APPLICATION OF ORIENTATION CODE MATCHING FOR STRUCTURE FROM MOTION

T. Anai [a] *, N. Fukaya [a], N. D'apuzzo [b], N. Kochi [a]

[a] Imaging and Measuring Laboratory, R&D Center, Topcon Corporation,75-1 Hasunuma Itabasi Tokyo, Japan
[b] HOMOMETRICA CONSULTING, Culmannstrasse 59, CH-8006 Zurich, Switzerland
t.anai@topcon.co.jp

Commission V, ICWG V/I

KEY WORDS: Feature, Matching, Extraction, Video, Exterior Orientation, Bundle, GPS

ABSTRACT:

In recent years, in order to perform the 3D measurement using video image sequences from the freely moving plat-form, the exterior orientation technique that uses video image sequences has widely been investigated in the field of computer vision. This problem is called as “Structure from Motion” or “Simultaneous Localization and Mapping”. In these techniques, the estimation of exterior orientation is performed by tracking feature points extracted from video images sequences, and the exterior orientation parameters can be obtained continuously. Therefore, robustness of tracking of feature points often becomes important problem.
In this paper, we describe about an application of the robust template matching based on "Orientation Code” image processing algorithm for the natural feature point tracking procedure in Structure from Motion.

1. INTRODUCTION

Generally, In order to perform exterior orientation from video image sequences by using “Structure from Motion” (SfM) technique in wide area applications such as mobile mapping and real-time machine control of construction equipment, many problems must be considered. The typical SfM technique using video image sequences estimates exterior orientation by tracking natural feature points extracted from video image sequences, and exterior orientation parameters can be obtained continuously. However, if the information of global coordinate is not given, the exterior orientation parameters are obtained in model space which has arbitrary position, scale and rotation. Another important problem is that estimated parameters generally include the accumulative error based on tracking error. Moreover, the tracking of natural feature points often becomes fail by change of environment. One solution for these problems is combination with another positioning sensor. In order to obtain the exterior orientation parameter of video camera in global axis without accumulative error, the RTK-GPS system is widely used in many applications of SfM technique. An important problem of RTK-GPS system is that the accuracy of RTK-GPS positioning is depending on GPS satellite position and condition of the environment. RTK-GPS systems can perform RTK-fixed solution positioning only at the sky opened area. Therefore, the positioning data of RTK-GPS generally include RTK-float solution positioning data due to the integer ambiguity is not obtained correctly. It should also be noted that GPS positioning data of RTK-float solution often include large outliers that cannot be judged as outliers only from the GPS accuracy information. Also, RTK-GPS positioning can’t perform in much situation at city road.
For these problems mentioned above, we have already proposed an Extended Bundle Adjustment method for video image sequences that uses both SfM technique and RTK-GPS data with considering accuracy of RTK-GPS in previous research (T. Anai et al. 2009). This proposed Extended Bundle Adjustment method minimizes both re-projection error for video images and positioning error for RTK-GPS simultaneously. The outlier in RTK-GPS positioning data is removed by robust regression method that embedded in proposed Extended Bundle Adjustment method. In opposite, accumulative errors of SfM technique caused by tracking errors are suppressed by using RTK-GPS data. However, robustness of natural feature points tracking in the SfM technique is still important issue because our proposed method can't estimate exterior orientation parameters correctly when natural feature points tracking becomes fail. The most important point of robust tracking procedure is robustness of matching algorithm for the changing of video image condition, such as brightness change or occlusion of object.
On the other hand, In order to perform robust template matching in real-time applications, Ullah and Kaneko have suggested “Orientation code matching” (OCM) algorithm in 2001. In this algorithm, the “Orientation Code” (OC) image has been used instead of the original gray-scale images of video image sequence. Takauji et al. have investigated the natural feature point extraction algorithm from OC image called “Orientation Code Richness” (OCR) in 2005. Moreover, Kitamura et al. have investigated the application of OCM and OCR for the automatic stereo matching procedure on human body measurement system in 2009. They have evaluated the robustness and processing speed between OCM and other matching algorithm such as SSDA and NCM, and they have concluded that OCM have the better robustness and processing speed.
From these circumstances mentioned above, we have been investigating the robust natural feature point tracking method for the SfM procedure which based on OCM and OCR algorithm. Moreover, the application of the SfM procedure based on Extended Bundle Adjustment under the outdoor environment has investigated.
This paper is structured as follows. Section 2 describes OC image processing. Section 3 describes SfM procedure based on Extended Bundle Adjustment that using OC image processing. Some example of OC image processing for SfM procedure is also described. Finally, Section 4 describes conclusion and future work.

2. OC IMAGE PROCESSING

This section describes about “Orientation Code” (OC) image processing. The most unique point of OC image processing is that the coded images are used instead of the original gray-scale images of the video sequence. Details of OC image processing are as follows.

2.1 Orientation Code (OC)

In the OC image processing, every pixel in a gray-scale image (with values ranging from 0 to 255) is coded into a number of Orientation Codes (OC). The orientation codes refer to a quantization of the direction of the maximal intensity change of the pixel in the neighborhood around interest pixel. The OC of each pixel is defined as following equation:

$$c_{xy} = \begin{cases} \left[\dfrac{\tan^{-1}\left(\dfrac{\Delta I_x}{\Delta I y}\right)}{\Delta\theta} \right] & if \quad |\Delta I_x| + |\Delta I_y| \geq \gamma \\ N = \dfrac{2\pi}{\Delta\theta} & otherwise \end{cases} \tag{1}$$

In this equation, ΔI_x and ΔI_y show horizontal and vertical gradient of pixel (x, y). We use Sobel operator for calculating gradient of each pixel. N is the quantization level of direction. N is set to 16 typically (Figure 1). The "γ" is the threshold value for the suppressing of small gradient pixel. Therefore, if gradient value is larger than γ, OC code is obtained as 0~15 values. Otherwise, OC code is 16 that mean unreliable code.

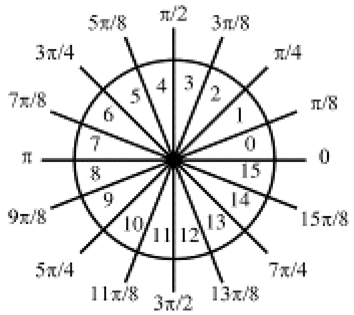
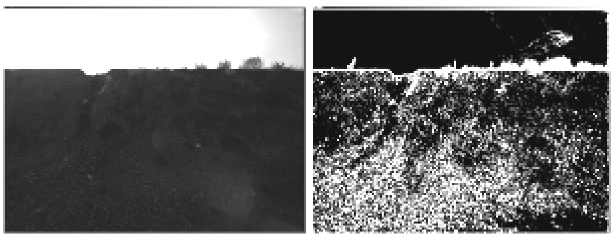


Figure 1. Orientation Code (N = 16)

Figure 2 shows the example of OC image. From this image, it is understood that the OC image is independent from the changing brightness and also OC image describe the important feature of image.



a) Grayscale Image b) OC image

Figure 2. Orientation Code Image

2.2 Orientation Code Richness (OCR)

In order to extract natural feature points from OC image, we use Orientation Code Richness (OCR) that extracts the pixels which have high entropy of OC. The high entropy of OC means the important texture for template matching.

In order to evaluate the entropy of each pixel on OC image, we calculate the relative frequency from the local limited area of the pixel size M-by-M region centered at the interest pixel (x, y) as follows,

$$P_{xy}(i) = h_{xy}(i)/M^2 - h_{xy}(N) \tag{2}$$

Where $h_{xy}(i)$ (i =0,1,,,,,,N-1) means frequency of OC of M-by-M pixel size region. Because the value of the $h_{xy}(N)$ is the unreliable code, we leave it out from the relative frequency. As a result, the entropy of OC, 0 ~ N-1, is shown by the following equation.

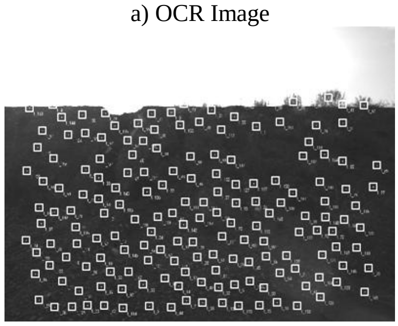
$$E_{xy} = \sum_{i=0}^{N-1} P_{xy}(i) \log_2 P_{xy}(i) \tag{3}$$

When each OC goes with uniform distribution P_{xy}(i)=1/N, the maximum value of entropy E_{max} is log2N. Thus the Richness R_{xy} is defined as

$$R_{xy} = \begin{cases} \dfrac{E_{xy} - \alpha_e E_{\max}}{E_{\max} - \alpha_e E_{\max}} & if \quad E_{xy} \geq \alpha_e E_{\max} \\ 0 & otherwise \end{cases} \tag{4}$$

Where, the threshold value α_e is defined to remove low entropy area. We made OCR image by calculating the richness of the entire picture and covered it with grids of the predetermined size, and detected the highest Richness of each grid as the natural feature points. Figure 3 a) shows the OCR image and Figure 3 b) shows the detected natural feature points (171 points).

a) OCR Image



b) Natural Feature Points by OCR (171 points)

Figure 3. Natural Feature Points by OCR

Figure 4 shows the result of the HARRIS operator for original gray-scale image as an example of other feature point detector (51 points). From these images, it is understood that the detection of natural feature points from the low contrast area by the HARRIS operator is difficult, but can find the good natural feature points from the OC image by OCR because the OC image is independent from intensity.

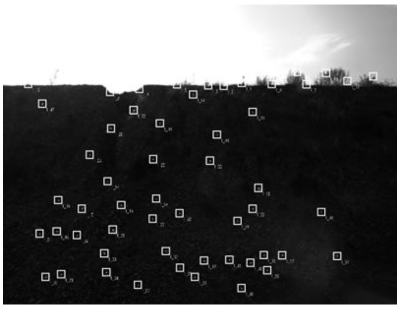
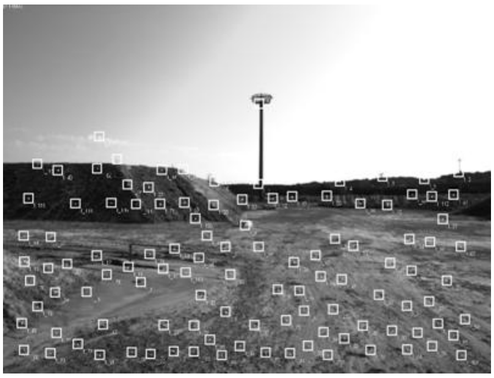


Figure 4. Natural Feature Points by HARRIS (51 points)

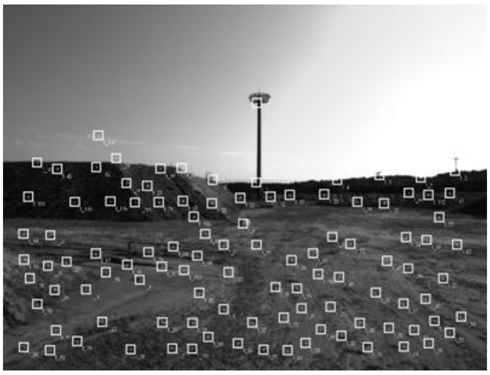
2.3 Orientation Code Matching (OCM)

The OCM process is similar to other simple image based template matching algorithms. The difference inside template image patches is considered as a measure to minimize in order to determine the best match. The major difference to other algorithms is that in the case of OCM, OC images are employed instead of the original gray scale images.

More in detail, the difference between a template image patch from OC image (O_t) and search OC image (O_i) is defined as following equation.

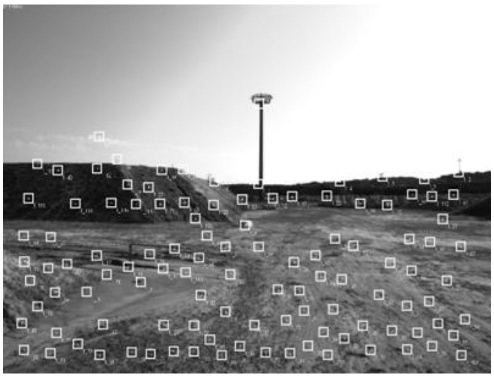


a) Start Frame



b) Next Frame

Figure 5. Matching Result of OCM



a) Start frame

b) Next frame

Figure 6. Matching result of SSDA

$$D = \frac{1}{M}\sum_{M} d(O_t, O_i)$$
$$d(a,b) = \begin{cases} \min\{|a-b|, N-|a-b|\} & if \quad a \neq N, b \neq N \\ N/4 & otherwise \end{cases} \tag{5}$$

where D = difference between O_t and O_i
M = Size of template image patch
d = difference between a pixel and b pixel

Figure 5 shows the example of OCM. Figure 6 shows the result of SSDA template Matching using Gray Scale image. Because the sun was covered with a cloud to test image of Figure 5 and 6, there is a brightness change between a start and the next frame, but the camera position does not move. Therefore, the correct correspondence points are natural feature points that do not move on next frame. From Figure 5 and 6, it is understood that the OCM perform robust template matching under the brightness changing condition.

3. SFM PROCEDURE USING OC IMAGE PROCESSING WITH EXTENDED BUNDLE ADJUSTMENT

This section describes an application of OC image processing for SfM procedure based on Extended Bundle Adjustment.

Figure 7 shows the main flow of SfM procedure of this investigation. The SfM procedure of this investigation consists from following steps.

1. At the first, gray scale image from video image sequences is converted to the OC image.

2. Many natural feature points are extracted from OC image by using OCR processing.
3. Natural feature points are tracked for successive image frames by using OCM template matching.
4. Extended bundle adjustment is performed by using both natural feature points and the RTK-GPS positioning data.
5. Outliers in tracking result of natural feature points and RTK-GPS positioning data are eliminated by using M-estimator.

By iterating theses processes from the first frame to the last frame, exterior orientation parameter of each video frames and 3-D coordinate of every natural features points are determined.

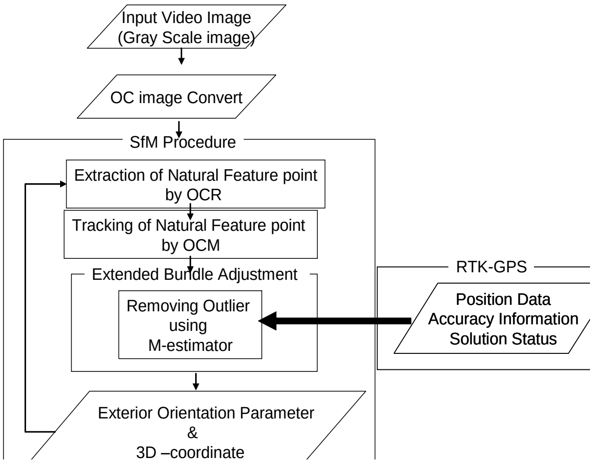


Figure 7. Main flow of SfM Procedure

3.1 Extended Bundle Adjustment

The exterior orientation parameters for each video frames and 3D coordinate of every natural feature points from SfM procedures are estimated as the bundle adjustment by minimizing re-projection error of natural feature points. When the RTK-GPS positioning data is obtained the 3D coordinate of camera position for each video frames are also obtained from RTK-GPS positioning data. The position differences between SfM and RTK-GPS positioning data are defined as the weighted error function that considering the accuracy of RTK-GPS positioning data. Therefore, the extended bundle adjustment is performed by minimizing both re-projection error for video image and positioning error for RTK-GPS simultaneously as the following equation.

$$E' = \sum_f \sum_p W_{fp} \left| x_{fp} - x'_{fp} \right| + \sum_{f'} W_{f'} \left| G_{f'} - G'_{f'} \right| \quad (6)$$

Where
f'= f'-th frames obtained RTK-GPS positioning data
$W_{f'}$= weighting coefficient of reliability for RTK-GPS positioning data at f'-th frame
$G_{f'}$=3-D coordinates of f'-th frame obtained from RTK-GPS
$G'_{f'}$= 3-D coordinates of f'-th frame obtained from SfM

3.2 Removing Outlier

Generally, tracking result of natural feature points and RTK-GPS positioning data include outlier. In this investigation the outlier in these observation data are removed by M-estimator embedded in Extended Bundle Adjustment.
The M-estimator is defined as the least-square method that the each weighting coefficient is decided to suppress the outlier. Weighting coefficient of M-estimator w^{eff} is obtained from following equation.

$$w^{eff} = [\psi(z)/(z)]w$$
$$z = v/\sigma \quad (7)$$

Where w = weighting coefficient of least-square method
v = residual error of each observation equation
σ = RMS of each observation equation.
z = normalized residual error
ψ = influence function

The influence function ψ of this research is Tukey's Biweight that defined as follows.

$$\psi(z) \equiv \begin{cases} z\left[1-(z/c)^2\right]^2 & \text{when} \; |z| < C \\ 0 & \text{when} \; |z| \geq C \end{cases} \quad (8)$$

The constant C in equation (8) is selected as 5~9.
Therefore, the removing outlier using M-estimator is performed as the minimizing the following error equation.

$$E'' = \sum_f \sum_p w^{eff}_{fp} \left| x_{fp} - x'_{fp} \right| + \sum_{f'} w^{eff}_{f'} \left| G_{f'} - G'_{f'} \right| \quad (9)$$

Where w_{fp}^{eff} = Weighting coefficient of M-estimator for re-projection error
$w_{f'}^{eff}$ = Weighting coefficient of M-estimator for positioning error

3.3 Experiment

In order to estimate the robustness of SfM procedure using OCM and OCR under the outdoors environment, the experiment using outdoor test field have been performed. Figure 8 shows the test field of this experiment. Figure 9 shows the Test system of this experiment. The video camera and RTK-GPS antenna are fixed on the vehicle as one body. The specification of video camera and RTK-GPS are shown in Table 1 and 2.
This Test filed is sky opened area. Thus, almost RTK-GPS positioning data of this experiment are acquired as the RTK-Fix solution data. The test vehicle ran around 200m construction earth at about 10km/h speed while watching a wall of the construction earth in a video camera on the test vehicle.
Figure 10 shows the example of video image sequence of this investigation. As seen in Figure 10, brightness and contrast of image is no good condition. Figure 11 shows the OC image of Figure 10. Also, Figure 12 shows the tracking result of natural feature points using OCM and OCR. As mentioned above, the OC image describes the important feature of image. Therefore, the OC image processing can perform the tracking of natural feature points precisely even if the condition of the image such as brightness or contrast is no good.
Moreover, in order to compare the robustness OC image processing and other method, the SfM procedure using grayscale image processing was performed. In this SfM procedure using grayscale image processing, natural feature points were

extracted by HARRIS operator and the tracking of natural feature points was performed by SSDA template matching. Figure 14 shows tracking result of natural feature points using HARRIS and SSDA. From Figure 13, it is under stood that tracking of natural feature points in the low brightness and contrast area is difficult using grayscale image processing.

Figure 8. Test Field

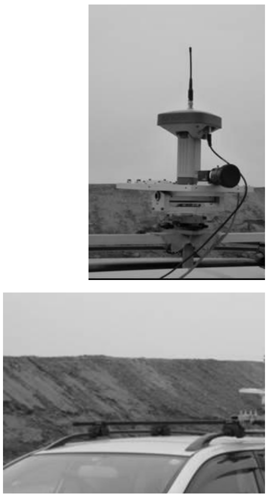


Figure 9. Test system

Name	PointGray Research Grasshopper
Resolution	200M pixel (1600×1200 pixel)
fps.	30fps
Lens	f=5mm

Table 1. Specification of Video Camera

Name	TOPCON GR-3
Signal	GPS/GLONASS L1/L2/L5 C/A P Code & Carrier
Accuracy	10mm + 1 ppm horizontal 15mm + 1 ppm vertical (RTK/ Kinematic)

Table 2. Specification of RTk-GPS

Figure 10. Example of video image sequence

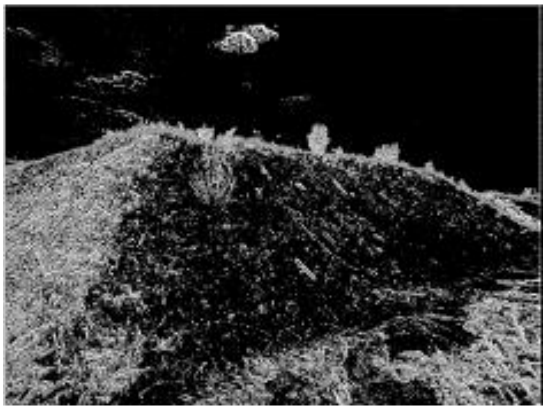


Figure 11. Example of OC image

Figure 12. Tracking Result using OCM and OCR

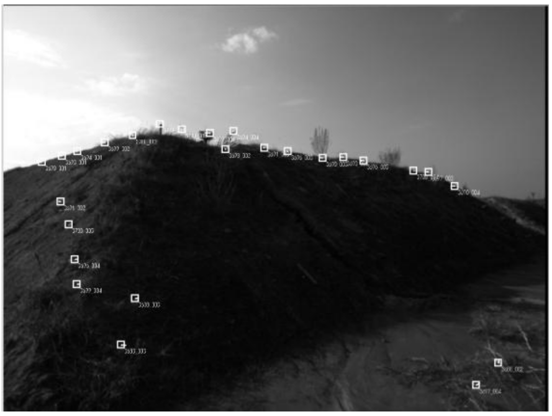


Figure 13. Tracking Reslut using HARRIS operator and SSDA Template Matching

Figure14 shows the result of SfM procedure using OCM and OCR proposed in this investigation. The trajectory of vehicle and 3D coordinate of natural feature points were obtained correctly due to the OCR and OCM can perform tracking of natural feature points regardless of environment.
Figure 15 shows the result of SfM procedure using HARRIS operator and SSDA template matching. In this SfM procedure though we used RTK-GPS as mentioned above, SfM procedure stopped on the way because not enough natural feature points are tracked as seen in Figure 13.
From this experiment, it is confirmed that the SfM procedure using OCM and OCR can perform robust exterior orientation at the outside environment.

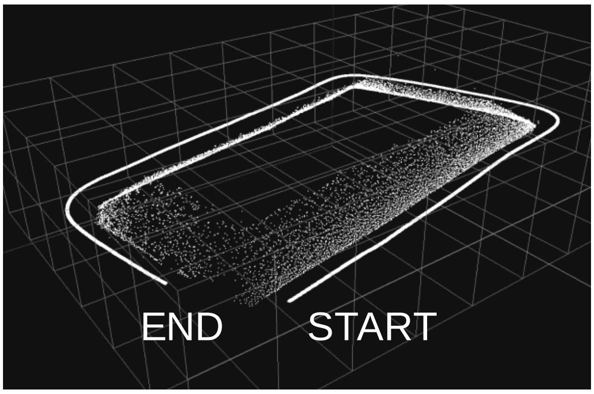


Figure14. Result of SfM procedure using OCM and OCR

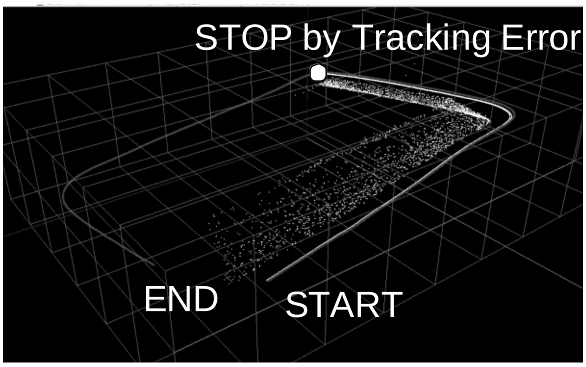


Figure15. Result of SfM procedure using HARRIS operator and SSDA Template Matching

4. CONCLUSION

In this investigation, in order to perform robust exterior orientation under the outdoor environment, the authors have developed the robust natural feature point tracking method using OC image processing for SfM procedure. The proposed robust natural feature point tracking method consists from natural feature point extraction procedure using OCR, template matching procedure using OCM. Moreover, exterior orientation was performed using Extended Bundle Adjustment that uses both tracking result of natural feature points and RTK-GPS data simultaneously. In this Extended Bundle Adjustment, evaluations of outlier in correspondence of natural feature points and RTK-GPS positioning data have been performed by using M-estimator. From these procedures, we can obtain the correct correspondence of natural feature points in each video frames, also the exterior orientation parameter in global coordinate without accumulative error for each video frames can be obtained. In experiments, the application for outdoor test field was performed and the effectiveness of the proposed method was confirmed.
However, the estimation of the reliability of proposed method is still not enough. Therefore, experiments of another situation such as city road are needed for the practical use of this method. Moreover, the processing speed of this proposed SfM procedure is not enough for real-time application. Thus the hardware development for real-time application is important next problem

REFERENCES

Anai, T., 2009. Exterior Orientation Method for Video Image Sequences with Considering RTK-GPS Accuracy. In: *9th Conference on Optical 3-D Measurement Techniques*, Vienna, Austria, July 1–3 pp 231-239.

Ullah, F., Kaneko, S., and Igarashi, S.,2001 Orientation code matching for robust object search. *IEICE Trans. of Inf. & Sys, E84-D(8)*, pp.999-1006.

Takauji, H., Kaneko, S. and Tanaka, 2005 Robust tagging in strange circumstance, *The Transactions of the Institute of Electrical Engineers of Japan (C)*, vol.125, pp.926 (in japanese)

K. Kitamura, 2009 Human Body Measurement by Robust Stereo-Matching , *In: 9th Conference on Optical 3-D Measurement Techniques*, pp 254-263, Vienna, Austria, July 1–3