Proc. VIIth Digital Image Computing: Techniques and Applications, Sun C., Talbot H., Ourselin S. and Adriaansen T. (Eds.), 10-12 Dec. 2003, Sydney

Consistent Presentation of Interactive Virtual Objects in Real Space with 3D Markers – Interactive Virtual Interior Design –

Yoshihiro Yasumuro[1], Yu Ishikawa[2], Koichi Minami[1], Masataka Imura[1], Yoshitsugu Manabe[1], and Kunihiro Chihara[1]

[1] Graduate School of Information Science, Nara Institute of Science and Technology, 8916-5, Takayama, Ikoma, Nara, Japan
{yasumuro, koich-mi, imura, manabe, chihara}@is.aist-nara.ac.jp
http://chihara.aist-nara.ac.jp/
[2] Matsushita Soft-Research, ltd,
4-5-15, higashi-shinagawa, shinagawa-ku, Tokyo, 140-8632, Japan
yu-ishi@msr.mei.co.jp

Abstract. For presenting virtual objects in a real space, mixed reality (MR) technique is very effective to simulate an overview of a scenery. In interior design, for instance, the coordinates of the arrangements and the colours of the items need to be determined based on various layout plans. We propose a new framework for interactively arranging and rendering virtual objects in real space by manipulating physical 3D markers. A 3D marker consists of 2D codes and a spherical mirror for acquiring both geometric and photometric conditions by single camera instantaneously. Experimental results showed that the proposed approach is effective for providing appropriate geometry and local lighting conditions to present each virtual object in a simple system. The proposed system allows the user to experiment arranging of virtual furniture and verify the created scene which is visually consistent with the physical space.

1 Motivation

Coordinating furniture and interiors for living space gathers continuing interests in general. However, planning of colour coordinates and organising interior items often require expert knowledge. How to show the design plan is important to communicate between the expert and a customer. Designers usually paint perspective figures for conveying concepts of their plans as well as technical drawings. However, hand-painted figures are basically subjective expression and are not always intuitive to the customers. Moreover, it takes some time to paint the figures and instant modification over the discussion is impossible.

In the contrast, computer-generated scenery using software for CAD (computer aided design) and CG (computer graphics) is capable of not only presenting objective perspective figure based on technical drawings but also allowing users to change viewpoints and the contents to present. Since every single component in the scene must be digitised and modelled to be rendered, synthesising scenery also spends time and effort. Synthesised image has a limitation in terms of reality expression as well. Especially in the case of reforming or rearranging the interior items in an existing room or a space, it is difficult to present realistic scenery including physical objects and added CG components as if they coexist.

Instead of rendering all of the components using CG, this paper is addressed to how to present virtual objects in existing spaces, introducing MR (mixed reality) techniques. We propose a method for automatically maintaining geometry and photometric consistency between CG objects and physical spaces for realtime MR presentation.

2 Related Work

Mixed reality technique enables virtual objects to be superimposed on real environment scenery. To present virtual and real objects seamlessly, geometric and photometric conditions between virtual and physical world need to be consistent. Moreover, if realtime rendering of MR scene is achieved time consistency among the user's behaviour, physical conditions and presentation of the virtual objects is supported. This allows the users to manipulate virtual objects intuitively and modify the MR scene interactively.

One of the major methods to achieve geometric consistency is vision-based approach. Even a passive marker such as a 2D code printed on a piece of paper works for indicating objects' geometry effectively [1–5]. For acquiring dense geometry information, the use of natural features in a scene is employed. Since automatic and dynamic detection of natural feature points has limitation in reliability, hybrid use of markers and natural features are also proposed[6]. Another method is to employ positioning sensors such as the ones using magnetic field. Since the sensors must be prepared as many as the objects to be registered and thus the system tends to be expensive and complicated. Hybrid methods of optical markers and positioning sensors is also used for MR [7].

As for acquisition of photometric conditions for MR presentation, a wide angle image sensor such as a camera with a hyperboloidal mirror or a fish-eye lens is used to measure light source distribution effectively [8, 9]. However, local lighting condition is not always reflected correctly and required computational cost is not suitable for dynamic change of surroundings.

The objective focused on in this paper is simultaneous support of both geometric and photometric consistency for creating MR scenes. There is no decisive method to respond this requirement yet. Andrei [7] proposed hybrid method to use an optical marker and a positioning sensor to register a virtual object in real space. Andrei demonstrated environment mapping onto virtual object, but the acquired photometric information is not sufficient for generic rendering CG models. We proposed a method to support geometric and photometric consistency, using a physical 3D marker with a spherical mirror and multiple 2D codes [10]. Kambara [11] also used a combination of a small metal ball and a simple 2D code for measuring rough estimation of lighting condition and the marker posture. In this paper, we extend the 3D marker approach to construct flexible

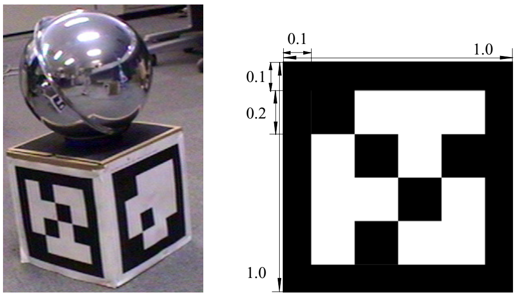

Fig. 1. Examples of 3D Marker (left) and 2D Code (right); Digits Stand for the Size Ratio.

MR space in which the user can manipulate multiple virtual objects in realtime.

3 Proposed Method

Since MR scenery is visually appreciated from a user's viewpoint eventually, we settle the base coordinate system on a camera, considering a use of video-see-through manner. Each virtual object has its coordinate system described relative to the camera coordinate system. 3D markers are positioned as the user wants to arrange the virtual objects. The camera capture the position of the marker and how the marker is illuminated in the physical space. Geometric and Photometric information of the 3D marker make the virtual object to be presented as the marker is placed physically. We compose a 3D marker with a spherical mirror and multiple 2D codes. 2D codes allow the camera to acquire the relative geometry from variety of view points. In addition, 2D patterns of the code can be assigned uniquely for identifying multiple markers in the same scene. Spherical mirror shows how the 3D marker is illuminated by the surrounding to the camera at any viewpoints. These geometric and photometric conditions can be captured in a perspective of single camera simultaneously and thus the system constitution is expected to be simple. Once preparing 3D makers with specified standard such as code patterns and the size, the user can use variety of virtual objects as many as the number of code patterns permit. The only device that requires calibration is a camera. Considering the use for existing spaces, the proposed approach is suitable because of easiness for handling and portability.

3.1 3D Marker

We put a 2D code on each side of a cube and a spherical mirror on the top (Fig.1). We assign a unique pattern to each 2D code to use it to detect which side is seen from the camera. In this paper, used $4 \times 4 = 16$ bits pattern. Considering that a 3D marker needs four 2D markers and excluding rotationally symmetric

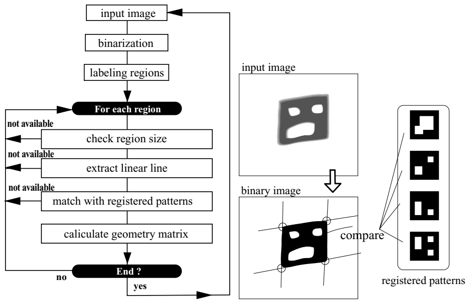

Fig. 2. Process for Detecting and Decoding 2D Code

patterns, $(2^{16} - 2^8)/(4 \times 4) = 4080$ kinds of virtual objects can be assigned for 16 bits 2D code.

3.2 Acquisition of Geometry Information from 2D Code

After detecting a 2D code, its position and posture are calculated based on perspective projection model and expressed in camera coordinate system[2, 12]. Firstly, the system binarizes input image and labels the foreground regions. Checking the size of each region and its contour, the regions which has proper size and a contour consists of 4 linear lines are selected. Template matching is performed for detecting 2D codes by comparing selected regions and registered patterns (Fig.2). In our implementation, we used ARToolKit library[2] for 2D code detection and geometry transformation calculation.

3.3 Acquisition of Lighting Condition from Spherical Mirror

We assume that lighting condition at the position of the 3D marker contains direct light from the light source, reflected light from specular surface in the environment and ambient light from distant view [8, 9]. As shown in Fig.3, when a spherical mirror is observed as a circle whose radius is r pixels and center position is (x_c, y_c), normal vector $\overrightarrow{\mathrm{N}}$ on the sphere surface at the position (x, y) observed in the spherical mirror image is expressed as (Eq. 1).

$$\begin{aligned} \overrightarrow{N} &= [e_x, e_y, e_z] \\ e_x = (x - x_c)/r, e_y &= (y - y_c)/r \\ e_z &= \sqrt{1 - (e_x^2 + e_y^2)} \end{aligned} \tag{1}$$

Light coming through the focal point O into a pixel position (x, y) is emitted from a direction which is regular reflection line relative to normal vector $\overrightarrow{\mathrm{N}}$ on

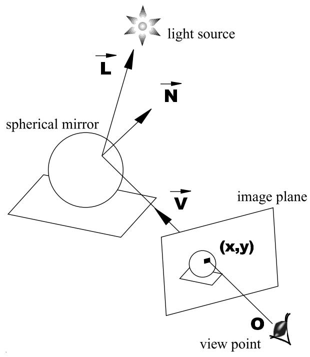

Fig. 3. Schematic Geometry of Viewpoint and Light Source

the sphere. In the same way, the direction of the light source to illuminate each pixel on the image plane can be calculated.

$$\overrightarrow{\mathrm{L}} = \overrightarrow{\mathrm{V}} - 2(\overrightarrow{\mathrm{N}} \cdot \overrightarrow{\mathrm{V}})\overrightarrow{\mathrm{N}} \tag{2}$$

We divide the image of the spherical mirror into small regions and sample the intensity of the light source which illuminate each region. Averaged values for R, G and B are used for the light intensity. For calculation of the direction of the light sources, we used centers of gravity of the regions.

3.4 Rendering

We assume that the colours of the object surface are determined according to the model of Lambert and Phong (Eq. 3) for rendering virtual objects.

$$\begin{aligned} I &= Specular + Diffuse \\ &= M_s \frac{\sum_r \sum_\theta I_s \left[\frac{<n \cdot h_{r,\theta}>}{|n||h_{r,\theta}|} \right]^\mu}{\sum_r \sum_\theta \left[\frac{<n \cdot h_{r,\theta}>}{|n||h_{r,\theta}|} \right]^\mu} + M_d \frac{\sum_r \sum_\theta I_d \frac{<l_{r,\theta} \cdot n>}{|l_{r,\theta}||n|}}{\sum_r \sum_\theta \frac{<l_{r,\theta} \cdot n>}{|l_{r,\theta}||n|}} \end{aligned} \tag{3}$$

Here, n stands for normal at a vertex, $h_{r,\theta}$ is the intermediate vector between incoming radiation vector and view point vector. $< \cdot >$ stands for inner product. An angle set of r and θ specifies an incoming radiation vector $l_{r,\theta}$. I_s and I_d are intensities of specular and diffuse component of the light source, respectively. M_s and M_d are coefficients for specular and diffuse reflection on a vertex of model surface to be rendered. μ is a parameter for beam width of specular reflection. Applying (Eq. 3) to the colour component of R, G and B, each vertex of the virtual object is rendered.

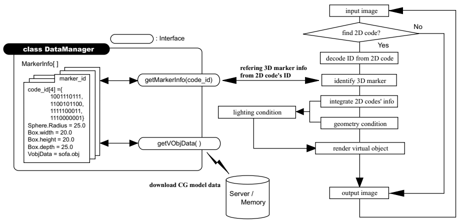

Fig. 4. Process Flow

Table 1. Devices and Materials for Experimental Setup

PC	CPU:Intel Pentium4 1.8GHz Memory:512MB Graphic Board:nVIDIA GeForce3
camera	WAT-202D by WATEC (F:1.4, f:6mm)
3D marker	paper cube 20cm on a side monochrome printed (2D code)

4 Implementation

Fig.4 shows the process flow. If a 2D code is detected on an image frame captured by the camera, geometrical information including its distance and posture is calculated from the appearance of the contour of the code region [2]. Embedded ID number is acquired by decoding the 2D code. Each ID number is associated with a CG model stored in a virtual furniture library. CG furniture are polygon models which contain 2000-8000 vertices on which diffuse and specular reflection parameters are properly designed. Table 1 shows the devices and materials used for our implementation.

4.1 Acquisition of Light Source Distribution

Based on the captured geometry of 2D code(s), a spherical mirror image for each 3D marker is automatically extracted as a circle region. We divide the region respect to radius and angle after normalising as 50 pixel radius size, The

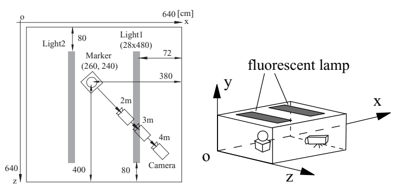

Fig. 5. Layout of the Room (left) and Coordinate System (right)

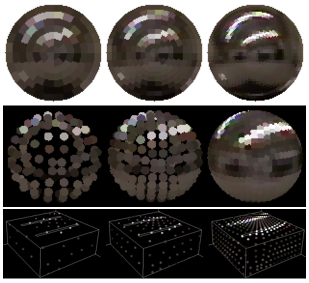

Fig. 6. Different Division for Lighting Map — Spherical Mirror Image (top), Lighting Map (middle), 3D Arrangement of the Lighting Map for Comparison (bottom), Sampling Number for Light sources:112(left), 220(center), 480(right)

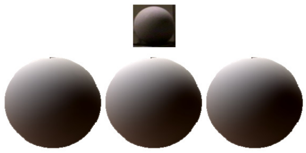

Fig. 7. Rendering Results of Diffused Surface Sphere with Different Lighting Sampling — Real Plaster (upper), 80 Polygon CG Image (sampling number for light sources:112(left), 220(center), 480(right)

circle region is divided equally into K rings and the k-th ring (counting from the center) is divided into $4k$ arcs. Fig. 5 depicts a layout of a room and its sealing lights for an experiment. We tried three different sampling numbers, $112(K = 7)$, $220(K = 10)$ and $840(K = 20)$. Divided spherical mirror image is shown in Fig. 6(top). Fig. 6(middle) indicates sampled lighting map. Fig. 6(bottom) shows physical light arrangement for comparison. Rendering results of diffused surface sphere model (80 polygons) are shown in Fig. 7. A real plaster ball in the same condition is shown in Fig. 7(top) for comparison. The result shows that lighting map resolution does not affect crucially rendering diffused objects. However, expression of specular items may require higher resolution lighting map to reflect. Since computational cost for colour integration (Eq.3) depends on the number of lights and vertices, we prepare several resolutions of lighting map so that the user can switch them according to rendering quality of the glossy items

4.2 Rendering Results

Since proposed method is capable of identifying each 3D marker, multiple virtual objects are rendered individually taking into account associated 3D marker

Fig. 8. Effect of Local Lighting Condition — A partition put far from markers (left) and put closer (right)

information. Fig. 8 shows three virtual chairs using three 3D markers. As a real partition moves, shade and reflection change photometric condition locally. Each virtual object expresses local photometric effects individually. Using four 3D markers, Fig.9 demonstrates a scenery of a room in different hours of a day. A desk, a sofa, a chair and a computer(on a real desk) are virtually arranged. Daylight and fluorescent lamp have totally different characteristics of illumination colour, intensity and distributions. However, the same system setup achieves harmonised rendering results for each occasion. In our implementation, image capturing rate is 30 [frame/sec]. Rendering rate is 0.5-10 [frame/sec], depending on the number of objects and vertices of the CG models. Comparing to existing renderer for interior design, rendering time is so short that the user can try some different layout and check it interactively(Fig. 10). A set of vertical 2D codes helps to enhance visibility of the 3D marker comparing to that of existing MR systems [2, 11].

5 Discussion and Conclusion

3D marker we used was 20 cm cube, which allows the system to capture its position and direction with measurement error within a few percentage in the range up to 8m from the camera. Considering that the user is supposed to move around carrying a camera, possible working range of MR space is not limited as 8m. Observed problem is that visibility of the 2D marker itself is affected by the physical lighting condition. If the room is dark or shade is casted on the marker, geometric measurement error became larger or even missed to detect the marker sometimes. Introducing invisible light such as near-infrared light and reflection paint marker can be considered for improvement. Another problem is caused by limitation of the dynamic range of the camera, which is not able to cover both intense and dim light. Utilising several cameras with different lens diaphragm[8] as well as a camera with hi-dynamic range is expected to improve dynamic range of camera system.

We proposed an interactive MR system which supports geometric and photometric consistency for presenting multiple virtual objects. Our future work is

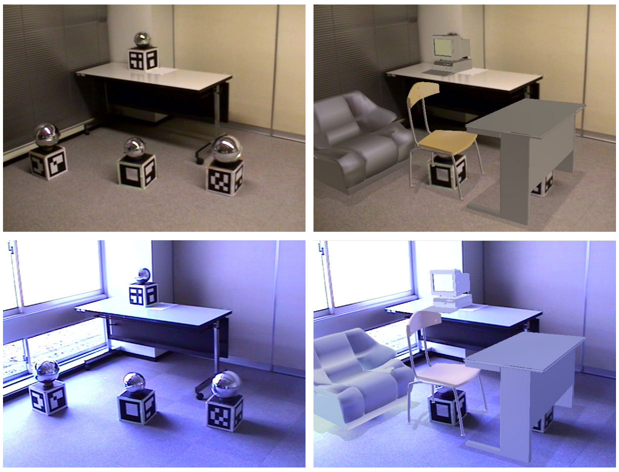

Fig. 9. Different Lighting Condition for a Whole Room — fluorescent lamp (top), day light from windows (bottom); input image (left) and output image (right)

addressed to constructing a collaborating space over the network for sharing the proposed MR space.

Acknowledgement

This research is (partly) supported by Core Research for Evolutional Science and Technology (CREST) Program "Advanced Media Technology for Everyday Living" of Japan Science and Technology Agency (JST).

References

1. B.Thomas : "An outdoor/indoor augmented reality first person application", 4th International Symp. Wearable Computers (ISWC 2000), pp. 139–146, (2000)
2. H.Kato and M.Billinghurst : "Marker tracking and hmd calibration for a video-based augmented reality conferencing system", IEEE International Workshop on Augmented Reality, pp. 125–133, (1999)
3. N.Navab, E.Cubillo, B.Bascle, J.Lockau, K.D.Kamsties, and M.Neuberger: "a software platform for the creation and update of virtual factories", 7th IEEE International Conference on Emerging Technologies and Factory Automation, pp. 459–463, (1999)
4. R.M.Haralick, C.Lee, K.Ottenberg, and M.Nolle : "Review and analysis of solutions of the three point perspective pose estimation problem", International Jurnal of Computer Vision, Vol. 13, No. 3, pp. 331–356, (1994)
5. Ward Mark, Ronald Azuma, Robert Bennett, Stefan Gottschalk, and Henry Fuchs : "A demonstrated optical tracker with scalable work area for head-mounted display systems", Proceedings of 1992 Symposium on Interactive 3D Graphics, pp. 43–52, (1992)

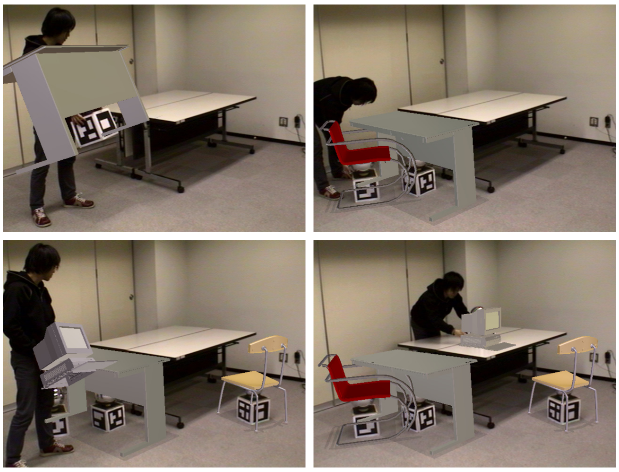

Fig. 10. Interactive Manipulation

6. Tomokazu Sato, Masayuki Kanbara, Naokazu Yokoya and Haruo Takemura "3-D modeling of an outdoor Scene by multi-baseline stereo using a long sequence of images", Proc. 16th IAPR Int. Conf. on Pattern Recognition (ICPR2002) , Vol. III, pp. 581-584, August. (2002)
7. Andrei State, Gentaro Hirota, David T. Chen, William F. Garrett, and Mark A. Livingston : "Superior augmented-reality registration by integrating landmark tracking and magnetic tracking", Proceedings of SIGGRAPH 96, pp. 429–438, (1996)
8. Paul Debevec : "Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography", Proceedings of SIGGRAPH 98, pp. 189–198. 1998.
9. I.Sato, Y.Sato, and I.Ikeuchi : "Acquiring a radience distribution to superimpose virtual objects onto a real scene", IEEE Transactions onVisualization and Computer Graphics, Vol. 5, No. 1, pp. 1–12, (1999)
10. Yu Ishikawa, Koichi Minami, Masataka Imura, Yoshihiro Yasumuro, Yoshitsugu Manabe, Kunihiro Chihara: "Interactive Rendering of Virtual Objects in Real Scenes Using 3D Marker", Proc. of the 46th Annual Conf. of ISCIE, pp.317–318, (2002) (in Japanese)
11. M. Kanbara and N. Yokoya: "Geometric and photometric registration for real-time augmented reality", Proc. IEEE and ACM Int. Symp. on Mixed Augmented Reality(ISMAR2002) , pp. 278-279, (2002).
12. Diego Lopez de Ipiña, Paulo Mendonca, and Andy Hopper : "a low-cost vision-based location system for ubiquitous computing", Personal and Ubiquitous Computing, Vol. 6, No. 3, pp. 206–219, (2002)