MVA2000 IAPR Workshop on Machine Vision Applications, Nov. 28-30, 2000, The University of Tokyo, Japan

7—2

Robust Image Registration Using Selective Correlation Coefficient

Yutaka Satoh, Shun'ichi Kaneko and Satoru Igarashi *
Faculty of Enginnering, Hokkaido University

Abstract

We propose a new method for robust image registration called 'Selective Correlation Coefficient (SCC)' in order to search images under ill-conditioned illumination or partial occlusion. A correlation mask image is generated for selecting pixels of a image before matching. The mask image can be derived from a binary-coded increment sign image defined from any object image and the template image. The masking rate of occluded regions is theoretically expected to be 50%, while unoccluded regions have much lower rate than 50%. We realize robustness under ill-conditioned environment since inconsistent brightness of occluded regions can be omitted by the masking operation. We propose, furthermore, the mask expanding procedure to get more stable robustness. The effectiveness of probabilistic masking increases by the procedure, resulting the rate around 70% of masking of occluded regions. This paper includes a theoretical modeling and analysis of the proposed method and some experimental results with real images.

1 Introduction

A new method for image registration or image matching based on a halftone image called 'Selective Correlation Coefficient (SCC)' is proposed in order to obtain robustness against ill-conditioned illumination and partial occlusion. In pattern recognition, image registration techniques to judge image similarity based on a rational quantitative measure is one of the oldest and fundamental image processing techniques. These construct an important foundation that should be a kernel of image processing in wide fields such as the assembly lines in FA, industry applications that have been oriented toward automation such as a robot vision for visual inspection, some environment monitoring systems such as in roads and parking lots, and general-purpose image-based measurement systems. High-speed and reasonable cost PCs and image processing hardware can be recently of popular utilization. And then rather expensive algorithms for image processing with a lot of computation can become feasible even in real world applications. Sum of square difference (SSD), correlation coefficient (CC), increment sign correlation (ISC) have been proposed as template matching techniques so far. CC has been used widely because it can cope with noisy images and change in brightness of objects. But, when the scene involves partial occlusion, highlight, the brightness values in the object region, this product-based correlation technique sometimes fails to get good results, while it is good at measuring some slight change in brightness. SSD evaluates a sum of square difference in brightness. Though it is an efficient algorithm, it is very weak against change in brightness and occlusion. ISC was developed with the aim of the image registration method which can cope with shadowing, occlusion, highlight. This is a method based on the tendency-based similarity measure which is evaluated by aggregating increment signs at each pixel. ISC is robust and efficient. But, due to information reduction by neglecting brightness values, there have been some problems such that sizes of templates cannot be reduced and it cannot discriminate slight change in brightness of images. In this paper, a novel approach for these problems is proposed. The proposed method is designed along the same approach as CC with extension by ISC which is so robust to the ill-conditions abovementioned. In this method, an increment sign (IS) image transformed from the original image is used for masking irrelevant portion to a template image. Because of the statistical nature of IS, any parameters are not needed to realize the masking operation. The masking operation can automatically select consistent part to the template and consequently decrease wrong influence to taking cross correlation between the template and an object image, and then it becomes possible to find defective pixels in the object image.

2 Selective correlation coefficient

2.1 Definition

Fig. 1 shows the overview of the proposed mechanism for robust image registration. It generates a correlation mask for each pixel before calculating correlation. The mask can be derived from the binary-coded 'Increment Sign (IS)' between a object and a template image. Then correlation is calculated from not only the template and the object

---

* Address: Kita-13 Nishi-8 Kita-ku, Sapporo 060-8628 Japan. E-mail: `satoh@mee.coin.eng.hokudai.ac.jp`

image but also the mask image based on IS which can be assumed as pre-judgement of consistency between matched pixels. Pixels in the mask image take value 0 or 1 and consequently they can be used to choose which brightness is adopted for similarity evaluation between two images. Simple subtraction and checking their absolute difference causes pseudo masks corresponding to brightness change in object images due to illumination change which occurs often in real situations. This is one of our motivations of this work.

The conventional correlation coefficient (CC) or the normalized cross-correlation function is considered first. A one-dimensional ordered list into which a two-dimensional image is rearranged in some order is utilized here for simplicity. Images are generally and often decomposed into each row or column and then stacked in their order. Let $F = \{f_n\}_{i=0}^{N-1}$ and $G = \{g_n\}_{i=0}^{N-1}$ define a template image and an object image which is of the same size N as the template and extracted from the scene. CC can be expressed by the formula as,

$$r_{cc} = \frac{\sum_{n=0}^{N-1}(f_n - \bar{f})(g_n - \bar{g})}{\sqrt{\sum_{n=0}^{N-1}(f_n - \bar{f})^2}\sqrt{\sum_{n=0}^{N-1}(g_n - \bar{g})^2}} \quad (1)$$

where $\bar{f}$ and $\bar{g}$ mean the averages of them.

SCC is defined by the next equations that extend the above expression.

$$r_{scc} = \frac{\sum_{n=0}^{N-1} c_n(f_n - \bar{f})(g_n - \bar{g})}{\sqrt{\sum_{n=0}^{N-1} c_n(f_n - \bar{f})^2}\sqrt{\sum_{n=0}^{N-1} c_n(g_n - \bar{g})^2}} \quad (2)$$

In this expression, brightness values from the template and the object are multiplied through the coefficient c_n, and then the sum is normalized by standard deviations of each image in which the coefficients are introduces as well as in the numerator.

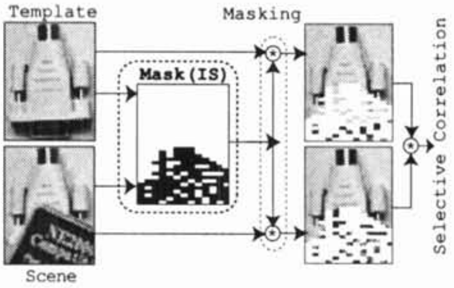


Figure 1: Overview of method.

The mask coefficient c_n is defined by

$$c_n = \begin{cases} b_n \cdot b'_n + (1 - b_n) \cdot (1 - b'_n) & (n = even) \\ c_{n-1} & (n = odd) \end{cases} \quad (3)$$

where the increment sign b_n for F is defined by[6]

$$b_n = \begin{cases} 1 & (f_{n+1} \geq f_n) \\ 0 & (\text{otherwise}) \end{cases} \quad (4)$$

and b'_n is also defined for G similarly. Increment signs b_n and b'_n are binary digit which show a tendency of change in brightness in the neighborhood at each pixel. The value 1 and 0 correspond to non-negative and negative signs of increment between adjacent pixels.[1] The mask coefficient c_n is constructed through the rule that it takes 1 if b_n and b'_n are coincident and 0 otherwise. Consequently c_n can mask pixels with probable inconsistency out of the calculation of correlation. Two adjacent coefficients have the same value according to the definition. As for the pixel of the basic mask image, 2 pixels that adjoin it are set up in the same value by the definition. This is because of the area of defining b_n and b'_n each of which needs two original pixels. Increment sign correlation (ISC) [6] has been proposed for robust image registration in case of occlusion and/or illumination change. The correlation function is known as a statistics that have a binomial or normal distribution under a reasonable approximation. From statistical investigation, ISC has the following fundamental characteristics:

- The deviation is generally small.
- The mean is the probability of sign reversal between b_n and b'_n.
- It has an expected value 0.5 for uncorrelated images.
- It has an expected value more than 0.5 for correlated images, which depends on an amount of noise included in the scene.

These statistical characteristics of the image correlation based on increment signs guarantee the performance of the basic mask image M which is expected to exclusively choose probably consistent pixels and omit inconsistent ones. The masking rate of occluded regions is theoretically proved to be expected as 0.5 because of ISC characteristics for uncorrelated images. On the other hand, for unoccluded regions, it can have the possible masking rate below 0.5 in case of correct pair of images.

2.2 Calculation cost

The calculation cost of SCC is examined in comparison with CC. From Equations (1) and (2), SSC

---

[1] It is possible to call them as signs of spatial difference, but in this paper, we call increment signs in order to appeal they can be defined for even indifferentiable signal.

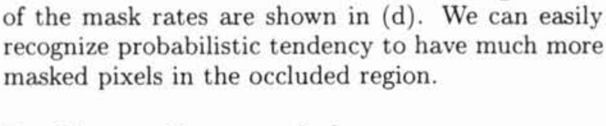
needs summations of three-term products while CC has ones of two-terms. Because that c_n has 1 or 0 values in nature in these summations, the cost of multiplication is not essentially increased. SSC has a rather small number of products since it masks pixels in some degree. The prominent issue of computational costs is the one for variance computations. In CC, the variance or the self-correlation of the template can be calculated only once in offline manner for it does not change throughout the calculation of a CC, but in SSC, three correlations should be calculated every time. SCC needs an extra calculation cost for obtaining signs c_n at every position. The cost for it, however, can be small because it depends on comparing operations between two brightness values.

2.3 Characteristics of basic mask image

Fig. 2 shows an template image in (a) and the similar object image with occlusion in (b), and they generate a basic mask image in (c) in which two kinds of mask rates can be computed as shown in (d). The scene in (b) includes about 24% occlusion region at the lower right corner. In the mask image shown in (c), pixels in black show masked pixels of the value 0, and pixels in white unmasked pixels of the value 1, which should be used in computation of the correlation value. And then the average values of the mask rates are shown in (d). We can easily recognize probabilistic tendency to have much more masked pixels in the occluded region.

3 Expanding mask image

3.1 Expanding procedures



(a)Template (b)Scene (c)Mask (d)Mask rate

Figure 2: An example of correlation mask.

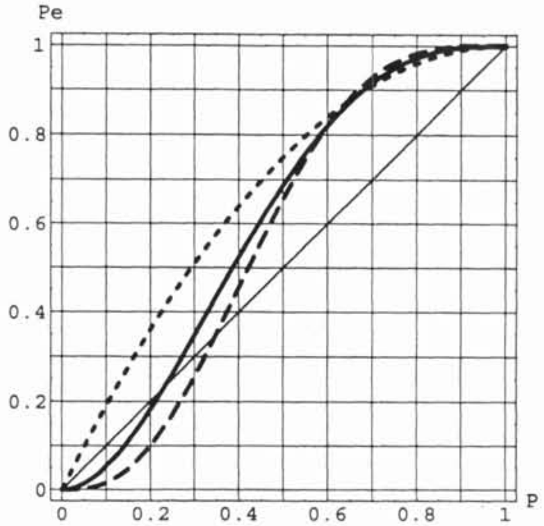


Figure 3: Mask extension curves.

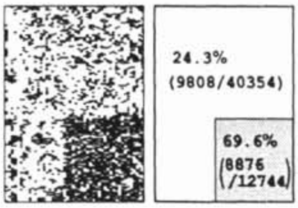


Figure 4: Extended mask image.

As shown in Fig. 2, in real images, the mask rates of unoccluded regions are generally more than 0 because of quantization and additive noise introduced by an imaging process and illumination fluctuation. On the other hand, in occluded regions, since the nature of correlation is weak, influence of noise is small and the mask rates are around 0.5 as expected. When noise is lager, however, the difference of mask rates between the regions become small and then should be improved by an expanding procedure. The following procedure for mask expansion is introduced. The procedure is fundamentally based on a majority rule: there are two groups of masked pixels and unmasked ones, and the neighborhood of a pixel of interest on a basic mask image, which is defined in the previous section, has its majority. The minority pixels should be changed to $c_n = 1$ or 0 according to the value of the majority pixels. Fig. 3 illustrates expanding characteristics according to two(dotted line), four(solid line) and six(dashed line) pixel neighborhood, respectively. Let p and p_e be a mask rate of a basic mask image and one of an extended mask image, respectively. p_e can be obtained, for example for the four pixel rule, as

$$p_e \quad = \quad 1-(1-p)^4-4(1-p)^3 \cdot p \qquad (5)$$

We adopt the four-pixel version of the majority rule for the mask expansion of basic mask images in this paper. Using this rule, the mask rate increases by around 20% in occluded regions and the one in unoccluded regions is almost unchanged for the example shown in Fig.4, where p_e is expected to change from 0.223 to 0.217 in unoccluded regions and form 0.5 to 0.688 in occluded ones according to Equation (5). Actually, the corresponding mask rates are observed as 0.243 and 0.696.

3.2 Characteristic about the occlusion rate

Tendency of correlation in case of occlusion is investigated. The template image and the scene that were used in Fig. 2 were exchanged for generality. Occlusion rate was changed form 0% to 100%. Fig. 5 shows the tendency correlation versus the occlusion rates. We use SCC, CC, and SCC+EM for the selective correlation coefficient with the extended-mask image. SCC+EM, SCC and CC have the similar value of -0.3 at 100% occlusion rate, and so the original images can be considered to be uncorrelated. When the occlusion rates increase, SCC+EM always holds the highest correlation and SCC is also maintaining high correlation compared with CC. The rate of decrement for both SCC and SCC+EM decrease slowly as increasing occlusion rates. Compared with SCC and CC, SCC+EM is very robust. We call hereafter SCC+EM by SCC for simplicity.

4 Experiments

4.1 Occlusion

The effect of SCC in the real scene containing occlusion is investigated. Fig. 6 shows the result of an experiment. Occlusion rates increase from the left to the right images. SCC detected proper positions while CC failed because of the occlusion rates around from 20% to 40%. In (b), (c) and (d), CC failed to report the correct positions as shown in the figure. Also shown in the right upper corner in each of the scenes, SCC could make the extended mask images that could probably capture the correct occluded parts. Table 2 shows the actual values of correlation for SCC and CC, where SCC could retain reasonable values for matching while CC suffered in some cases and lost the correct positions. Fig. 7 indicates the profile of the correlation values in the neighborhood of the proper position in Fig.6(b). Fig. 7(a) shows the profile of SCC, which the prominent maximum peak could be recognized at the correct position, while in (b) the profile of CC also seemed to have such kind of peak at the same part, however, it was not the maximum over the scene. Since it has been known that ISC has the similar characteristics to SCC [6], SCC is expected to have a neutral nature between ISC and CC. Thus, we can utilize SCC in the case that the scene has occlusion and simultaneously we want to test the brightness of the objects which keeps even important information for matching. The average calculation time of SCC and CC were about 823 and 844 seconds, respectively.

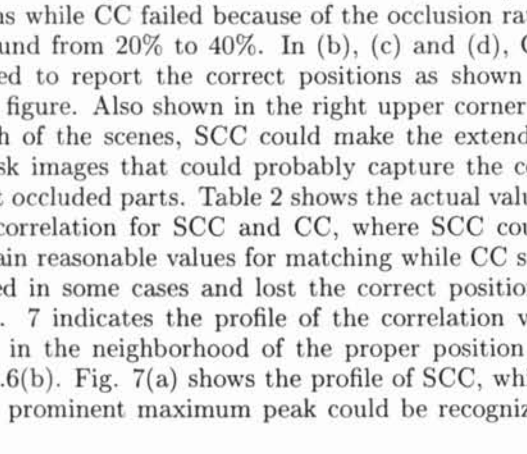


Figure 5: Correlation versus occlusion rates.

Table 1: Specifications of experiments.

Brightness data	8bit grayscale
TP size	120 × 92
Scene size	640 × 320

Table 2: Correlation values.

	(a)	(b)	(c)	(d)
SCC	0.82	0.58	0.58	0.70
CC	0.69	0.35(0.42)	0.36(0.42)	0.39(0.42)

4.2 Highlight

Fig. 8 is an example image of a glossy package including highlight. The highlight or saturation by specular reflection of illumination can be often occurred in the real world scene, for example, many wrapped industrial products. Highlight scenes can be dealt with by SCC as well as in the case of occlusion. Results of SCC and CC were shown by black frames in Fig. 8(a),(b). In the both figures, SCC could report the correct objects while CC reported the wrong portion with the similar distribution of brightness to the one of the template. ISC is known to have robustness as well as SCC has [6], but it sometimes fails to search the correct instances of the template in the case of few gray level images like binary images as shown in Fig. 8. In such cases, SCC can be expected to have much more robustness because it is based on both robust codes of brightness increment and brightness itself. The average calculation time of SCC was about 21 seconds, and the one of CC was about 24 seconds.

5 Conclusions

An image matching method by the selective correlation coefficient called by SCC, which has the nature of robustness against occlusion and ill-conditioned illuminations, has been proposed. By using basic mask images based on increment signs, brightness based correlation is effectively improved, reducing influence of occlusion. Furthermore, a simple procedure for expanding a basic mask image has

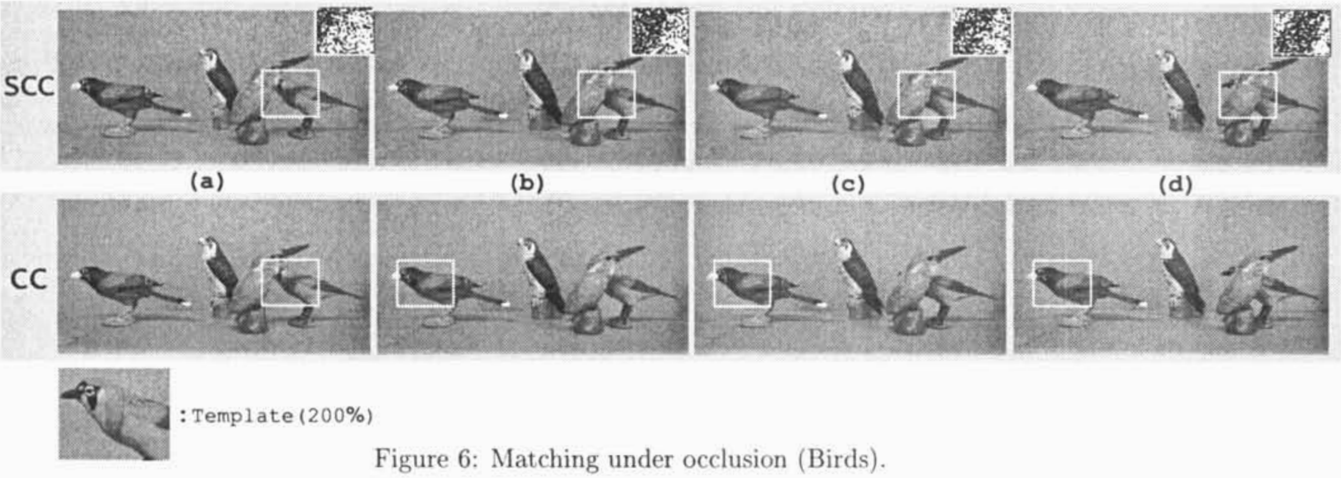


Figure 6: Matching under occlusion (Birds).

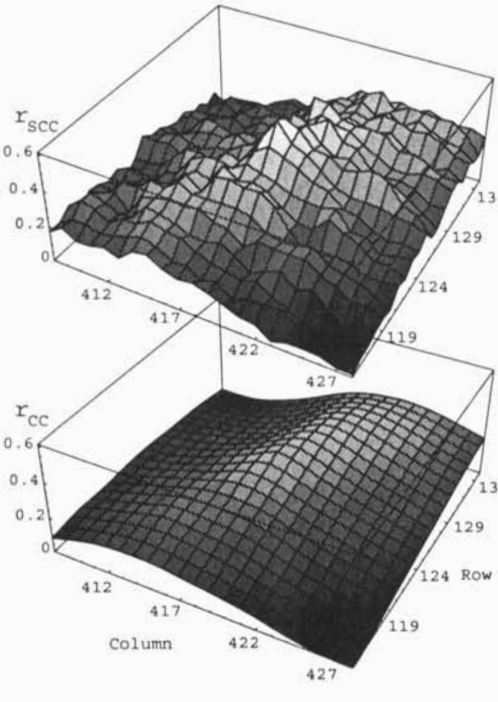


Figure 7: Profiles of correlation values.

(a)Label1

(b)Label2

Figure 8: Matching of glossy object.

also been proposed. It enables us to enhance the difference between mask rates in occluded regions and unoccluded ones which is able to estimate the degree of enhancement by use of a mathematical model. The calculation cost of SCC is almost the same as the one of CC even if it includes a step of generating mask images. By the experiments with the real images, the good performance on robustness against occlusion and highlight was observed.

References

[1] A.K.Jain, R.P.W.Duin and J.Mao: "Statistical Pattern Recognition: A Review", IEEE Trans., Vol.PAMI-22, No.1, pp.4-37, 2000.

[2] J.C.Russ, ed.: "The Image Processing Handbook", CRC Press, Tokyo, 1994.

[3] L.G.Brown: "A Survey of Image Registration Techniques",ACM Computing Surveys, Vol.24, No.4, pp.325-376,1992.

[4] J.K.Aggarwal, L.S.Davis and W.N.Martin: "Correspondence processes in dynemic scene analysis", Proceeding of IEEE, Vol.69, No.5, pp.562-571,1981.

[5] D.I.Barnea and H.F.Silverman: "A Class of Algorithms for Fast Digital Image Registration", IEEE Trans. Computer, Vol. c-21, No.2, pp.179-186, 1972.

[6] I.Murase, S.Kaneko, S.Igarashi: "Robust Matching by Increment Sign Correlation", IEICE Trans.(D-II), Vol.83-D-II, No.5, pp.1323-1331, 2000. (in Japanese)