O.area.size = (w, h),

*O.area.base* = (x, y), *O.subs* =  $\{B, G, W\}$ . We should prepare different *O.layout* for each of b, c, d, and e. For f, *O.area.size* should be changed also.

#### 2. BASIC CONCEPTS

## 2.1 Ill-configured Object

Assume O denotes a geometrical object contained in an image I which can be *approximately* represented by a tuple:

 $O \equiv (O. attrib, O. area, O. subs, O. layout),$ where O. attrib is the attribute, such as the name, of O.O. area = (x, y, w, h) denotes the minimum bounding rectangle (MBR) of O (Samet, 1993). Occasionally, we use O. area. sizeand O. area. base to represent the size (w, h) and the base point (x, y) (north-west corner) of the MBR in I, respectively. *O.subs* denotes the set of components (sub-objects) of *O. O.layout* is the layout of *O.subs*.

In the case in which *O* is a terminal object (having no subobjects), *O.subs* = *O.layout* =  $\phi$ , where  $\phi$  denotes a null set. As we are concentrating on two-level hierarchical objects, *O.subs* are composed of terminal objects. We refer to *O* as wellconfigured (or not ill-configured) if *O* can be represented by a small set of such models. For example, trademarks and human faces are well-configured objects. However, as shown in Figure 1, several objects of concern to us are not wellconfigured, and these objects are referred to as ill-configured object (ICO). As discussed before, traditional representations such as graphs and others are not useful for ICOs because the number of models required in representing an ICO becomes very large (as shown later) and/or the cost of model application (matching) to real data becomes very large.



Figure 2. Neighbour set representation (NSR)

(a) School image; (b), (c), and (d) neighbor set of the building, water, and ground, respectively, when their placements are fixed (non-ICO); (e), (f), (g) neighbor set when the building is fixed at left but other components can move (ICO); (h), (i) and (j) neighbor set when all three components can move freely (ICO). In this example, model size (=|O.subs|) = 3.

## 2.2 Neighbor Set Representation

First, we introduce the neighbor set representation of an object O, denoted as NSR(O).

**Definition 1:** *NSR(O)*, the neighbor set representation of an object of *O*, is defined as follows:

 $NSR(O) \equiv (O. attrib, O. area, O. subs, O. nset),$ 

Note that, the term O.layout in the original definition of O is changed to O.nset. O.nset is a neighbor set of O. In other

words, *O.nset* is a set of MBRs, each element of which represents the possible extension of *O* around each child element in *O.subs*. See Figure 2. More specifically,

$$O.nset = \{MBR^{*}(c, O.area) \mid c \in O.subs\},\$$

where  $MBR^*(c, O. area)$  is a union of all the possible

extensions of O.area around c.

For NSR to be an effective representation scheme of ICO, the following properties must be satisfied.

(1) The size of NSR(O) for ICO representation is not so large.

(2) Computationally efficient ICO search is possible.

In the following, Theorem 1 assures (1). Theorem 2 and 3 give fundamental properties of NSR upon which we can assure (2). Proofs are omitted due to page limitation.



Figure 3. *O. area* is always identical to the core of *O. nset* (a) Non-ICO (*O. nset* is Fig.2: b, c, d); (b) and (c) ICOs having partially fixed components (*O. nset* is Fig.2: e, f, g); (d), (e), (f) and (g) ICOs having no fixed components (*O. nset* is Fig.2: h, i, j).

**Theorem 1:** The size of *O.nset* which must be specified in order to represent an ICO is O(|O.subs|) where |O.subs| denotes the size of *O.subs*.

**Discussion 1:** See Figure 2. It should be noted that, if we resort to a relational model (see Figure 1) to represent all the possible layouts of ICO, we are forced to prepare O(|O.subs|!) models in

*O.layout*. In order to verify this, select an object *O* having *O.area.size* = (nw, mh) and |O.subs| = nm (each element of *O.subs* has different attribute but are identically shaped, i.e., width = w, height = h, and  $w \neq h$ ). Further suppose that r of mn elements are located at fixed positions and that the remaining mn - r elements can be placed at any of the mn - r positions. Then, the number of possible configurations becomes O((|O.subs| - r)!), or O(|O.subs|!).

**Theorem 2:** Any object *O* satisfies the following set theoretic equation:

 $O.area = \bigcap O.nset (\equiv \bigcap_{c \in O.subs} MBR^* (c, O.area)).$ 

We refer to the right-hand side of this equation as the *core* of *O.nset*.

**Discussion 2:** Theorem 2 states that *O.area* is identical to the core of *O* for both non-ICOs and ICOs. See Figure 3 for examples.

**Theorem 3:** Let  $MBR^{**}(c, O.area)$  be an expansion of  $MBR^{*}(c, O.area)$  satisfying

 $MBR^*(c, O.area) \subset MBR^{**}(c, O.area)$ , then we obtain

 $O.area \subset \bigcap_{c \in O.subs} MBR^{**}(c, O.area)).$ 

**Discussions 3:** Theorem 3 states that we can enclose the *O.area* in the core of the enlarged *O.nset*. See Figure 4 for some examples.

#### 2.3 ICO Recognition using NSR

In this section we discuss the method of recognizing ICO using the above definitions and Theorems. Since real-world ICOs have a number of variations, we discuss methods for dealing with these variations.

**The Problem:** Let *I* and *SEG(I)* denote an image and its segmented version, respectively. We assume each segment in *SEG(I)* is approximately represented by an MBR. Thus, *SEG(I)* is a set of records of the form  $\{(s.attriv, s.area) | s \in SG\}$ , where *SG* denotes the set of segments in *I. s.attriv* is the attribute (e.g., water, ground, house, etc.) of the segment *s*, and *s.area* = (*x*, *y*, *w*, *h*), where (*x*, *y*) and (*w*, *h*) denote the base point (e.g., northwest corner) and the size (e.g., width and height) of the MBR of *s* in *I*, respectively.

The goal is to find task objects TO in I.

The model set (MS) for *TO* is assumed to be provided using the NSR scheme. Thus,

 $MS = \{(O.attrib, O.area, O.subs, O.nset) | O \in TO\}.$ 

Notice that in *MS*, the base point parameters (x, y) in *O.area* are undefined, because they can not be determined until the model is instantiated (matched to an object) in *SEG(1)*.



Figure 4: Effect of expanding the neighbor set

(a), (b) and (c) Expanded neighbour set (dotted line); (d) the core of the expanded neighbour set fully enclose the fastened layout of O.subs; (e) and (f) the expanded core can also contain unfastened layout of O.subs.

**Finding Objects:** An object finding function can be designed using Theorem 2. Suppose we seek an object *O*. Theorem 2 states that we can determine *O.area* by logically conjuncting every element of *O.nset*, i.e.,  $MBR^*(c, O.subs)$ , for all  $c \in O.subs$ . Therefore, we select a first candidate component  $c_1 \in O.subs$  in SEG(I) and find the next component  $c_2 \in O.subs$  in  $SEG(I) \cap MBR^*(c_1, O.subs)$ .

The  $MBR^*(c_1, O.subs)$  used here is defined in *MS*. Repeating this process, we finally find *O.area* containing all elements of *O.subs*. This process is referred to as *narrowing*. Figure 5 illustrates object finding via narrowing.

**Coping with Object Shape Variation:** In real images, instances of *O* having the same *O.attriv* may have different *O.area.size* due to intrinsic size variations and/or the scale change of the image *O*. This problem can be addressed using Theorem 3, which states that if a revised *O.nset* containing an expanded  $MBR^{**}(c, O.subs)$  is used, we can trap the *O.area* in the core  $\bigcap_{c \in O.subs} MBR^{**}(c, O.subs)$ .

Therefore, expanded  $MBR^{**}(c, O. subs)$  should be used in defining the model set *MS*. After a successive narrowing, we can determine the MBR of the found *O. subs* as an estimation of *O. area*. Furthermore, it is often the case that objects in an image *I* are rotated from their models in *MS*. We can cope with this problem by extending the *MS* to *MS*<sup>\*</sup> containing rotated models.



Figure 5. Finding ICOs by a successive narrowing The neighbour set of Figure 2: h, i, j is used. (a) Segmented image; (b) narrowing by using the neighbour of B; (c) narrowing by using the neighbour of G; (d) narrowing by using the neighbour of W.



Figure 6. Classification of component shapes (a) and (b) The minimum area MBR among vertical and the diagonal

 $(45^{\circ})$  candidates is useful to enhance the approximation accuracy of the component shape (area and aspect ratio); (c) classification of components by area and aspect ratio. This classification is effective for computerization of infinitely varying shapes.

**Coping with Component Variations:** Instances of *O* having the same *O.attriv* may not have identical *O.subs* due to component variations. Typical variations include changes in size and component deletions/additions. The size change problem is addressed by adding shape attributes to each

component, as shown in Figure 6. The shape type is defined so as to reflect the area and aspect ratio of the component. In order to enhance the precision, we enclose the original component by two MBRs that are rotated by  $0^{\circ}$  (non-rotation) and  $45^{\circ}$ , and select the MBR having the lesser area (Figure 6: a, b). The area and aspect ratio of the chosen MBR is used to define the component shape type (Figure 6: c). By using this extended attribute in component identification, we can extend the applicability of *MS* to objects composed of components with changed size. The problem of component deletion/addition is addressed by introducing a likelihood measure between the model *O* and the found object *FO*. More specifically, let *O.subs* and *FO.subs* be their components, respectively, classified by their attributes. Then, we introduce the following likelihood function to measure their similarity:

 $L(FO, O) \equiv (FO.subs \cap O.subs)/(FO.subs \cup O.subs).$ 

Note that L(FO, O) = 1.0 when FO has components identical to those of O.subs. This function can be refined by adding a measure of area similarity between O.area and FO.area. An example refinement LL is (using | FO | to denote the area of FO, etc.):

 $LL(FO, O, FO.a.s, O.a.s) \equiv$ 

 $L(FO, O) \times \min(|FO|, |O|) / \max(|FO|, |O|).$ 

Note that LL(FO, O, FO. a. s, O. a. s) = 1.0 when both FO and O have identical components and sizes.

#### 3. APPLICATIONS

#### 3.1 Object Recognition

**Purpose:** ICO recognition in artificial and real images is performed in order to verify that the NSR can find highly varying ICOs using only a few models.

**Method:** We perform two experiments, OR.ex1, and OR.ex2. In OR.ex1, an artificial color map *I* contains three types of ICOs:  $O_1.attriv = stadium$ ,  $O_2.attriv = park$ ,  $O_3.attriv = apartment$ ,

each having four instances of varying configurations (Figure 7). Image segmentation is performed using pixel color classification to yield SEG(I). Uppermost objects are used to build the NSR model set  $MS^* = MS \cup rot(90^\circ, MS)$ , where MS is the topmost three objects in the figure and  $rot(90^\circ, MS)$  is a rotated version

of MS So the model set  $MS^*$  contains two models for each ICO. Assuming that every configuration of components is possible and *O.area.shape* varies, the

expanded  $MBR^{**}(c, O.subs)$  of the maximum extension is used (see Figure 2 (h, i, and j) and Figure 4).

In OR.ex2, *apartment* recognition is performed using an actual aerial image. The small window in Figure 8 shows the site used to define *MS*. Including the rotated variant,  $MS^*$  contains two models  $(MS^* = MS \cup rot(90^\circ, MS))$ .

**Results and Considerations:** Figure 7 shows the result of OR.ex1. All 12 ICOs were successfully recognized using only two models for each ICO. In addition, for OR.ex2, Figure 8 shows that although not completely free of recognition failure, most of the visually recognizable *apartments* could be recognized automatically using only two models.



Figure 7. OR.ex1

Artificial map including 12 ICOs (stadiums, parks, and apartments). Uppermost three ICOs are used to define NSR models. All 12 ICOs are successfully recognized. For each ICO, the MBR covering *O.subs* is shown as the estimated *O.area*. Components not in *O.subs* might hang out of the MBR.

## **3.2** Automatic Matching of Highly Deviated Landmarkless Images

**Purpose:** In a number of applications of aerial/satellite image analysis, the image matching function is fundamental. Land cover change monitoring, hazard map generation, and map revision are only a few examples. Since adjustment of camera conditions, such as the height and the direction is very difficult when the images are taken on different occasions, we are given two images of differing shift, scale, and rotation. Moreover, landmarks that are usable in image matching are not usually provided in normal images. Therefore, the problem of automatic matching of two highly deviated landmark-less images must be solved in order to automate these tasks. We examine the applicability of NSR to this open problem.

**Method** Two images,  $I_1$  and  $I_2$ , can be automatically matched using NSR. First, we automatically extract NSR models  $MS(SEG(I_1))$  out of  $SEG(I_1)$  and find them in  $SEG(I_2)$ . Notice that we must find varying (shifted, scaled,

and rotated) ICOs out of  $SEG(I_2)$  using the model extracted from  $SEG(I_1)$ . We then try to define similar triangles in  $SEG(I_1)$  and  $SEG(I_2)$  using modeled and found objects, respectively. In generating  $MS(SEG(I_1))$ , we introduce a regular grid in  $SEG(I_1)$  and generate one NSR object out of each partition. We choose a fixed number of components (we used three components having larger areas) out of  $SEG(I_1)$  contained in the partition, which were then used as *O.subs*. Since this problem is sensitive to image rotation, we used the minimum area MBR among rotated candidates, as described in Figure 6 in determining component attributes. Using  $MS(SEG(I_1))$ , we find objects  $FO(SEG(I_2))$  in  $SEG(I_2)$  and then attempt to find similar triangles in  $SEG(I_1)$  and  $SEG(I_2)$  having an identical object arrangement on three vertices. If this fails,  $I_1$  and  $I_2$  are judged not to match. In fact, we used 3 similar triangle pairs in order to enhance the matching accuracy, and if none of the three pairs are judged to be similar, matching failure has occurred. Otherwise, we determine the average translation (shift, scale, and rotation) of  $I_2$  from  $I_1$  using found (at least two) similar triangles. Finally, the two images are matched using the estimated parameters.



Figure 8. OR.ex2

Apartments in a real aerial image (Tama New Town in Tokyo) are recognized. The small rectangle is the site used to define an NSR model.

**Results and Considerations:** Figures 9 (a and b) are a sample pair of task images, which are generated by cutting two overlapping images out of an aerial image, and applied an artificial translation of (scaling, rotation) =  $(1.25, 60^{\circ})$  to the image of Figure 9 (b). Figure 9 (c and d) are segmented versions. The most similar triangle pair found is shown with their gravity centers and the matching result is shown in Figure 9 (e). In this example, we obtained a very accurate estimation

of (scaling, rotation) =  $(1.24, 60^{\circ})$ . Ten experiments were performed using different task images and the result of which

are summarized in Table 1. The accuracy of scaling/rotation parameter estimation was very high. The average error rate was 0.51% / 1.99%.

Case	Scale	Rotation	S. Error	R. Error
1	0.51/0.5	45.8/45	2.0	1.8
2	0.80/0.8	51.1/50	0.0	2.2
3	1.60/1.6	5.4/5	0.0	8.0
4	1.25/1.25	310.1/310	0.8	0.0
5	2.00/2.0	311.0/315	0.0	1.3
6	0.63/0.63	354.9/355	0.2	0.0
7	1.98/2.0	43.4/45	1.0	3.6
8	1.24/1.25	60.0/60	0.8	0.0
9	0.63/0.63	15.4/15	0.3	2.7
10	0.50/0.5	316.0/315	0.0	0.3
Mean			0.51	1.99

TABLE 1. Accuracies Scale and Rotation
Estimated/Real data, rotations are in degrees, and errors are in %.

#### 4. CONCLUSIONS

We introduced the concept of the ill-configured object (ICO) and proposed the concept of neighbor set representation (NSR) of an object to represent the ICO. Several important properties of NSR were clarified mathematically, especially the possibility of characterizing an ICO (including non-ICO) as a solution (fixed point) of a set theoretic equation of NSR of the object. Using this property, we proposed an iterative algorithm by which to find an ICO in an image. In addition, we reported two applications of NSR. The first being ICO object recognition in artificial and real images, and the second being automatic matching of highly deviated landmark-less images. In the former, we illustrated that ICO objects of varying configurations can be recognized using only a small NSR model set. In the latter, we illustrated that highly deviated landmark-less images can be automatically matched with high accuracy. This function provides a foundation of automatic land cover change analysis using satellite/aerial images obtained under different camera conditions. Future research includes extension and applications of the NSR concept to a wider range of media data.

#### REFERENCES

Watanabe, T. and Sugawa, K., and Sugihara, H., 2002. A new pattern representation scheme using data compression. *IEEE trans. PAMI*, 24(5), pp. 579-590.

Shapiro, L. G. and Haralick, R. M., 1981. Structural descriptions and inexact matching. *IEEE trans. PAMI*, 3(5), pp. 504-519.

Shapiro, L. G. and Haralick, R. M., 1982. Organization of relational models for scene analysis. *IEEE trans. PAMI*, 4(6), pp. 595-603.

Shapiro, L. G. and Haralick, R. M., 1985.

A metric for comparing relational descriptions. *IEEE trans. PAMI*, 7(1), pp. 90-94.

Vosselman, G., 1992. Relational matching. LNCS 628, Springer-Verlag, Berlin.

Kim, W. Y and Kak, A. C., 1991. 3-D object recognition using bipartite matching embedded in discrete relaxation. *IEEE trans. PAMI*, 13(3), pp. 224-251.

Oommen, B. J. and Loke, R. K. S., 2001. On the pattern recognition of noisy subsequence trees. *IEEE trans. PAMI*, 23(9), pp. 929-946.

Geman, S. and Geman, D., 1984. Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images. *IEEE trans. PAMI*, 6, pp. 721-741.

Young, T. Y. and Fu, K. S. Eds., 1986. *Handbook of pattern recognition and image processing*. Academic Press, San Diego.

Ohta, Y., 1985. Knowledge-based interpretation of outdoor natural color scenes, Pitman Publishing, Boston.

Chang, S. K., Shi, Q. Y., and Yan, C. W., 1987. Iconic indexing by 2-D strings. *IEEE trans. PAMI*, 9(3), pp. 413-428.

Mohan, A., Papageorgiou, C., and Poggio, T., 2001. Example-based object detection in image by components. *IEEE trans. PAMI*, 23(4), pp. 349-361.

Suto, T., Nishikawa, Y., Watanabe, T., and Sugawara, K., 2000. Recognizing objects in aerial images. *Proc. JSPRS 2000*, pp.263-266. (in Japanese).

Nishikawa, Y., Kubo, T., Watanabe, T., and Shinada, T., 2001. MLDI: Matching largely deviated landmark-less images. *Proc. JSPRS 2001*, pp.149-152. (in Japanese).

Vapnik, V., 1998. Statistical learning theory. John Wiley & Sons, Inc., New York.

Samet, H., 1993. *Applications of spatial data structures*, Addison-Wesley.

## ACKNOWLEDGEMENTS

T. Suto, Y. Nishikawa and T. Kubo, students in our laboratory, are acknowledged for prototyping several applications of the NSR theory. Mr. F. Komura at Hitachi is acknowledged for discussions on aerial image registration problems. Prof. Dr. A. Gruen at ATH, who suggested to post this work to PCV'02, is also acknowledged.



Figure 9. Matching experiment using real images (a) and (b) Task images;

(a)  $(408 \times 606)$  pixels. (scaling, rotation) =  $(1.00, 0^{\circ})$ ,

(b) (scaling, rotation) =  $(1.25, 60^{\circ})$ ,

(c) and (d) the most similar triangles found in two segmented images are shown. The vertices are gravity centers of found ICOs. Gravity centers of triangles are also shown;(e) matching result