AUTONOMOUS AND ROBUST STRUCTURING OF REAL ENVIRONMENT BY SEARCHING COMPLEX REGIONS

Hidenori Takauji, Io Nakayama, Shun'ichi Kaneko and Takayuki Tanaka

Graduate School of Information Science and Technology, Hokkaido University
Kita-14, Nishi-9, Kita-ku, Sapporo, Japan
uji@ssc.ssi.ist.hokudai.ac.jp, nakayama@ssc.ssi.ist.hokudai.ac.jp, kaneko@ssi.ist.hokudai.ac.jp, ttanaka@ssi.ist.hokudai.ac.jp
http://www.ssc-lab.com

KEY WORDS: complex regions extraction, image searching, image scaling, orientation code entropy, mobile robot

ABSTRACT:

This paper aims to propose a fast image searching method from environmental observation images even in the presence of scale changes. A new scheme has been proposed for extracting feature areas as tags based on a robust image registration algorithm called Orientation code matching. Extracted tags are stored as reference images and utilized in tag searching. As the number of tags grows, the searching cost becomes a serious problem. Additionally, change in viewing positions cause scale change of an image and matching failure. In our scheme, richness in features is important for tag generation and the entropy is used to evaluate the diversity of edge directions which are stable to scale change of the image. This characteristic contributes to limitation of searching area and reduction in calculation costs. Scaling factors are estimated by orientation code density which means the percentage of effective codes in fixed size tag areas. An estimated scaling factor is applied to matching a scale of reference images to one of observation images. Some experiments are performed in order to compare computation time and verify effectiveness of estimated scaling factor using real scenes.

1 INTRODUCTION

We have proposed a new scheme for estimating scaling of object images compared to reference images which are registered in a image database off line, and developed a prototype system for searching images in the scene. Orientation codes(F.Ullah et al., 2001) is utilized for designing formalizations necessary for the estimation, which are called as orientation code density and conditional entropy. In this study, references are called *tags* which bear isolated objects with pictorial prominency measured by the conditional entropy.

Image registration techniques have been utilized in many fields, for instance, machine vision for production lines where we need positioning, appearance and defect inspection, video media processing for advanced contents analysis, or web-oriented mining systems for image retrieval. This is one of the most fundamental techniques in image processing field for practical applications. As they recently have a wide range of applications, robust methods available in the real world scene are needed, where we have some irregularities or ill-conditions, such as, local change in brightness, highlighting, shadowing and occlusion(S.Kaneko et al., 2002, F.Ullah and S.Kaneko, 2004). Our proposed method is based on a template matching and consists of an extraction process and a matching process where feature areas suitable for the matching are extracted and stored as reference images and then utilized in image searching. Such a system has wide applications, for example, correspondence search for stereo vision, estimation of optical flow in motion-image sequence analysis and landmark recognition in robot vision. Moreover, the distribution of feature quantity used in the extraction process becomes rich in case of objects which are isolated from background patterns and usually used in conventional template matching. Therefore, our proposed method may have applicability to a image search for a predefined object.

Filters for feature extraction have been proposed which extracts feature points from observation image such as Moravec operator (H.Moravec, 1977), SUSAN operator (S.M.Smith and J.M.Brady, 1997), Harris operator (C.Harris and M.Stephens, 1988) and Kanade-Lucas-Tomasi feature tracker (C.Tomasi and T.Kanade, 1991, J.Shi and C.Tomasi, 1994). Since these methods use a distribution of brightness in itself as input data, they have a problem about robustness against ill-conditions such as changes in brightness. A method based on distribution of autocorrelation have been proposed for a fast image matching by fast hardwares (T.Mori et al., 1995). However, because of the sum of absolute difference used as a correlation calculation, the method is not robust against ill-conditions. The normalized cross correlation is useful for avoiding the mismatch due to illumination change over the whole image. However, local change in brightness causes the miss registration.

This paper aims to propose a new scheme for extracting candidate areas and selecting feature areas from them based on the orientation code representation and matching which is a robust image registration algorithm against local change in illumination, occlusion, highlighting and shadowing. A reference image is searched from various scenes where some irregularities happen such as translation, rotation, image scaling, etc. In this paper, we deal with image scaling which is frequently occurred in real systems such as robot vision, image measurement, etc. Methods for estimating scaling factors based on geometric features have been proposed. For example, methods for extracting geometric features using Hough transform(D.Ballard, 1981) and methods based on curve moment (D.M.Zhao and J.Chen, 1997). These methods adapt to scale changes and affine transformations too. However, because of edge features with strong contrast which have been thought as one of the robust features, the result excessively depends on edge extraction algorithms and the accuracy of estimation decreases in case of areas include few edge features. Methods which directly uses image based features to estimate scaling factors has not been found so far. Since the robust tagging is based on template matching, searching with resized reference images are needed for adapting to large scale changes and call for much calculation costs. In this paper, we propose a fast tag searching method based on scale estimation by an image feature: the density of orientation code in a tag area, and limita-

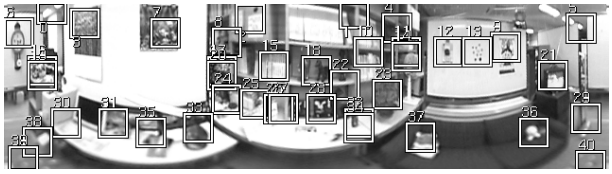


Figure 1: Tags based on orientation code entropy.

Figure 2: Example regions.

tion of searching area.

The paper consists as follows: Section 2 introduces tags as references with prominency and the method for extracting them. In Section 3, the definition of orientation code density is provided and a novel approach is proposed to extract tags using the orientation code density, which is applied to estimating scaling factors in Section 4. In Section 5, a fast tag searching method is described. In Section 6, we perform some experiments to verify effectiveness of the proposed method, and finally in Section 7, we gives the concluding remarks.

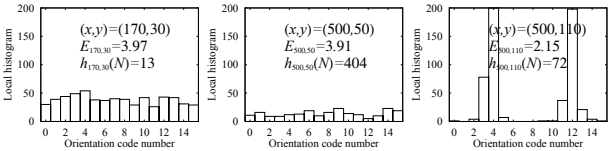
2 TAG DEFINITION

In this study, we define a *tag* as a reference or reference image having prominent pictorial features, which are necessary for discriminating itself from the background, and then we propose a novel scheme for searching tags in the real world scene by utilizing a scale estimation method.

Orientation code matching, hereafter OCM, has been based on local characteristics of orientational or gradient structures, different from brightness itself, which can be robust against local illumination changes(F.Ullah and S.Kaneko, 2004). Orientation codes are quantized numerical codes which can be robust for additive noise in brightness observation in the real world scene. For realizing effective and efficient search method, we want to propose a novel scheme to utilize this robust OCM for extracting tags as prominent or richer regions which are full of prominency from the environmental scene. We have to consider about what areas should be tags extracted as richer areas in information from the scene. Edge features of strong pictorial contrast have been thought as one of the robust features and utilized in many methods, but we focus our attention to regions of rich texture by evaluating orientation code distributions. In order to quantify the prominency around any pixel in terms of the orientation code distributions, an information entropy (S.Kaneko et al., 1992) could be utilized. Entropy E_{xy} was calculated through a histogram of Orientation codes in any neighborhood area centered at the position (x, y). Next, we define *richness* by normalizing the entropy beyond a pre-specified threshold and by clipping the smaller values into 0. In our robust tagging, the richness is calculated at each position (x, y) and local maxima can be extracted as tags. Figure 1 shows an example of tag extraction from one of the real world scene. We find that in the entropy distribution higher values correspond to richer areas in the scene and it is applicable to autonomous extraction of tags. In Section 3, we improve the definition of richness in order to obtain a more precision in matching, and then in Section 4, we define new features based on the revised richness for estimating scaling factors.

3 ORIENTATION CODE DENSITY

Any effective orientation codes are required to be calculated at pixels having not less contrast values enough for derivative computations. We discriminate the futile codes N by use of a threshold values Γ. In order to design an effective richness, it is necessary to take density of codes into account in definition of the richness. Figure 2 shows three example regions in the indoor scene: a part of poster, a decoration, and a desk top, all of which are of size 25×25, respectively at $(x, y) = (170, 30)$, $(500, 50)$, and $(500, 110)$.



(a)Poster (b)Decoration (c)Desk

Figure 3: Local histograms of three regions.

Figure 3 shows the histograms $h_{xy}(i)$ $(i = 0, 1, \cdots, 15)$ in the three regions shown in Figure 2. Because of the diversity of orientation code distribution, the poster region (a) and the part of decoration (b) have high values of entropy E_{xy} such as $E_{170,30} = 3.97$ and $E_{500,50} = 3.91$. On the other hand, the desk top region (c) includes many similar codes and then the entropy $E_{500,110} = 2.15$ becomes low. From the point of view of entropy, (a) and (b) are equally richer areas in information, while (c) is not rich compared with them. Figure 4 shows their coded images of (a), (b) and (c) through the orientation coding, in which effective codes are shown as white pixels and futile codes as black ones. Since the number of futile codes are $h_{170,30}(N) = 13$ in (a) and $h_{500,50}(N) = 404$ in (b), we find the image of the poster (a) has more effective information than in (b). In order to evaluate the fraction of effective codes, we improve the richness definition based on their density in local but whole domains. Letting I define an event where an orientation code is i $(i = 0, 1, \cdots, N)$ in a local domain of the size $M \times M$ centered at (x, y), and $\{h_{xy}(i)\}$ a histogram of the codes, we define a relative frequency $P^{\rho}_{xy}(i)$ of the event I as

$$P^{\rho}_{xy}(i) = \frac{h_{xy}(i)}{M^2} \quad (\text{where } i = 0, 1, \cdots, N) \tag{1}$$

We also define an event J where an orientation code is effective $(j = 0, 1, \cdots, N-1)$ or futile $(j = N)$, therefore we get a relative frequency $Q^{\rho}_{xy}(j)$ by

$$Q^{\rho}_{xy}(j) = \begin{cases} \frac{M^2 - h_{xy}(N)}{M^2} & \text{if} \quad j = \text{Effective code} \\ \frac{h_{xy}(N)}{M^2} & \text{if} \quad j = \text{Futile code} \end{cases} \tag{2}$$

In this study, the fraction of effective codes is utilized to evaluate actual effectivity of the prominency designated by the entropy by introducing a conditional entropy. Table 1 shows joint probabilities of the events I and J, and then Table 2 shows conditional probabilities of I given J. Therefore, the conditional entropy is

(a)Poster (b)Decoration (c)Desk

Figure 4: Coded images (effective code = white, futile code = black).

Table 1: Joint probabilities of event I where the code becomes i, and event J where it is effective or futile.

$P^{\rho}_{xy}(i,j)$		J		$P^{\rho}_{xy}(i)$
		Effective code	Futile code	
I	0	$\frac{h_{xy}(0)}{M^2}$	0	$\frac{h_{xy}(0)}{M^2}$
	1	$\frac{h_{xy}(1)}{M^2}$	0	$\frac{h_{xy}(1)}{M^2}$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	$N-1$	$\frac{h_{xy}(N-1)}{M^2}$	0	$\frac{h_{xy}(N-1)}{M^2}$
	N	0	$\frac{h_{xy}(N)}{M^2}$	$\frac{h_{xy}(N)}{M^2}$
$Q^{\rho}_{xy}(j)$		$\frac{M^2-h_{xy}(N)}{M^2}$	$\frac{h_{xy}(N)}{M^2}$	1

given as follows:

$$
\begin{aligned}
E^{\rho}_{xy} &= \sum_j Q^{\rho}_{xy}(j) \sum_{i=0}^{N} \left\{ -P^{\rho}_{xy}(i|j) \log_2 P^{\rho}_{xy}(i|j) \right\} \\
&= -\frac{M^2 - h_{xy}(N)}{M^2} \sum_{i=0}^{N-1} P_{xy}(i) \log_2 P_{xy}(i) \\
&\qquad \text{where } P_{xy}(i) = \frac{h_{xy}(i)}{M^2 - h_{xy}(N)} \\
&= \frac{M^2 - h_{xy}(N)}{M^2} E_{xy} \qquad (3)
\end{aligned}
$$

Here, we define the orientation code density ρ_{xy} which means a density of effective codes in a local domain by

$$\rho_{xy} = \frac{M^2 - h_{xy}(N)}{M^2} \qquad (4)$$

Then the conditional entropy E^{ρ}_{xy} is represented as the product of the orientation code density ρ_{xy} and the entropy E_{xy}.

$$E^{\rho}_{xy} = E_{xy}\rho_{xy} \qquad (5)$$

The supremum value $E^{\rho}_{\sup}$ of the conditional entropy E^{ρ}_{xy} can be assumed to be realized when there is no futile code ($h_{xy}(N) = 0$) in the domain and we have the uniform distribution $P^{\rho}_{xy}(i) = 1/N$ ($i = 0, 1, \cdots, -1$).

$$E^{\rho}_{\sup} = -\sum_{i=0}^{N-1} \frac{1}{N} \log_2 \frac{1}{N} = \log_2 N \qquad (6)$$

Next, using a parameter $\alpha_e (0 < \alpha_e < 1)$, we calculate a richness R^{ρ}_{xy} by normalizing the conditional entropy beyond $\alpha_e E^{\rho}_{\sup}$ and

Table 2: Conditional probabilities of event I given J.

$P^{\rho}_{xy}(i|j)$		J	
		Effective code	Futile code
I	0	$\frac{h_{xy}(0)}{M^2-h_{xy}(N)}$	0
	1	$\frac{h_{xy}(1)}{M^2-h_{xy}(N)}$	0
	$\vdots$	$\vdots$	$\vdots$
	$N-1$	$\frac{h_{xy}(N-1)}{M^2-h_{xy}(N)}$	0
	N	0	1

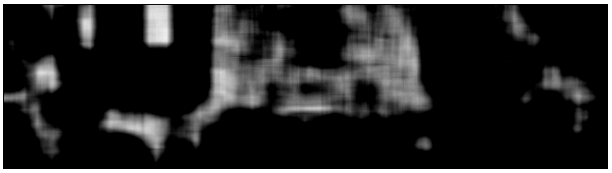


Figure 5: A richness image based on orientation code density ($M = 25$, $\alpha_e = 0.65$).

clipping the smaller values into 0 as follows:

$$R^{\rho}_{xy} = \begin{cases} \frac{E^{\rho}_{xy} - \alpha_e E^{\rho}_{\sup}}{E^{\rho}_{\sup} - \alpha_e E^{\rho}_{\sup}} & \text{if } E^{\rho}_{xy} \geq \alpha_e E^{\rho}_{\sup} \\ 0 & \text{otherwise} \end{cases} \qquad (7)$$

Figure 5 shows an example of the richness distribution on the richness image corresponding to Figure 2 with $\alpha_e = 0.65$. Since the low density in the case of Figure 2(b) such as $\rho_{500,50} = (25 \cdot 25 - 404)/(25 \cdot 25) = 0.35$, the evaluated value decreases from $E_{500,50} = 3.91$ to $E^{\rho}_{500,50} = 0.35 \cdot 3.91 = 1.38$. Otherwise, since the density in the case shown in Figure 2(a) keeps high, the value of the modified entropy is $E^{\rho}_{130,70} = 3.62$, and then Figure 2(a) is represented as bright areas in the richness image. Furthermore, because of the biased distribution of orientation codes shown in Figure 2(c), which means the brightnesses distribution in the region have not so much randomness, the evaluated value keeps low as $E^{\rho}_{500,110} = 1.90$. Figure 6 shows an example image of tag extraction which includes local maxima in R^{ρ}_{xy}. In comparison with the conventional tags without modification shown in Figure 1, the regions are not extracted, where miss-matching is apt to be occurred because of the low population of effective orientation codes.

4 SCALE ESTIMATION

In this section, we apply the orientation code density to estimating scaling of reference images. In Figure 6, we have recognized the posters, the dolls on the sofa, the ornaments on the table and the wall clock could be extracted as tags which were isolated from the background patterns. We call them isolated objects in this paper. Isolated objects have high richness, because of the diversity of orientation code distribution in their images. Tags tend to be extracted at the center of isolated objects, due to the introduction of the orientation code density. Since in tag regions the orientation codes tend to have a uniform distribution, the percentage of effective codes in the fixed sized domain or the orientation code density of isolated objects can be considered to have some relationship with scaling of the images. We find this fact can be

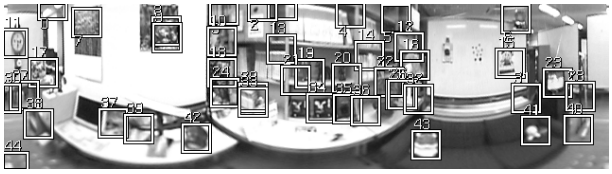
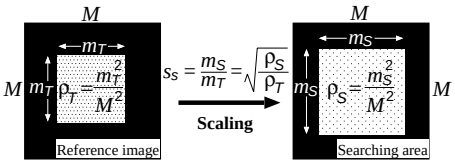
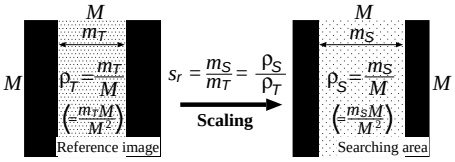


Figure 6: Tags based on orientation code entropy and density.



(a) Square model.



(b) Rectangle model.

Figure 7: Scale estimation based on orientation code density.

utilized for scale estimation. In scale estimation, we divide isolated objects into two types or models. The first model shown in Figure 7(a) is called a square model where an isolated object is regarded as a square which locates inside a tag region, while the second model in Figure 7(b) is called a rectangle model where it is regarded as a rectangle, the two short sides of which are placed outside the region. In the square model, when an image is scaled from $m_T \times m_T$ to $m_S \times m_S$, the scaling factor is represented as $s_s = m_S/m_T$. Each orientation code density is given by $\rho_T = m_T^2/M^2$ and $\rho_S = m_S^2/M^2$, and then we estimate the scaling factor s_s as

$$s_s = \frac{m_S}{m_T} = \sqrt{\frac{\rho_S}{\rho_T}}, \tag{8}$$

where ρ_T and ρ_S are given by the number of futile codes such as $\rho_T = \{M^2 - h_{xy}^T(N)\}/M^2$ and $\rho_S = \{M^2 - h_{xy}^S(N)\}/M^2$. Next, in the rectangle model, each orientation code density are given by $\rho_T = m_T M/M^2 = m_T/M$ and $\rho_S = m_S M/M^2 = m_S/M$, and then the scaling factor s_r is estimated in the same way as

$$s_r = \frac{m_S}{m_T} = \frac{\rho_S}{\rho_T} \tag{9}$$

Our scheme of scale estimation is considered as useful for estimating scaling factors for isolated objects projected as tag regions. There may be object images which can not be handled by the proposed method for scale estimation, such as the one shown in Figure 2(a), which are almost wholly filled with effective codes. Hence, in the proposed method, tags can be divided into the three groups by the orientation code density. One group is represented as a square model, the other is a rectangle model and the rest is useless for scale estimation. Figure 8 shows a selection algorithm for the useful tags. First of all we select tags $t_T(k)$ from a set of all of the tags $T_T = \{t_T(k)|k = 0, 1, \cdots, K_T - 1\}$, the orientation code density $\rho_T(k)$ of which is lower than a predefined threshold ρ_{TH}. Next, we define the orientation code density in the upper, bottom, left and right area of size $M \times M/10$ as

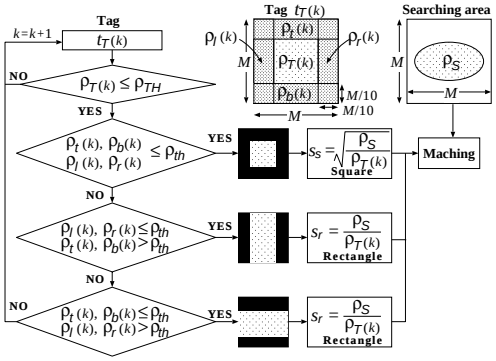
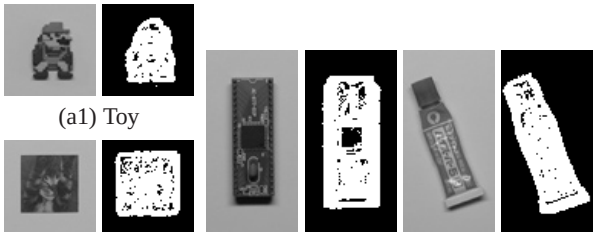


Figure 8: Selecting template images suitable for scale estimation.

$\rho_t(k)$, $\rho_b(k)$, $\rho_l(k)$ and $\rho_r(k)$. In case all of them are lower than predefined threshold ρ_{th}, the square model is applied to scale estimation and then the estimated scaling factor is used in tag search. In another case only $\rho_t(k)$ and $\rho_b(k)$ are lower or only $\rho_l(k)$ and $\rho_r(k)$ are lower, the rectangle model is applied. In the other case, the tags are useless for scale estimation. Therefore, a scaling factor estimated by another tag is used to search them.

5 FAST ALGORITHM FOR TAG SEARCH

Estimation of scaling factor is primarily introduced for making a search algorithm faster than usual, however, in this section, the reduction of computation cost can be enhanced by limiting searching areas. Let us discuss the search problem in the scene S which includes scaled instances of tags or references $T_T = \{t_T(k)|k = 0, 1, \cdots, K_T - 1\}$. In the fast tag searching, T_T is divided into two groups and then the searching area is limited. One group is $T_{T1} = \{t_{T1}(k)|k = 0, 1, \cdots, K_{T1} - 1\}$ including only isolated objects and we apply the square or rectangle models to them. The other is $T_{T2} = \{t_{T2}(k)|k = 0, 1, \cdots, K_{T2} - 1\}$, all of which are not suitable for scale estimation. In case of isolated objects, the positions of local maxima of orientation code density can not be expected to change extremely even in the presence of scale changes. Therefore, extracting isolated objects based on the orientation code entropy and density is stable to scale change of the image. This characteristic can be utilized for limiting candidate positions in pre-scanning of the scene. We extract tags $T_{S1} = \{t_{S1}(u)|u = 0, 1, \cdots, U_{S1} - 1\}$ of size $M \times M$ from S using the orientation code entropy and density and then limit searching areas $L_{S1} = \{l_{S1}(u)|u = 0, 1, \cdots, U_{S1} - 1\}$ to the neighboring area of T_{S1}. We set $l_{S1}(u)$ in this study as the area of size $M/2 \times M/2$ centered at $t_{S1}(u)$ in consideration of position errors of extracted tags. For searching $t_{T2}(k)$, a scaling factor is also necessary to match a scaled version of $t_{T2}(k)$ to S. However, from the definition, they can not be useful for scale estimation. We substitute a scaling factor $s_{\min}$ estimated by $t_{T1}(k)$ where matching error between $t_{T1}(k)$ and S obtains minimum value. Orientation code distribution in fixed size tag area changes by scale changes except for areas of isolated objects, and then different areas of size $M \times M$ can be extracted from a scaling scene image S. Therefore, for searching T_{T2}, we extract tags $T_{S2} = \{t_{S2}(u)|u = 0, 1, \cdots, U_{S2} - 1\}$ of size $s_{\min}M \times s_{\min}M$ and then limit searching areas $L_{S2} = \{l_{S2}(u)|u = 0, 1, \cdots, U_{S2} - 1\}$ to the neighboring area of T_{S2}, the size of which is $s_{\min}M/2 \times s_{\min}M/2$.



(a1) Toy

(b1) Photo (c1) IC (d1) Tube

Figure 9: Reference images and their coded images used for scaling estimation.

(a) 50[cm] (b) 55[cm] (c) 60[cm] (d) 65[cm] (e) 70[cm]

Figure 10: Scaled images for camera distances 50[cm] through 70[cm].

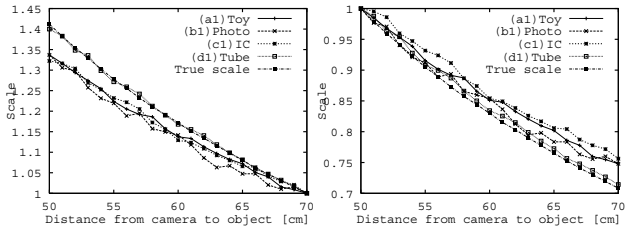
6 EXPERIMENTS

6.1 Experiments of scale estimation

In the experiments for verifying the effectiveness of estimation of scaling factors, the object distances were changed every one[cm] from 50[cm] to 70[cm]. Figure 9 shows the reference images as tags and their coded images for isolated objects such as (a1) Toy, (b1) Photo, (c1) IC and (d1) Tube at the distance of 60[cm]. Figure 10 shows extracted tags ($M = 60$) using (d1) which could been stable in detection and matching even in the presence of scale changes. In scale estimation, they were divided into two models where $\rho_{TH} = 0.8$ and $\rho_{th} = 0.4$, and then the square model was applied to (a1) and (b1), while the rectangle model to (c1) and (d1). Figure 11 shows the experimental results of scale estimation. Figure 11(a) shows the estimated scaling factors in enlargement or scaling-up, using the image at 70[cm] as a standard, while Figure 11(b) shows the ones in scaling-down with the image at 50[cm] as a standard. In each experiment using (a1) $\sim$ (d1), from the both case of scaling-up and down, the high correlations were obtained as (a1) 0.999, (b1) 0.995, (c1) 0.997 and (d1) 0.999. Figure 12 show the dissimilarities between the standard images and each experimental image. The case (a) and (c) show the results of OCM, while the case (b) and (d) show the results of the proposed scalable matching in which the estimated scaling factors are introduced into matching the scaled version of the standard images with the one of each object images. From the figures, all of the dissimilarities by the proposed method were kept lower than the results without scaling estimation.

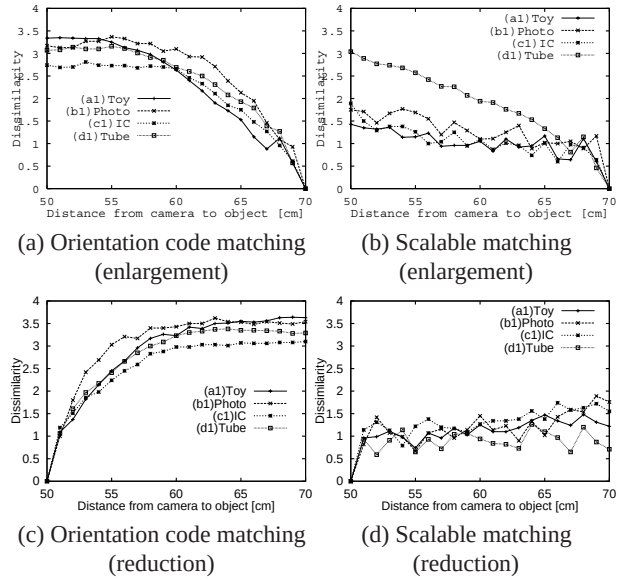
6.2 Experiments in fast tag search

Some experiments have been performed in order to compare the computation time required for searching a tag with the system of Pentium M 1.8GHz by Linux. Figure 13 shows the reference images (isolated objects) of size 60×60 pixels which are used in Subsection 6.1 and the scene image of size 640×480 pixels. The references were searched in the enlarged scene image under illumination changes. We have compared the four ways of searching as (1) Conventional search with five reference images which were scaled offline from 0.6 to 1.4 at 0.2 intervals, (2) SSDA (Sequential similarity detection algorithm)-based quick search, (3) Quick search with five reference images in limited area mentioned in Section 5, and (4) Quick search with scale estimation in limited



(a) Case of enlargement (b) Case of reduction

Figure 11: Experimental results of scale estimation.



(a) Orientation code matching (enlargement) (b) Scalable matching (enlargement)

(c) Orientation code matching (reduction) (d) Scalable matching (reduction)

Figure 12: Comparison of dissimilarity between Orientation code matching and scalable matching.

area. Table 3 shows the experimental results for evaluating the searching time. In Figure 13, there have been 29 searching areas shown as squares written by broken line. In the conventional search (1), more than 270,000 times of matching for each reference had to be calculated, while in the proposed method (3), it was reduced into 23,000 times (8.5%) because of search area limiting. As for the average searching time, it was reduced into 12.8% in (3) in comparison with (1). These experimental results show the effectiveness of the limiting algorithm. Furthermore, in the proposed method (4), since no plural reference is necessary for matching, therefore the searching time was shortened by $1 \sim 2$[sec.] compared to (3) and was reduced to 7.8% of (1). In the cases of (3) and (4), it took about 1.7[sec.] to extract tags and to introduce some constraint to the search area. However, if in the case that the whole scene is uniformly scaled, this tag extraction has to be done just once even in multiple reference searching. To take this characteristic into consideration, the searching time of (3) and (4) were reduced into 8.7% and 3.6% of (1).

Next, we demonstrate an example of reference image which does not include any isolated object. Figure 14 shows reference images of size 60×60 pixels such as (a2) Lamp, (b2) Poster1, (c2) Umbrella and (d2) Poster2 and the scene of size 640×480 pixels, which was captured in a corridor environment. This experiment was performed on the assumption that a robot tries to search landmarks in a corridor environment. A scaling factor was estimated using (a2) which includes an isolated object and then (b2), (c2) and (d2) which include no isolated object were searched using the factor. Table 4 shows the experimental results of search-

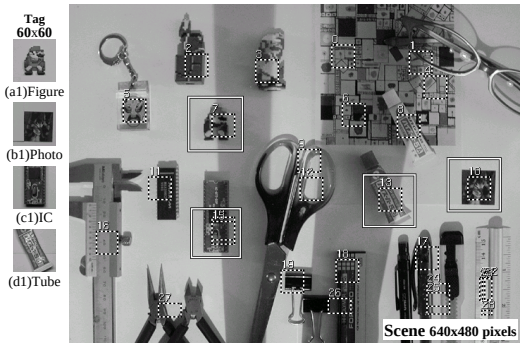


Figure 13: Reference images (a1)-(d1) and searched locations in the scene.

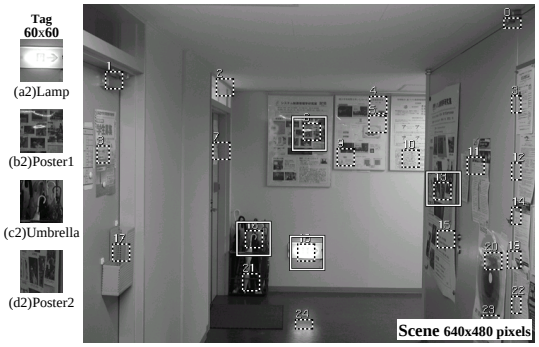


Figure 14: (a2) is a reference image which picks up an isolated object and used to estimate a scaling factor. (b2), (c2), (d2) are searched using the scaling factor in a scene image.

Table 3: Experimental results of searching time[sec.].

Searching method	Toy	Photo	IC	Tube
(1) Conventional search + Scaled offline images	25.20	33.94	35.00	32.75
(2) SSDA + Scaled offline images	17.69	22.32	25.60	24.11
(3) Quick search + Scaled offline images	3.51	3.83	4.49	4.25
(4) Quick search + Scale estimation	2.42	2.34	2.48	2.48

Table 4: Searching time[sec.] using the scene shown in Figure 14.

Searching method	Lamp	Poster1	Umbrella	Poster2
(1) Conventional	30.70	42.83	44.02	44.77
(2) SSDA	15.57	24.96	27.81	31.82
(3) Quick search	2.20	2.68	2.82	3.29 fault
(4) Quick search + Scale estimation	1.80	1.35	1.40	1.39

ing time. However, in the searching method (3), miss-matching was occurred at (d2). Different tag areas were extracted from the scene in the presence of scale changes because of the utilized tag size of 60×60 and did not include the true position of (d2). In this experiment, the scaling factor between the reference images and the scene was 0.76, and the estimated one by the rectangle model using (a2) were 0.71. In the case of (4), the tag size of $0.71 \cdot 60 \times 0.71 \cdot 60 = 43 \times 43$ pixels was used for extracting the other tag areas. There were 25 broken line squares in Figure 14 showing the limited searching areas. The average searching time of (4) were reduced into 5.2% of (1) and in the case of multiple reference searching it was 2.1%.

7 CONCLUSIONS

The new approach has been proposed to estimate scaling factors between reference images and the scene. Orientation code density has been defined to design the algorithm for the scale estimation. Some experiments with real world images have been performed and the effectiveness of the proposed estimation of scaling factors has been verified. Furthermore, the fast search method has been presented, which uses limitation of searching domain in the presence of scale changes and some experiments have been done in order to compare the variant combinations of conditions in their computation times. The experimental results have shown that the proposed method has reduced the computational time into less than 3.6% compared to the conventional searches. As future problems, fast schema for the scale estimation has to be designed for the case of independent images with different scalings in the scene.

REFERENCES

C.Harris and M.Stephens, 1988. A combined corner and edge detector. In: Proc. 4th Alvey Vision Conf., pp. 147–151.

C.Tomasi and T.Kanade, 1991. Detection and tracking of point features. In: CMU Tech. Rep. CMU-CS-91-132.

D.Ballard, 1981. Generalizing the hough transform to detect arbitrary shapes. Pattern Recognition 13(2), pp. 111–122.

D.M.Zhao and J.Chen, 1997. Affine curve moment invariants for shape recognition. Pattern Recognition 30(6), pp. 895–901.

F.Ullah and S.Kaneko, 2004. Using orientation codes for rotation-invariant template matching. Pattern Recognition 37(2), pp. 201–209.

F.Ullah, S.Kaneko and S.Igarashi, 2001. Orientation code matching for robust object search. IEICE Trans. on Inf. & Sys. (8), pp. 999–1006.

H.Moravec, 1977. Towards automatic visual obstacle avoidance. In: Proc. 5th Int. Joint Conf. Art. Intell., p. 584.

J.Shi and C.Tomasi, 1994. Good features to track. In: Proc. IEEE Conf. Comput. Vision Patt. Recogn., pp. 593–600.

S.Kaneko, H.Ohmachi and T.Honda, 1992. Introducing structurization of edge segment groups based on prominency entropy to binocular depth calculation. IEICE Japan J75-D-II(10), pp. 1649–1659. (in Japanese).

S.Kaneko, I.Murase and S.Igarashi, 2002. Robust image registration by increment sign correlation. Pattern Recognition 35(10), pp. 2223–2234.

S.M.Smith and J.M.Brady, 1997. SUSAN-a new approach to low level image processing. Int. J. of Comput. Vision 23(1), pp. 45–78.

T.Mori, Y.Matsumoto, T.Shibata, M.Inabe and H.Inoue, 1995. Trackable attention point generation based on classification of correlation value distribution. In: Rroceedings of JSME Annual Conference on Robotics and Mechatronics(ROBOMEC '95), pp. 1076–1079. (in Japanese).