'90 KACC 1990. 10. 26～27

An Image Association Technique Employing Constraints among Pixels

Seiji Ishikawa, Tomokazu Goda, Kiyoshi Kato

Department of Electrical, Electronic and Computer Engineering
Kyushu Institute of Technology
Tobata, Kitakyushu 804 JAPAN

Abstract The present paper describes a new technique for associating images employing a set of local constraints among pixels on an image. The technique describes the association problem in terms of consistent labeling which is an abstraction of various kinds of network constraint problems. In this particular research, a pixel and its gray value correspond to a unit and a label, respectively. Since constraints among units on an image are defined with respect to each n-tuple of pixels, performance of the present association technique largely depends on how to choose the n-tuples on an image plane. The main part of this paper is devoted to discussing this selection scheme and giving a solution to it as well as showing the algorithm of association. Also given are some results of the simulation performed on synthetic binary images to examine the performance of the proposed technique, followed by the argument on further studies.

1. INTRODUCTION

There are increasing demands on an intelligent robot to be employed in a factory that it should recognize those objects on a conveyer belt or in a bin precisely to grasp them properly by robot hands. In such three-dimensional object recognition problems, typical difficulties include the case where objects reveal themselves in various appearances, and the case where an object to be recognized is occluded by other objects resulting in insufficient information with the shape of the object concerned. The former difficulty can be solved in principle by the employment of model-based recognition techniques[1,2]. Some studies[3,4] propose appearance-oriented model description and they keep in the memory those features obtained from every aspect of a model. Aspects are provided from sorting all appearances of an object from various view directions by fixing eyes on topological changes. The concept of aspects(or appearance-oriented models) with respect to three-dimensional objects seems promising, since it is likely that man keeps in mind three-dimensional object's shape in the form of aspects. Moreover it can be applied to automatic generation of recognition procedure[5] where a classification tree is produced by a program which examines features with each aspect of a model object. Model-based recognition, however, requires a lot of time in matching process and still needs study before it is put to practical use.

Occlusion is a constant difficulty in three-dimensional object recognition. When object A is partially hidden from sight by object B, object A is said to be occluded by object B. This relation, of course, disappears if object B is taken away, but, before that, it is necessary in advance to judge whether the occluded object is of present concern or not. This suggests the importance of associating whole shape of the occluded object from its visible part. The present paper therefore concentrates on an image association problem. Already proposed association techniques are those based on neural networks[6,7,8] which often suffer from interference between memorized images, *i.e.*, individual images cannot be separated from others when recalling. This happens when the number of memorized images increase.

The present paper devotes itself into associating two-dimensional objects such as characters. The principle of association employed here can, however, be applied to the three-dimensional case[9]. The association is performed employing local constraints among pixels which is nothing to do with neural networks at the moment. This idea comes from [10] which deals with distorted images matching, but the idea is described here as a consistent labeling problem[11] and its solution corresponds to an associated image. A new technique is also proposed for accelerating search for the solution. Performance of the proposed association technique including the avoidance of interference problem largely depends on the selection scheme of local sets of pixels that constrain one another on an image. The algorithm of how to choose these sets of pixels is given and it is supported by theoretical arguments. Some results of the simulation employing synthetic images are also shown and further studies are discussed including a noisy image case.

2. PROBLEM DESCRIPTION AND THE SOLVING TECHNIQUE

Let us take a digitized image I which has $M \times N \equiv m$ pixels. A gray value of a pixel on I is denoted by f_{ij}. If a new subscript $i \equiv (i-1)N+j$ is introduced, it is written as f_i $(i=1,2,\ldots,m)$. Then an image I is equivalently represented by $I \equiv \{(i, f_i) \mid i=1,2,\ldots,m\}$. The first component of the pair is understood as a unit and the second as a label to be given to the unit. If the set of the units and the set of the labels are denoted by U and L, respectively, $U=\{1,2,\ldots, m\}$ and $L=\{1, 2,\ldots,l\}$, where l is the maximum gray value given *a priori*. The unit constraint set T whose elements are n-tuples of pixels is chosen in the manner stated later.

Now suppose that K images $I^{(k)} \equiv \{(i, f_i^{(k)}) \mid i=1,2,\ldots, m\}$ $(k=1,2,\ldots,K)$ are to be memorized as reference data. For each image, there exists a unit-label constraint subset R_k based on T, *i.e.*, if $(u_1, u_2, \ldots, u_n) \in T$, there exists $(u_1, l_1^{(k)}, u_2, l_2^{(k)}, \ldots, u_n, l_n^{(k)}) \in R_k$ where $(u_\cdot, l_\cdot^{(k)}) \in I^{(k)}$. The unit-label constraint set R is then defined by

$$R = \bigcup_{k=1}^{K} R_k \ , \quad R \subseteq (U \times L)^{2n},$$

which provides the local constraints among pixels. It is clear that the consistent labeling problem with respect to the quadruplet (U, L, T, R) has all the memorized images $I^{(k)}$ $(k=1,2,\ldots, K)$ as its solutions. Unit-label constraint set R is the knowledge the proposed association technique possesses.

An unknown image X acquired by an image acquisition device is also represented in the form of a unit-label constraint set based on T. Considering that image X can partially be incomplete because of occlusion, its unit-label constraint set (also denoted by X) is defined in the following way:

$$X=\{(u_{x1}, l_{x1}, u_{x2}, l_{x2},\ldots, u_{xn}, l_{xn}) \mid (u_{x1}, u_{x2},\ldots, u_{xn}) \in T, \text{ not } ((l_{x1}=nil) \wedge (l_{x2}=nil) \wedge \ldots \wedge (l_{xn}=nil))\}.$$

Note that the state a pixel has unknown gray value is expressed as '*nil*' in the above definition.

Since image X agrees with one of those images stored in R (or does not agree with any of them), some of (or all of) the elements in R are inconsistent with those in X, *i.e.*, there is disagreement in labels with respect to the same units between R and X. Therefore once X is given, the set R receives screening by X. This screening divides the elements in R into two classes according to if they are consistent with those elements in X. The set of the elements in $R \equiv R(0)$ which survived this screening is denoted by $R(1)$. The following algorithm **REMOVE**[12] performs this screening:

REMOVE

begin
  copy R to $R(1)$;
  for all $\mathbf{r} \equiv (u_{r1},l_{r1},u_{r2},l_{r2},\ldots,u_{rn},l_{rn}) \in R$ *do*
    for all $\mathbf{x} \equiv (u_{x1},l_{x1},u_{x2},l_{x2},\ldots,u_{xn},l_{xn}) \in X$ *do*
      if $\mathbf{r}|_T = \mathbf{x}|_T$
        then if for a certain i
          $(l_{ri} \neq l_{xi})$ and $(l_{xi} \neq nil)$
            then remove $\mathbf{r}$ from $R(1)$
end.

Note that, if $\mathbf{s} = (u_{s1},l_{s1},u_{s2},l_{s2},\ldots,u_{sn},l_{sn})$, $\mathbf{s}|_T=(u_{s1},$

$u_{s2},\dots,u_{sn}$).

By solving the consistent labeling problem specified by the quadruplet (U, L, T, $R_{(1)}$), an image associated from X is obtained as its solution. The ordinary technique for solving consistent labeling problems is depth first search [11]. Here, instead of employing depth first search, we introduce a new technique based on the following definition concerning labels assignment:

Definition

$$u \leftarrow l \quad \textit{iff} \quad \forall\, \mathbf{t} \in T(u),\ \exists\, \mathbf{r} \in R \quad \text{s.t.}\ \ \mathbf{r}|_T = \mathbf{t}\ ,\ \ \mathbf{r} = (\dots, u, l, \dots)$$

Here $T(u)=\{\ \mathbf{t} \mid \mathbf{t} \in T,\ \mathbf{t}$ contains $u\}$, and the assignment of label l to unit u is denoted by $u \leftarrow l$.

Once the reduced set $R_{(1)}$ is obtained, some unknown units on input image $X \equiv X_{(0)}$ are given labels (if any) according to $R_{(1)}$ and the above definition, which yields $X_{(1)}$. The augmented image $X_{(1)}$ then gives constraint to $R_{(1)}$, and by applying **REMOVE** to $R_{(1)}$, $R_{(2)}$ is produced. This $R_{(2)}$ in turn contributes to assigning labels to some of the unknown units in $X_{(1)}$, which results in $X_{(2)}$. In the same way, this process continues until final $X_{(\bullet)}$ is obtained which is the image associated from X. Note that, for a positive integer i, $\mathrm{n}(R_{(i)}) \geq \mathrm{n}(R_{(i+1)})$ and $\mathrm{n}(X_{(i)}) \leq \mathrm{n}(X_{(i+1)})$. Here $\mathrm{n}(S)$ denotes the number of the elements in the set S. **Figure 1** shows how this reduction and augmentation process goes.

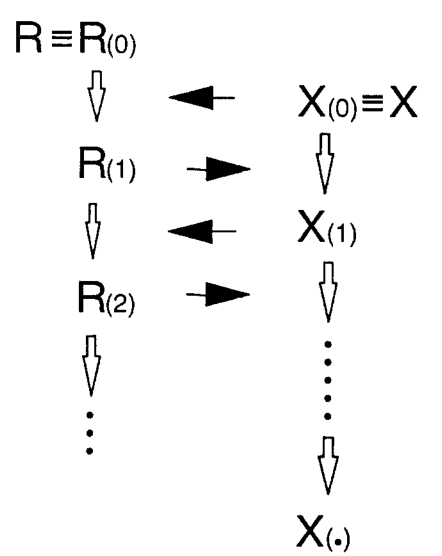

Fig. 1 Reduction and augmentation process of unit-label constraint set R and input image X. Final associated image is given by X(•).

3. SELECTION SCHEME

It is of great importance to choose the elements of unit constraint set T carefully. As stated before, each element in T is composed of a set of n pixels(an n-tuple of pixels) and these pixels constrain each other on a given image. It is argued[13] that all the elements in T do not necessarily have the same value of n. In this particular research, however, they are assumed to have the same value of n.

In the following argument, an image is supposed to be a binary image whose gray values are referred to as black and white. A set of black pixels on a binary image I is denoted by I_F which, as usual, corresponds to the shape of an object(or objects). Note that subscript F comes from 'Figure' used in gestalt psychology.

A constant c is defined with respect to K images, $I_i (i=1,2,\dots,K)$, as

$$c = \max_{i<j} \mathrm{n}(\{\ \mathrm{e} \mid \mathrm{e} \in I_{F,i} \cap I_{F,j}\}).$$

This gives the largest number of black pixels shared between each pair of images among $I_1, I_2, \dots, I_K$. Employing the constant c, a disconnector between images I_i and I_j(denoted by $\tau_{c+2,i|j}$) is defined as follows;

$$\tau_{c+2,i|j} = E_1 \cup E_2 \cup E_3$$

where

(i) $E_1 = \{\ \mathrm{e} \mid \mathrm{e} \in I_{F,i} \cap I_{F,j}\}$
(ii) $E_2 \subset \{\ \mathrm{e} \mid \mathrm{e} \in I_{F,i} \cap \overline{I_{F,j}}\}$, $\quad \mathrm{n}(E_2) \geq 1$
(iii) $E_3 \subset \{\ \mathrm{e} \mid \mathrm{e} \in \overline{I_{F,i}} \cap I_{F,j}\}$, $\quad \mathrm{n}(E_3) \geq 1$
(iv) $\mathrm{n}(E_1 \cup E_2 \cup E_3) = c+2$

It is actually a set of c+2 pixels(a (c+2)-tuple of pixels) and contributes to the separation among memorized images when recollecting. Another important set concerning image I_i is a descriptor $\tau_{c+2,i,\cdot}$ which is defined as

$$\tau_{c+2,i,\cdot} \subset I_{F,i}, \quad \mathrm{n}(\tau_{c+2,i,\cdot}) = c+2\ .$$

This is again a set of c+2 pixels. Examples of disconnectors and descriptors are shown in **Fig.2**.

The following theorem holds.

Theorem

If the following selection scheme is employed for establishing unit constraint set T, a recollected image is never connected with any other images in the memory.

Selection Scheme

(1) Set up a disconnector $\tau_{c+2,i|j}$ between each pair of images $I_{F,i}$ and $I_{F,j}$ among $I_1,I_2,\ldots,I_K$.

(2) Set up a set of descriptors $\{\tau_{c+2,i,u}|\ u=1,2,\ldots,N_i\}$, so that $n(\tau_{c+2,i,u} \cap \tau_{c+2,i,v}) = k+1$ and $\cup_{u=1,N_i} \tau_{c+2,i,u} = I_{F,i}$ holds.

(3) With a disconnector $\tau_{c+2,i|j}$ satisfying $n(E_1) \leq c-2$, if all the pixels in $\tau_{c+2,i|j}$ belong to an image $I_{F,k}$, set up a descriptor $\tau_{c+2,k,\cdot}$ composed of every pixel in $\tau_{c+2,i|j}$.

Note that, since disconnectors and descriptors become the element of the unit-label constraint set R, the selection scheme indirectly tells how to choose the element of T.

[*Outline of Proof*]

Since every element in T has $c+2$ pixels, $\tau_{c+2,i|j}$ and $\tau_{c+2,i,\cdot}$ are abbreviated as $\tau_{i|j}$ and $\tau_{i,\cdot}$, respectively, in this proof.

Here two main cases are given outlines of proof.

Case 1.

The first case is illustrated in **Fig.3**(a) where images I_i and I_j are overlapped. Disconnector $\tau_{i|j}$ is shown by a rectangle containing $c+2$ pixels including those in the overlapped part. Two patterns provided by this disconnector are given in Fig.3(b) where only dark part shows the presence of the image indicated in the figure. Now this is the case that there is no memorized image which contains all of the pixels in $\tau_{i|j}$ as its black pixels. Given part of image I_i as shown in Fig.3(a) by gray area, it is then connected in turn by descriptors $\tau_{i,1}$, $\tau_{i,2}$,....., and finally reach the other end of the image by τ_{i,N_i}. The connection continues beyond this end by descriptor $\tau_{j,1}$ of image I_j since τ_{i,N_i} and $\tau_{j,1}$ have common black pixels, and no inconsistency occurs in spite of this illegal connection. The presence of the disconnector, however, forbids

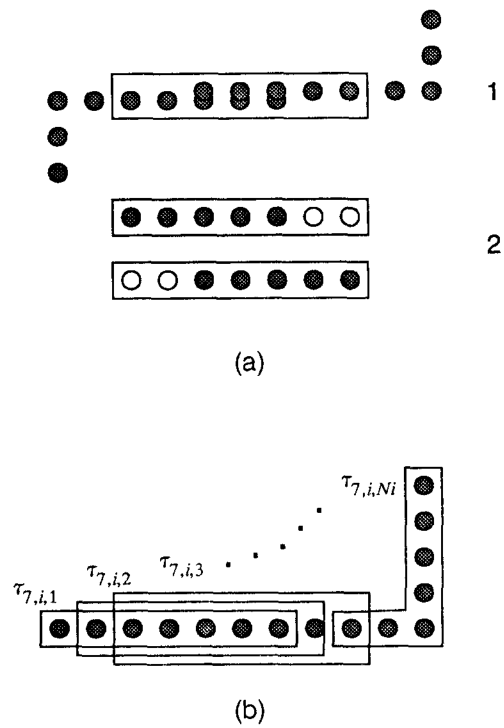

Fig. 2 Examples of disconnectors and descriptors: (a1)Two overlapped images and their disconnector, and (a2)Two patterns of unit-label constraint the disconnector provides concerning respective images; (b)A set of descriptors.

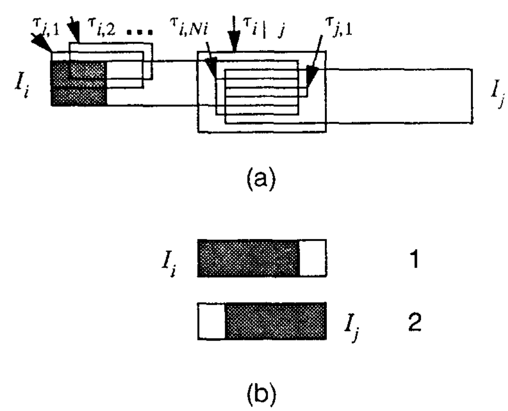

Fig. 3 Images and disconnectors: (a) The case that there is no other image which includes all the pixels in the disconnector $\tau_{i|j}$; and (b) the unit-label constraint patterns the disconnector provides.

this undesirable connection by the pattern Fig.3(b1) which indicates that image I_i does not exist in the right-hand side of the disconnector.

Case 2.

Even if there is $\tau_{i|j}$ between I_i and I_j, undesirable connection occurs if there is the third image I_k which has all the pixels in $\tau_{i|j}$ as black pixels as shown in Fig.3(c). Let us assume that this only happens between I_i and I_j and no other pairs. Then each of the three disconnectors $\tau_{i|j}$, $\tau_{j|k}$ and $\tau_{k|i}$ has three patterns of unit-label constraints as shown in Fig.3(d). Disconnectivity between I_i and I_k, and I_j and I_k are approved as they are the same cases as Case 1. Suppose part of image I_i be given as shown in Fig.3(c). Its descriptors $\tau_{i,1}$, $\tau_{i,2}$,....., τ_{i,N_i} reproduce unknown part of I_i to finally obtain its whole shape. But pattern (d3) in Fig.3 bridges I_i to I_j,

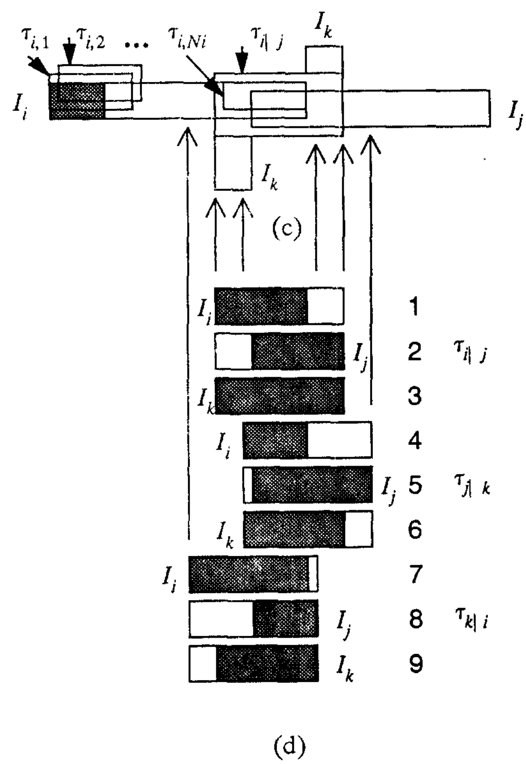

Fig. 3(cont'd) (c) The case that another image I_k includes all the pixels in $\tau_{i|j}$; and (d) the unit-label constraint patterns the disconnectors $\tau_{i|j}$, $\tau_{j|k}$, and $\tau_{k|i}$ provide.

which is undesired. This is actually forbidden since it contradicts to pattern (d7) in Fig.3. To settle this problem, mutually consistent three patterns chosen from respective disconnectors must be found on condition that image I_i is already recollected. Obviously the solution is (d1), (d4) and (d7) in Fig.3. Thus the expansion of image I_i stops after descriptor τ_{i,N_i} is employed.

4. EXPERIMENTAL RESULTS

The proposed association technique was programmed using C language, and its performance was examined employing synthetic binary images of 5×5 pixels. Memorized eight images are shown in **Fig.4**. Related data are as follows: c=2, n(T)=77 and n(R)=432. The descriptors and the disconnectors concerning this experiment are therefore all 4-tuples of pixels. **Figure 5** shows some results of the association process.

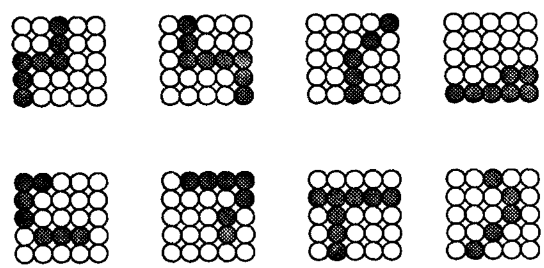

Fig. 4 Binary images employed for the experiment.

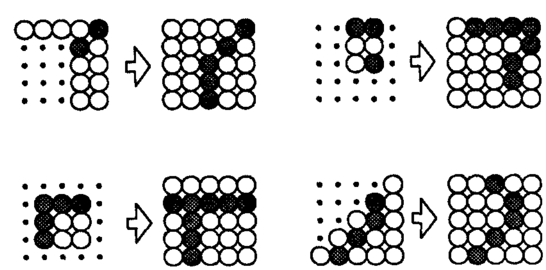

Fig. 5 Experimental results of association.

5. DISCUSSION

Experimental results show noninterference among memorized images when recollecting. This performance is realized by 77 4-tuples of pixels which compose the unit constraint set T, and they are only 0.61% of ${}_{25}C_4$=12,650 possible combinations of 4-tuples. Although the memorized images are not so many in the experiment as they should be in practice, the larger number of images will not disturb noninterference among them if only the unit constraint set T is defined according to the selection scheme.

As described before, image association can be formalized as a consistent labeling problem. Solutions are therefore found(or no solution is concluded) if only depth first search is applied to a given unit-label constraint set. This technique, however, causes time explosion in searching solutions when the size of the set is large. The filtering technique employed in this paper tries to escape from this problem.

The present algorithm for association does not take noisy environment into consideration yet. In the reported association technique [14] which also employs consistent labeling, the unit-label constraint set R is given filtering only once by input image X (which means longer execution time) to produce a reduced set R^*, and depth first search is employed to find consistent labeling from R^*. There, noisy images are taken into consideration by introducing admissible match in the filtering. This means that, if an element in R almost matches to one of the elements in X, it is not discarded. In the same way, the notion of admissibility can be introduced in the assignment of a label to a unit in the present research. The definition of the label assignment stated in **2** may be given modification in the future study in order to include noisy case in the proposed association technique.

6. CONCLUSION

A new technique for associating images was described which employed constraints among a set of pixels. The association problem was formalized as a consistent labeling problem and a filtering technique to find consistent labeling was explained. The selection scheme of the unit constraint set was then given in the form of a theorem and, according to the scheme, an experiment employing synthetic binary images was successfully performed. Extension of the present technique to a noisy case remains as a future study.

REFERENCES

[1] Perkins, W.A.:Model-based vision system for scene containing multiple parts, Proc. 5th Int'l Joint Conf. Artif. Intell., 678- 684 (1977).

[2] Horn, B.K.P., Ikeuchi, K.: The mechanical manipulation of randomly oriented parts, *Sci. Amer.*, **251**, 2, 100–111(1984).

[3] Koenderik,J.J., Van Doorn,A.J.:The internal representation of solid shape with respect to vision, *Biol. Cybern.*, **32**, 211–216(1979).

[4] Fekete,G., Davis,L.S.:Property spheres: A new representation for 3-D object recognition, Proc. Workshop on Comput. Vision, 192–201(1984).

[5] Ikeuchi,K.,Kanade,T.:Towards automatic generation of object recognition program, *Proc. IEEE*, **76**, 8, 1016–1035(1988).

[6] Nakano, K.:Associatron—A model of associative memory, *IEEE Trans. Syst. Man, Cybern.*, **SMC-2**, 3, 381–388 (1972).

[7] Kohonen,T.:Correlation matrix memories, *IEEE Trans.Computers*, **C-21**, 4, 353–359(1972).

[8] Hirai, Y.:A model of human associative processor(HASP),*IEEE Trans.Syst.Man, Cybern.*, **SMC-13**, 5, 851–857(1983).

[9] Gouhara, K.:Application of constraint analysis to three-dimensional object recognition, M.S. Thesis, Kyushu Inst. Tech., Dept. Comput. Eng. (1990).

[10] Ullmann,J.R.:Subset methods for recognizing distorted patterns, *IEEE Trans. Syst. Man, Cybern.*, **SMC-7**, 3, 180–191 (1977).

[11] Haralick, R.M., Shapiro, L. G. :The consistent labeling problem: Part 1, *IEEE Trans. Pattern Anal. Machine Intell.*, **PAMI-1**, 2, 173–184 (1979).

[12] Kurokawa,K.,Ishikawa,S.,Kato,K.:Associating patterns by consistent labeling, *Trans. Inform. Process. Soc. Japan*, **31**, 3, 511–515 (1990).

[13] Ullmann, J.R., Haralick, R.M., Shapiro, L.G. : Computer architecture for solving consistent labelling problems, *Comput. J.*, **28**, 2, 105–111 (1985).

[14] Kurokawa,K: Research on image understanding employing consistent labeling, M.S. Thesis, Kyushu Inst. Tech., Dept. Comput. Eng.(1990).