Developing A Moveable Augmented Reality System Using Projector for Supporting Guide in Museum

Kazuto Omori †
Graduate School of Integrated Science and Technology, Department of Informatics
Shizuoka University, Shizuoka, Japan
Tel: (+81) 80-6922-9896, Email: molmol178@sugilab.net

Takahiro Sugiyama
Graduate School of Integrated Science and Technology, Department of Informatics
Shizuoka University, Shizuoka, Japan
Tel: (+81) 53-478-1551, Email: sugi@inf.shizuoka.ac.jp

Abstract. In museum curator's guide, visitors can focus on an exhibit, listen to its guide and understand its detail, if diversified explanatory information can be placed closing to the exhibit at the same time. Therefore, in this study we develop a moveable AR system that can superimpose explanatory information around the exhibit for supporting the curators to apply the AR technology. In this system these requirements are mentioned as below. One is avoiding the exhibit from light directly for protecting from being tanned. Another is reaching a same mapping conclusion with a higher speed and accuracy not only from the confront location but also from the oblique direction. The proposed system consists three processing blocks to identify the exhibit, to estimate its region on the image and to project some explanatory information related to its exhibit. Usually visitors can see only one side of the exhibit due to the limitation of the layout in museum, thus, the relations between the confront exhibit image and an input image taken by a camera, have same topological attributes in image elements even its camera moved. We apply this property to propose a feature points detection method and a matching method under perspective distortion robustly.

Keywords: museum guide, augmented reality, topological attributes, feature points

1. INTRODUCTION

In many museums, as an effective way to understanding the exhibits, many kinds of PDA and KIOSK device are prepared for assisting in the acquisition of an exhibit's information to visitors. The visitors who use these devices cannot focus on the exhibit and the device screen in the place away from the exhibit at the same time [1][2][3]. Due to the visitors has to concentrate on the operation which is based on themselves, their attention is deflected from the exhibit [4][5]. If the visitors can focus on the information near the exhibit, it's helpful to understand the exhibit and have a better observation.

Additionally, in a more effective way to understand the exhibit, the visitors can listen to the guide by the curator in the museum. There are three merits as follows: firstly, it is possible that the curators can adapt the visitors with different attributes for the explanation based on the knowledge and experience of themselves. Secondly, the visitors can also listen to the explanation while they are looking at the exhibition. Thirdly, the visitors can obtain more details from hearing the guide of the curators. However, curator's guide has several matters which makes it difficult to verbalize. If the detailed explanation by the curator's guide is added to the visual information, the visitors can have a better observation.

Therefore, the visitors can focus on an exhibit without look away, listen to its guide and understand its detail, if multimodal explanatory information that is being prepared, such as text, photos, and videos can be projected on the periphery of an exhibit. What's more, when the visitors listen to curator's guide, it's easier to understand while the visual information is being added into the curator's guide.

Our study is on the presupposition that the visitors interact with the curator directly in the museum. The purpose of this study is to support the curator's guide by developing a moveable AR system that can superimpose explanatory information around the exhibit to apply the AR technology.

In order to realize the movable projection AR system which is used to present the information to the exhibit in the museum, these requirements are mentioned as follows:

- Avoiding the exhibit from light directly for protecting from being tanned.
- Reaching the same mapping conclusion with a higher speed and accuracy not only from the confront location but also from the oblique direction.
- Whether the location of the system is changed or not, the mapping location has to be recognized in a high speed and high accuracy.

The proposed system constitutes recognition of the exhibit, estimation of its region and projection of explanatory information related to the exhibit. Especially in this paper, referring the recognition and the region estimation of the exhibit, we suggest a feature points detection method and a matching method. Both of them are under perspective distortion robustly. Because between the confront exhibit image and an input image taken by a camera, there are the same topological attribute in image elements even the camera is moved. Furthermore, we compare the existing methods and this method for feature point detection accuracy and reveal the superiority this method.

2. RELATED WORK

As the research which use an AR system in museum, Nagamatsu et al. [6] developed a projection AR system with invisible markers. This system is difficult to install in existing museum, because this method needs to install many invisible markers in exhibition hall of the museum. Meanwhile, Jinbo et al. [7] used the tablet device to display AR contents on its screen. In this method, there is a burden for the visitors because they have to move their sight line from the exhibit to the tablet screen. In addition, this method is not appropriate for watching the exhibit in big crowds. Moreover, Kondo et al. [8] used a Head Mounted Display(HMD) to superpose 3DCG contents on the real exhibit. Although, the system using HMD can present the suitable information for each user, the interactions between the visitors and the curator or among the visitor are hindered. Comparing with general appreciation, the appreciation with a HMD also has the problems which are the narrow angle of vision and the physical burden to users.

Those existing methods of feature point detection are mentioned as follows: firstly, SIFT [9] detect scale and position of keypoint based on Difference-of-Gaussian processing. Secondly, SURF [10] use box filter which utilize integral image. Thirdly, FAST [11] detect corner depending on the decision tree. Since these methods rely on the geometric information of boundary region, they corresponded to scale and rotation transformation. However, they cannot be robust perspective distortion in the same time. As the research corresponding to the perspective distortion, there are also some methods for extraction a feature according to the Machine Learning [12]. This method need learning in advance which cost large amount of calculation due to the replacement of the object whose feature point has defected every time. There is also a method using the invariant features in perspective distortion which is calculated from the placement of the center the character's gravity [13]. This method can't deal with any image and is only for the character. The third method use a camera position and orientation information of the previous frame based on SURF [14]. This method has a large amount of calculation in the practical level because of solving PnP problem with the RANSAC algorithm.

In the field of the image matching, MSER [15] collect the areas from the image which has the similar illumination into one region, and detect it. In addition, ASIFT [16] is improved robustness to affine transformation by using SIFT. However, these methods have a large amount of calculation.

In this study, in order to detect feature point in any images and robust in perspective distortion, we focus on the relationship between different regions in one image. There is Hueckel operator [17] which use the information of the region. Due to the shape of the pattern which is used for the operator transforms according to the perspective distortion, the feature point can't be detected robustly.

Therefore, in this study we detect the feature point according to the calculation and discrimination of the topological attribute of the regions which constitute the image and their boundaries. The topological attribute has a characteristic which is they have the same topological attribute, regarding the same feature point in many conditions except the occlusion and some other specific ones, even if scale, rotation and perspective transformed in the image.

3. SYSTEM OVERVIEW

We develop a system in order to help curator's guide in museum. Specifically, we develop the system which overlap the explanatory information on the exhibit directly by using projector and camera. In addition, since the conversation between the visitor and the curator is becoming the largest part of the guide, we propose an interface that the curator can be operated easily and never disturb the guide at the same time.

Figure 1 provides example usage of this system. This system consists of a moveable tripod, a computer, a projector, and a camera. The curator operates this system by using the tablet in his hand. The angle range of the camera we prepared for the system is wider than the projection range of the projector. Furthermore, this system processes the projector in order to protect the exhibit from being tanned by the light directly.

After setting the system so the camera can recognize the exhibit in the angle of view, the curator connects the system with the tablet by operating it. If user select the connection destination, the system will recognize the exhibit and the tablet screen will transit as shown in Figure 2. The user can also switch from display to non-display explanatory information by tapping the checkbox on the upper left corner of it.

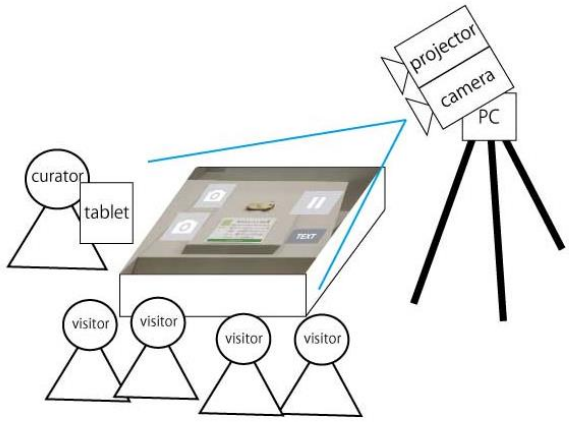

Fig.1. Example usage of this system

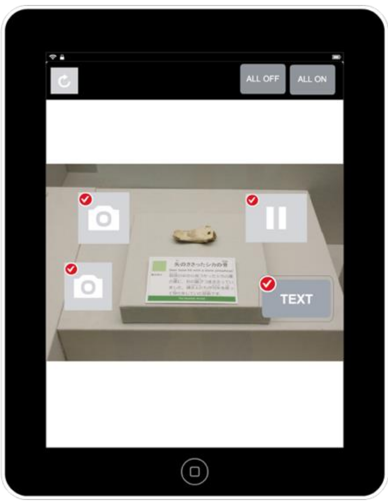

Fig.2. Guide view of this system

4. OUR METHODS

Through the feature point detection method proposed in this study, because of the different amount of the topical range (hereinafter called “patch”) in the region of the image, we consider some patterns which are composed by the regions. Figure 3 shows a sample of the image region in this study. In addition, we don’t consider the region which appears just like an independent island in the patch. The topological singularity won’t appear when the amount of the region in the patch is only one. In this study, we only deal with the patterns which consist two or three regions which are shown typically in Figure 4.

The pattern which consists two regions is as follows. Figure 4-2a shows the pattern that the different adjacent regions make the liner boundary by themselves, it is also called as ‘edge’. Figure 4-2b shows the pattern that the corner of the region is formed, it is also called as ‘corner point’. Figure 4-2c shows the pattern that the region has a part of the boundary of the curve. Figure 4-2d shows the pattern that the endpoint of a straight line with a width penetrates the patch. Figure 4-2e shows the pattern that a straight line with a width passes through the patch. Figure 4-2f shows the pattern that the curve with a width passes through the patch. Furthermore, Figure 4-2a~d are geometrically different, but they have a same topological attributes.

The pattern which consists three regions is as follows: Figure 4-3a shows the pattern that three regions pass through the patch. Figure 4-3b shows the pattern which consists one branch point. Figure 4-3c shows the pattern which makes T-junction. Figure 4-3d shows the pattern which consists two corner points. Figure 4-3e shows the pattern that two straight lines with a width pass through the patch. Figure 4-3f shows the pattern that two endpoints of a straight line with a width penetrate the patch. Furthermore, Figure 4-3b and Figure 4-3c are geometrically different, but they have same topological attributes.

In addition, we can also consider the pattern that form a more complex curve in the patch. However, it is conceivable the all patterns would match one of patterns in Figure 4 by making the patch as small as possible.

In this study, as the topological singular feature point, Figure 4-2b, Figure 4-2d, Figure 4-3b, Figure 4-3c are targeted in order to detect the feature point.Regarding the matching method, firstly we detect the feature point in contrast with the image of the template and the input image. Secondly we descript the feature of the detected feature point. Thirdly we do the matching by using the topological attributes.

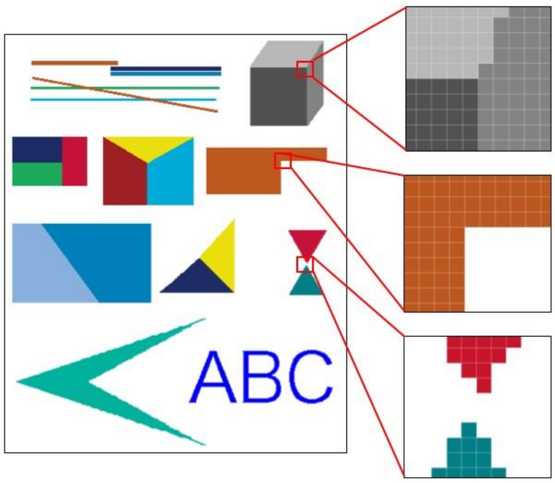

Fig.3. Sample of the image region in this study

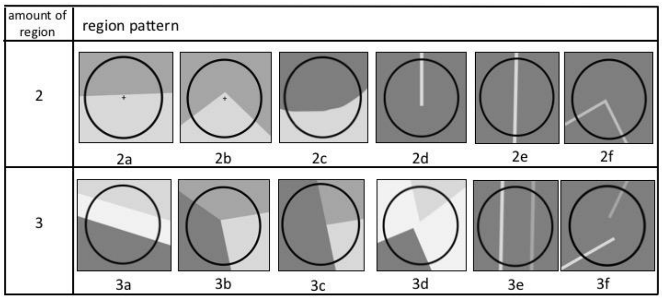

Fig.4. Region pattern of the local range

4.1 PROPOSED METHODS

In this study, we propose some methods which consist of image segmentation, topological feature calculation, feature point detection, feature point description and feature matching.

4.1.1 REGION SEGMENTATION USING VARIANCES OF GRAY IMAGE

By using the variances of the gray value of the histogram in two ranges with the histogram segmentation method, firstly we calculate a separability. Secondly, we segment the input image into regions. Thirdly, we make a label image.

At first, in Figure 5, the system makes the histogram of V value in HSV color space of the input image. We set width (T_1, T_2) to the left and right from x_n in the histogram, calculate the separability(η) between the two ranges. The ω_1 and ω_2 represent the proportions, the μ_1 and μ_2 represent the averages and the ${\sigma_1}^2$ and ${\sigma_2}^2$ represent the variances, of each region about the two regions.

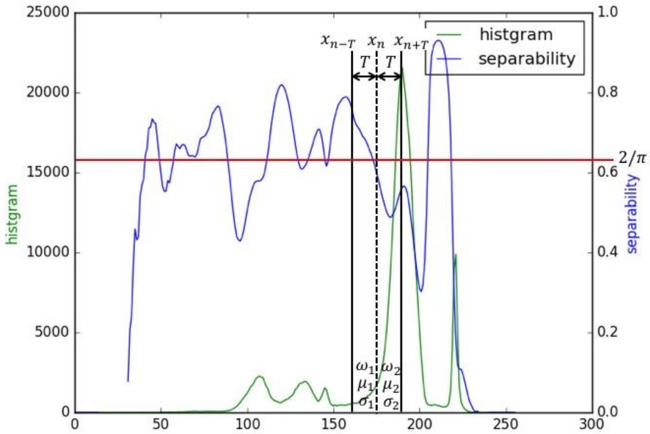

Fig.5. Histogram segmentation method

The total variances of the entire two regions are represented as Equation (1).

$$\sigma^2 = \omega_1\omega_2(\mu_1 - \mu_2)^2 + (\omega_1\sigma_1^2 + \omega_2\sigma_2^2) \qquad (1)$$

With the total variances σ^2, the separability η is represented as Equation (2)

$$\eta = \frac{\omega_1\omega_2(\mu_1 - \mu_2)^2}{\sigma^2} \qquad (2)$$

In the second place, the system calculates the separation point that in which point determines the histogram should be separated. From the preceding study [18], when the histogram is separated into two parts by threshold value and the distribution of the value in uniform region is Gaussian distribution at the same time, the maximum value of separability has been found to be $2/\pi$ regardless of the Gaussian distribution's variances. In this study, in order to separate the regions more clearly, when the separability η is the maximum value and over $2/\pi$, the point is regarded as the separation point.

At last, based on the separation point calculated, the system makes the label array and use it as a look-up table in order to output the label image which is translated from the input image.

4.1.2 FEATURE DETECTION METHOD USING TOPOROLOGICAL ATTRIBUTES

Firstly, the system scans the entire input image by using the 16 points circular patch (see Figure 6). The space of the scan is half a width of the patch. In addition, because the space has a big effect on the accuracy of the feature point detection, we will discuss this problem later.

Secondly, the system judges the regions in each patch if they have belonged to one of the patterns shown in Figure 4. Because of that, the system counts the amount of the label value which is composed of the pixel on the boundary of the patch. When the amount of the label value is one, it can't be judged as the feature point since there is only one region. When the amount of the label value is two and there are two the contiguous regions, also the angle of the region of which the label value of the patch's center is 120 degrees or less, it is judged as the feature point of the pattern in Figure 4-2b or Figure 4-2d. When the amount of the label value is three and there are three the contiguous regions at the same time, it is judged as the feature point of the pattern in Figure 4-3b or Figure 4-3c.

In specific process, the system converts the pixel on the boundary of the patch to a one-dimensional array and counts the partition of the label value shown in Figure 7. What's more, this array is processed as a circular array.

When the partition of the label value is two, it is judged as the feature point only if the amount of the pixels in which the same region as the label value of the patch's center, are the 1/3 or less of the amounts in the entire circumference. When the partition of the label value is three, it is judged as the feature point.

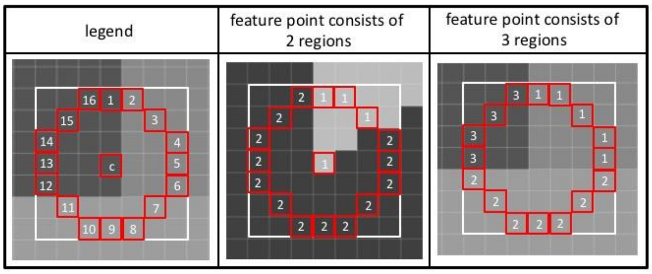

Fig.6. Feature points in this study

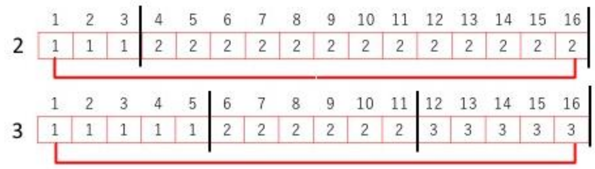

Fig.7. Feature point which became a one-dimensional array

4.1.3 MATCHING METHOD USING TOPOLOGICAL ATTRIBUTES

Through adapting the method of the feature point detection as mentioned to the image of the template and the input image, we descript the feature value and do the matching.

In this study, regarding the detected feature point, we interpret the feature point which has the same label value as a set. First of all, we do the matching of the set. Then, we do the matching of the feature point in respective sets. The speed and the accuracy of the matching can be improved by doing the two-step matching. Moreover, the improvement of the accuracy is estimated by using the geometrical relationship which is constituted by the feature point in the set of the label value.

In this study, we use the feature value as follows: regarding the feature quantities of the feature point, it is represented by the binary code which is created if the label value is included or not in the patch of the detected feature point and the average of the gray value in each region of the patch. Regarding the feature quantities of the set, it is represented by the label value of the set and the binary code which is created if it exists or not except for the label value of the set. These are shown in Figure 8.

We do the two-step matching by using these feature quantities.

The idea which is converted from the local region to binary also use in Local Binary Pattern (LBP) [19]. Although LBP is robust to illumination change, is not robust to other transformations. There are also some methods such as, rotated LBP [20] which has the rotation invariance, CS-LBP [21] in which calculation costs are reduced, Local Difference Binary(LDB) [22] which can recognize at a high speed and match in the large data base. However, these methods are not robust perspective distortion. Although MOBIL [23] which comes from LDB has the view point invariance, too many points which are not feature point have been detected by using ORB [24] technique. These become noises and have a negative effect on the result of detection.

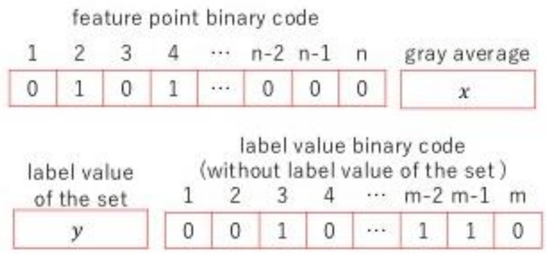

Fig.8. Feature value in this study

5. EXPERIMENT OF FEATURE DETECTION

In this experiment, in order to evaluate the detection performance of the feature point in our method, we use the images that the pattern of the unique feature point appears and the images that the perspective distortion appears from the former ones. The patch is the circumference which is constituted by the 16 pixels within the 7×7 pixels square. The space of the patch is 3 pixels. These are shown in Figure 9. In addition, we use FAST, SIFT, Harris corner detector [25] as comparison methods. We use the program, for comparison, which is published on the tutorial of OpenCV3.0 [26] [27] [28].

We show the results of the experiment as follows: the point which was detected as the feature point overlap two regions is the red rectangle. The point which was detected as the feature point overlap three regions is the green rectangle. The point which overlap two regions is the blue rectangle only if it isn't the feature point.

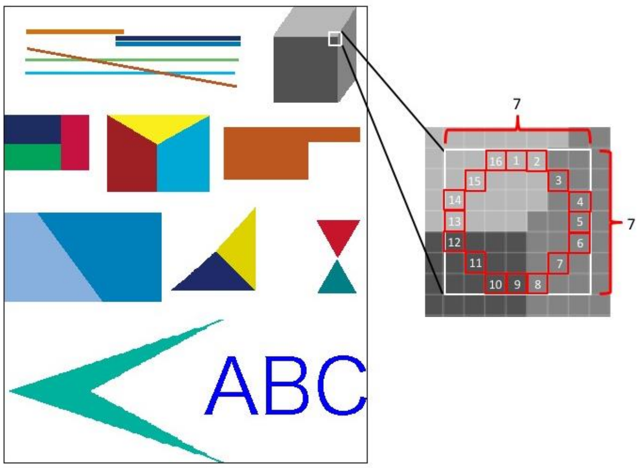

Fig.9. Image for the experiment and the patch

6. RESULT FOR THE PERSPECTIVE DISTORTION IMAGE

Regarding the template image and perspective distortion image, we show the result of the detected feature point as Figure 10 and Figure11. Furthermore, the result of the detected feature point regarding the perspective distortion image by using the comparison methods such as FAST, SIFT, Harris corner detector, are shown as Figure 12, Figure 13, and Figure 14. From this result, we prove that the feature point can be detected basically even for the perspective distortion. There are also several feature points which can't be detected in the corner pattern which is constituted by two regions. Those feature points can be found in an area that the center pixel doesn't overlap the small region in the region which overlaps the circumference of the patch shown as Arrow 1 in Figure 10 and Figure 11. The reason why feature points can't be detected is that there is an influence coming from space of the patches. It is confirmed that by setting the space at 1 pixel, those feature points can be detected as well. In addition, in the image of the result of the comparison methods, some points which are not singular, however have been detected as the feature points. As noises, these points have a negative effect on the process after feature detection such as feature matching. On the other hand, in the image of the result of our proposed method, it is confirmed that only the singular feature points have been detected exactly.

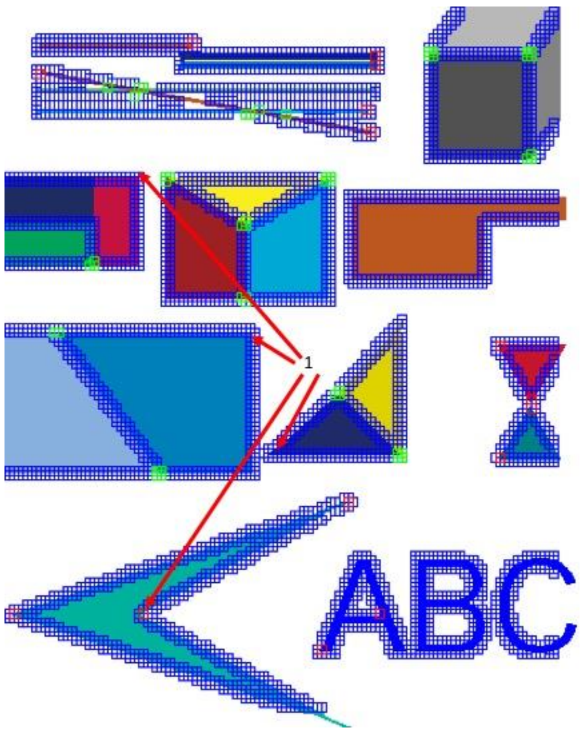

Fig.10. Result of the template image

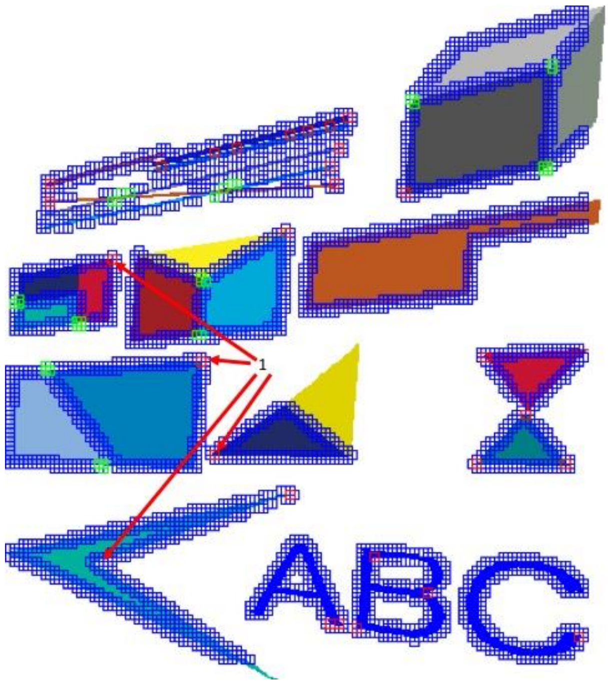

Fig.11. Result of the perspective distortion image

7. CONCLUSIONS

In this paper, we proposed three methods which are the detection, the description and the matching of the feature point by using the topological attributes of the image regions. According to the result of the experiment that the feature points can be detected robustly even towards an image with perspective distortion in the ideal image, the effectiveness of our method is revealed. In the future, we will improve our method in contrast with image in a real space and compare our method with the existing methods.

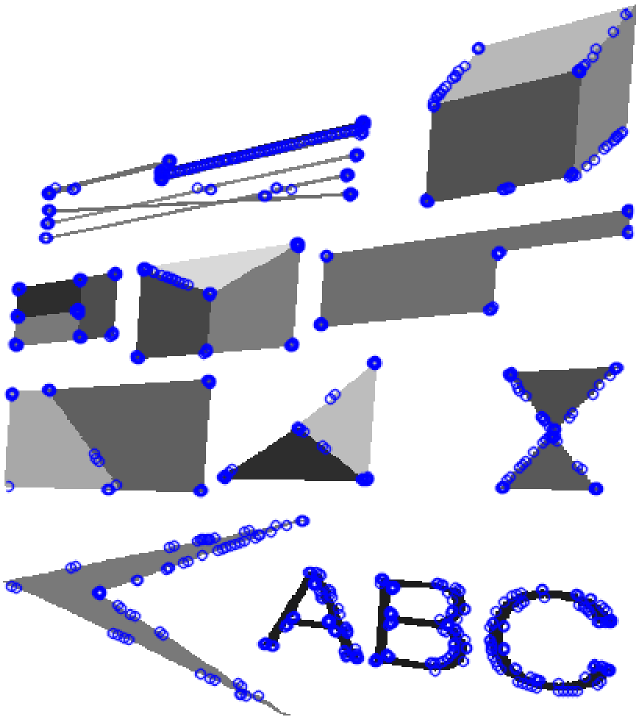

Fig.12. Result of the perspective distortion image using FAST

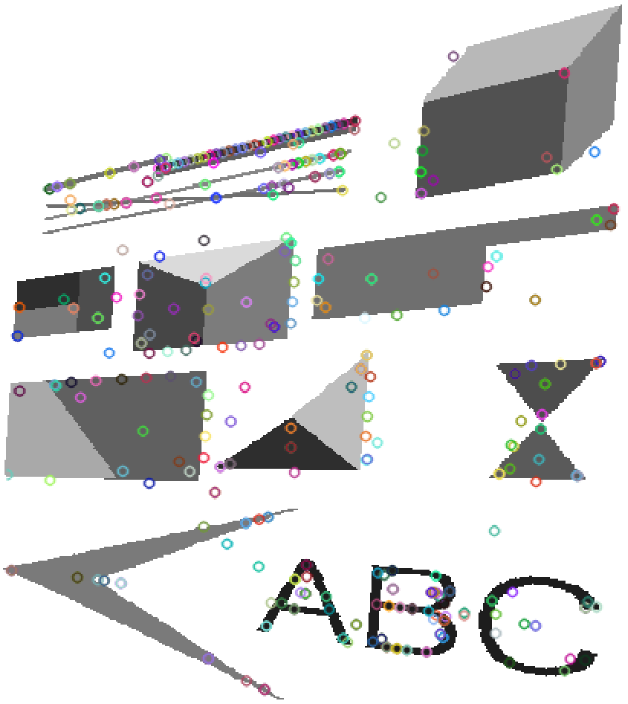

Fig.13. Result of the perspective distortion image using SIFT

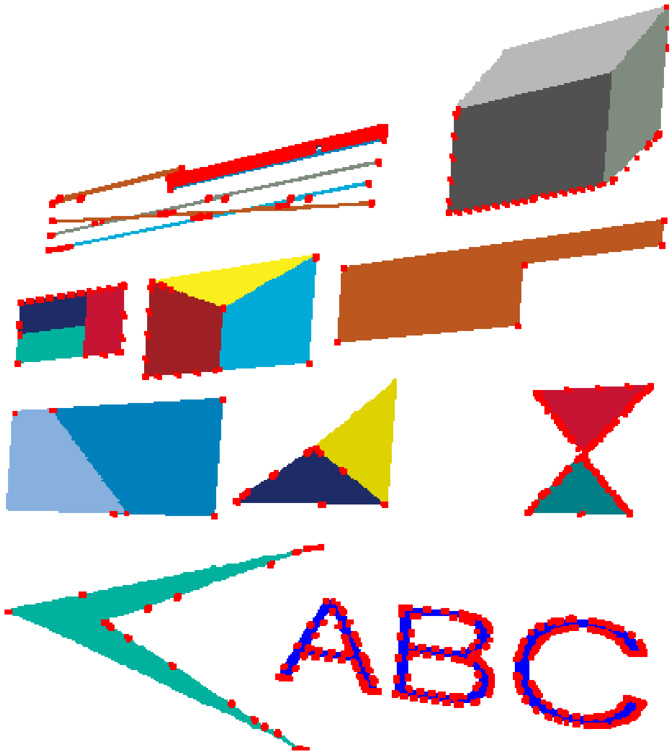

Fig.14. Result of the perspective distortion image using Harris corner detector

ACKNOWLEDGEMENT

This work was supported by Grant-in-Aid for Scientific Research(C)(15K01147)

REFERENCES

[1] The Japan Society for Exhibition Studies (2010), Exhibition Theory-Make the exhibit in the museum,Yusankaku(in Japanese)

[2] N. Watanabe(2015), Designing a tool for helping appreciation from the point of view of the visual information in museum, Graduation thesis, Department of Information Arts, Faculty of informatics, Shizuoka University(in Japanese)

[3] F. Kusunoki (2008), SoundSpot: Location-Bound Audio Guide System for Exhibition Supports in Museum, IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, Vol. J91-D No.2 pp.229-237(in Japanese)

[4] T. Yamaguchi(2004),Designing a tool to enhance children's in a museum: Pi_book, Japanese Society for the Science of Design,51 pp.48-49. (in Japanese)

[5] Y. Yagi(2011), Designing a tool to induce the observation that is enhanced the zoo's charm based on the analysis of the communication of parent and child. Master's Thesis, Department of Information Science, Graduate School of of Informatics, Shizuoka University. (in Japanese)

[6] A. Nagamatsu, Y. Nakazato, M. Kanbara and N. Yokoya (2009), AR Guide System Using Mobile Projector in Indoor Environment, Transaction of the Virtual Reality Society of Japan Vol.14, No.3, 283-293. (in Japanese)

[7] S. Jinbo, K. Anzai, Y. Saito and M. Nakamura (2014), Feasibility of AR Technology for Studying in Museum. Journal of Information Studies of Yokohama Campus in Tokyo City University (15), 16-22. (in Japanese)

[8] T. Kondo, J. Shibasaki, H. Arita, M. Manabe and R. Inaba (2006), A Proposal of a Mixed Reality Museum Display System, Journal of Japan Society for Educational Technology, Vol.30(Suppl.) 45-48. (in Japanese)

[9] D. G. Lowe(1999), Object Recognition from Local Scale-Invariant Features. Proc. International Conference on Computer Vision, 1150-1157.

[10] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool (2006), Surf: Speeded up robust features. In Proceedings of European Conference on Computer Vision, pp. 404–417.

[11] E. Rosten and T. Drummond (2006), Machine Learning for High-speed Corner Detection, Proc. European Conference on Computer Vision, 430-443.

[12] Vincent Lepetit Mustafa Ozuysal, Michael Calonder and Pascal Fua (2009), Fast keypoint recognition using random ferns. In Proceedings of the IEEE Trans- actions on Pattern Analysis and Machine Intelligence.

[13] H. Uchiyama, H. Saito (2009), Augmenting text document by on-line learning of local arrangement of keypoints. In Proceedings of the IEEE International Symposium on Mixed and Augmented Reality, pp. 95–98, Washington, DC, USA, IEEE Computer Society.

[14] K. Kouduki (2013), Robust feature tracking for Augmented Reality based on SURF under perspective distortion considering the camera pose information, Master's Thesis, Department of Information Processing, Graduate School of Information Science, Nara Institute of Science and Technology. (in Japanese)

[15] J. Matas, O. Chum, M. Urban and T. Pajdal (2002), Robust wide baseline stereo from maximally stable extremal regions, Proc. British Machine Vision Conf., pp. 384–393.

[16] J. -M. Morel and G. Yu (2009), ASIFT: A new framework for fully affine invariant image comparison, SIAM J. Imaging Sciences, 2, 2, pp. 438–469.

[17] H. M. Hueckel (1973), A Local Visual Operator Which Recognizes Edges and Lines, Journal of the ACM, Vol. 20, 1973m pp. 634-647.

[18] T. Sugiyama and K. Abe (2000), Edge Detection Method Based on Edge Reliability with Fixed Thresholds: Consideration of Uniformity and Gradation, Proc. of the 15th International Conference on Pattern Recognition, Vol.3. pp.660−663.

[19] Timo Ojala, Matti Pietikäinen, and David Harwood (1996), A comparative study of texture measures with classification based on featured distributions. Pattern Recognition, Vol. 29, No. 1, pp. 51–59.

[20] Matti Pietikäinen, Timo Ojala, and Zelin Xu (2000), Rotation invariant texture classification using feature distributions. Pattern Recognition, Vol. 33, No. 1, pp. 43–52.

[21] Marko Heikkilä, Matti Pietikäinen, Cordelia Schmid (2009), Description of Interest Regions with Local Binary Patterns, Pattern Recog, vol. 42, no. 3, pp. 425-436.

[22] Xin Yang, Kwang-Ting Cheng (2012), LDB: An Ultra-Fast Feature for Scalable Augmented Reality on Mobile Devices, IEEE International Symposium on Mixed and Augmented Reality (ISMAR) pp. 49-57.

[23] Abdelkader Bellarbi, Samir Otmane, Nadia Zenati, Samir Benbelkacem(2014), MOBIL:A Moments based Local Binary Descriptor, IEEE International Symposium on Mixed and Augmented Reality (ISMAR) pp.251–252.

[24] Ethan Rublee, Vincent Rabaud, Kurt Konolige, Gary Bradski, (2011) ORB: an Efficient Alternative to SIFT or SURF. In Proc. of ICCV'11, Barcelona, Spain.

[25] C. Harris and M. Stephens (1988), A combined corner and edge detector, Proc. 4th Alvey Vision Conf., pp.147–151, Manchester, U.K., Aug.

[26] http://docs.opencv.org/3.0-beta/doc/py_tutorials/py_feature2d/py_fast/py_fast.html#fast

[27] http://docs.opencv.org/3.0-beta/doc/py_tutorials/py_feature2d/py_sift_intro/py_sift_intro.html#sift-intro

[28] http://docs.opencv.org/3.0-beta/doc/py_tutorials/py_feature2d/py_features_harris/py_features_harris.html#harris-corners