Application of Visual Information Extraction from Images

Toru Kaneko and Atsushi Yamashita

Faculty of Engineering, Shizuoka University, 3-5-1 Johoku Hamamatsu 432-8561, Japan, TEL:& FAX +81-53-478-1070,
e-mail: tmtkane@ipc.shizuoka.ac.jp

Abstract

Visual information extraction from images plays a very important role for many applications using computers in various fileds. This paper introduces two examples of applications developing in our laboratory; one is checker-pattern screen chroma key for image composition, and the other is a method for digital archiving of book pages using a stereo vision system.

1. Introduction

Image processing techniques provide fundamental tools to handle images; noise reduction, histogram transformation, binarization, edge detection, smoothing, image enhancement, texture analysis, region extraction and so on. Computer vision provides a theoretical background for scene analysis, especially for interpretation of the 3-D world.

Sometimes we also use a word machine vision, and the following is a list example of machine vision application fields given in MVA2007[1], 10th IAPR (The International Association for Pattern Recognition) Conference on Machine Vision Applications:

- Factory Automation: Inspection, Diagnosis, Assembly
- Intelligent Transport Systems: Traffic Monitoring, Traffic Con trol, Driving Safety Assistance, Electric Toll Collection
- Robots: Home Robots, Mobile Robots, Service Robots, Space Robots
- Security Systems: Surveillance, Alarm Systems, Biometrics and Personal Identification, Home Security and Monitoring Systems, Forensic Systems
- Medical Systems: Medical Examination Assistance, Remote Medicine, Image-guided Surgery and Intervention, Analysis of 2D/3D Biomedical Images
- Geographic Information Systems: Map Processing, 3D Reconstruction from Photographs or Maps
- Multimedia: Image Retrieval, Document and Drawing Analysis, Digital Archiving, Multimedia Databases and Analysis
- Action Analysis and Recognition: Gesture Recognition, Behavior Recognition
- Human Computer Interaction: Face Recognition, Wearable Computing, Multimodal Interface, Personal Imaging
- Vision and Graphics: Augmented and Mixed Reality
- Others: Applications to Civil and Construction Engineering, Agriculture, Forestry, Fishery, etc

In our laboratory, we concentrate our effort on developing methods of image processing and machine vision applications. Our interests are on how to extract visual information from deteriorated or distorted images. For example, vision systems are usually assumed to be used in atmospheric environments where we do not need to consider the phenomenon of refraction. On the other hand, we are interested in underwater image processing where refraction causes image distortion[2]. We have been also studying how to remove stationary noises in images such as those caused by rain drops adherent on the camera lens protector. By using complementary images taken with a stereo camera system, we can interpolate image portions behind rain drops[3].

In this paper, we introduce two machine vision applications developing in our laboratory, The first one is checker-pattern screen chroma key for image composition, and the other is a method for digital archiving of book pages using a stereo vision system.

2. Checker-Pattern Chroma Key

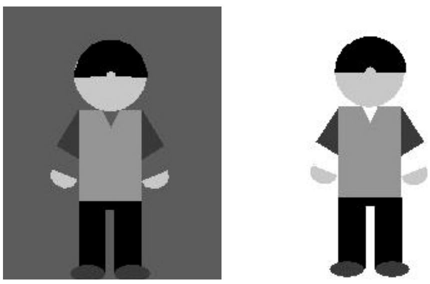
Image composition[4] is very important to creative designs such as cinema films, magazine covers, promotion videos, and so on. This technique can combine images of actors or actresses in a studio and those of scenery taken in other places.

To perform image composition, objects of interest must be extracted from images, and there are many studies about object extraction, e.g., pixel-based, area-based, edge-based, and physics-based ones.

Chroma key, which is also referred to as color keying or color-separation overlay, is a well-known image segmentation technique that removes a color from an image to reveal another image behind. Objects segmented from a uniform single color (usually blue or green) background are superimposed electronically to another background. This technique has been used for long years in the TV and the film industries.

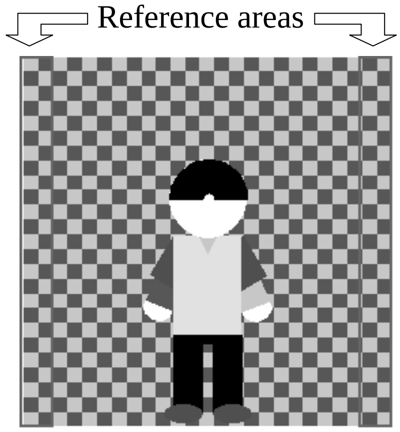
However, conventional chroma key techniques using a monochromatic background have a problem that foreground regions are regarded as the background if their colors are the same as the background, and these regions become transparent (Fig.1.).

To solve the above problem, we have developed a method for extracting foreground objects in any color by using a two-tone checker pattern background screen (Fig.2.)[5].



(a) (b)

Fig.1. Object extraction by chroma key.
(a) Foreground regions have the same color as the backgroun screen.
(b) Such regions become transparent.



(a) Checker pattern background.

(b) HLS color space.

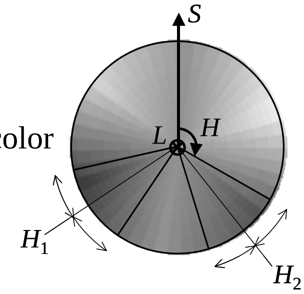


Fig.2. Chroma key with a checker pattern background screen.

<u>Background color extraction</u>

Candidate regions for the background are extracted by using a color space approach(Fig.2(b)). We select blue and green for the background two colors so that we can easily distinguish these two colors from each other and also from the human skin color. In the procedure, since the colors vary from the constant values according to the light conditions, we first examine the color components of the reference area, shown in Fig.2(a), where foreground objects do not exist.

<u>Background grid line extraction</u>

The digitized image of the checker pattern background has three types of regions; two types are for regions with pixels whose color values are either of the two tone colors, respectively, and the other type is for the boundaries of the checker patterns. The color of these boundaries has a different value from either of the two tones, because it is a mixture of the two colors.These boundary

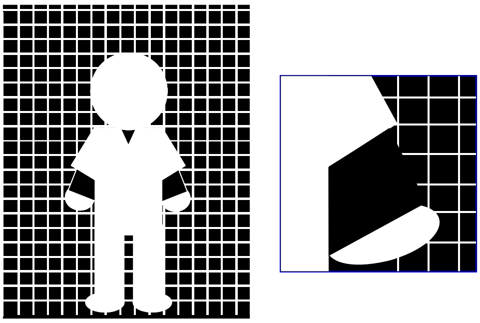


Fig.3. Endpoint detection of background grid lines to find foreground object contours.

regions form grid lines in the image, and they play a very important role to extract foreground objects.

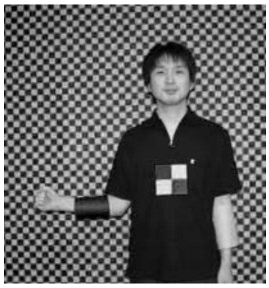
<u>Foreground extraction</u>

By executing the previous processing, pixels in the image are categorized into regions of the background color values, background grid lines, and regions of other color values than the background. Here, it should be noted that regions with the background color values do not always belong to the background, since foreground objects may have a color the same as the background color. Then, we judge whether regions with the background color values are really a background by checking their adjacency condition to background grid lines. As shown in Fig.3, inner regions of foreground objects do not connect to background grid lines, and contours (boundaries between foreground objects and background regions) touch endpoints of background grid lines.

Rough contour lines of the foreground object are given by simply connecting the background grid line endpoints on the contour. Finer contours are obtained by employing Snakes[6].

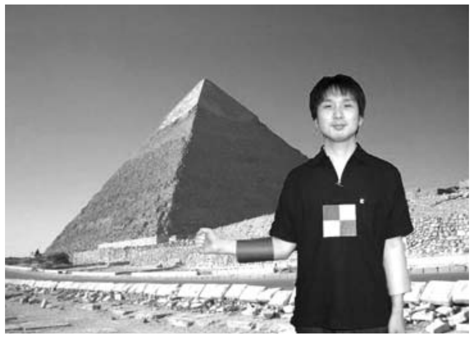
<u>Image composition</u>

The extracted foreground objects are combined to another background image to make a composite image. We determine the color $I(u, v)$ of a composite image at a pixel (u, v) by the following equation,



(a)Original image.

(b)Extracted object.



(c)Image composition.

Fig.4. Experimental result. (The contrast of checker pattern is modified so that readers can recognize the background easily in monochromatic printing.)

$$I(u, v) = \alpha(u, v) F(u, v) + (1 - \alpha(u, v)) B(u, v)$$

where $F(u, v)$ and $B(u, v)$ are the foreground and the background color, respectively, and $\alpha(u, v)$ is the so called alpha key value at a pixel (u, v). In chroma key, it is very important to determine the alpha value exactly. Methods for exact estimation of the alpha value have been proposed in applications of hair extraction, transparent glass segmentation, etc.

<u>Experiment</u>

To confirm the effectiveness of the proposed method, we made experiments using many images. Figure 4(a) is an original image example in which the foreground object wears paper sheets in the same color as the background. Figure 4(b) shows the extracted object, and Fig.4(c) is an example of image composition. The experimental results show the validity of our method.

3. Digitization of Book Pages

In recent years, digitization of existing books is performed actively, and flatbed scanners are widely used for entering page images into computer. However, in case of scanning book pages, shadings and distortions occur at the midst areas of two open pages as shown in Fig.5. We cannot obtain clear images by merely pushing the book on a scanner especially for thick books. In addition, pushing is not favorable from the viewpoint of document preservation.

To solve these problems, several 3-D methods have been proposed, e.g., based on shape from shading principle[6], based on structure from motion principle[7], and using a laser range finder[8].

In this section, we introduce a document digitization method using a stereo vision system. The method generates images without distortion by 3-D measurement[9] and also generates high-resolution, without specular reflection, and shading free images by integrating two stereo images. Figure 6 shows a system setup for stereo vision of the proposed method. The system is equipped with two cameras, two light sources and a personal computer. Book pages to digitize are opened and shot by two cameras.

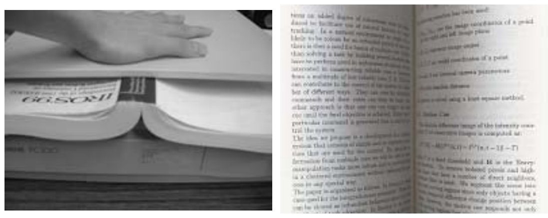


Fig.5. Scanning of book pages. The acquired image is shaded and distorted.

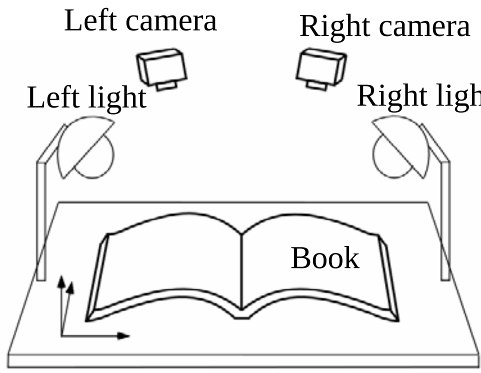


Fig.6. System setup for stereo vision.

3-D shape reconstruction

The method acquires 3-D coordinates of the book surface points by triangulation using stereo images. Then the method expresses the surface by NURBS curves which are generated by setting control points, control point's weight, and knot vectors(Fig.7). This curve fitting process is executed for two surface data sets independently, each of which is divided at inflection points between the two page surfaces.

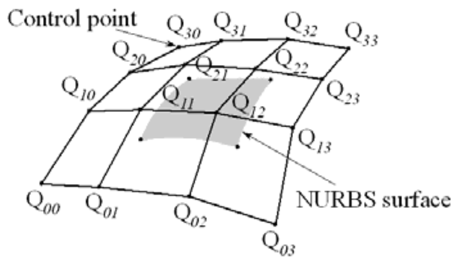


Fig.7. NURBS expression of surface.

Distortion correction

Distortion-free images of book surfaces are obtained by extending the NURBS curves to straight lines as shown in Fig.8. The pixel values along the NURBS curves are mapped onto the extended straight lines.

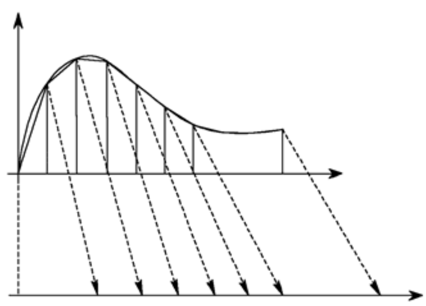


Fig.8. Extension of NURBS curves to straight lines.

Improvement of image quality

High resolution images are obtained by using the higher resolution image areas between the two camera images as shown in Fig.9.

Sometimes the obtained images have high lights owing to specular reflections. High light areas are detected by checking lighting and reflecting conditions, and the detected high light area are replaced by other images which are taken with a single light source to avoid specular reflection. The Shading effect is also removed by analyzing lighting and reflecting conditions of the book surfaces.

The difference of sensitivity between the two cameras causes discontinuity along the boundaries of combined two images. Smooth combination of images is realized by making the value of the corresponding pixels between the two images around the boundaries equal to each other.

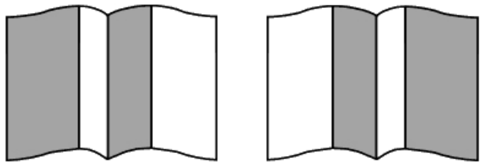


Fig.9. Integration of two camera images. The shaded areas have better resolution than those in the other image.

Experiment

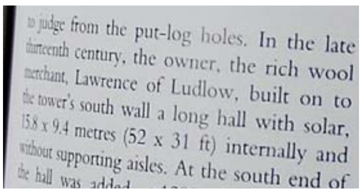
Figure 10 shows a stereo image pair acquired with digital cameras.

Figure 11 shows the 3-D measurement result. Most pixels on the book surfaces were measured appropriately. Using these 3-D points, the method expressed the two page surfaces by NURBS curves.

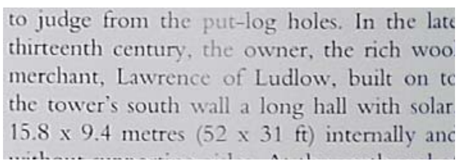
(a)Left camera image. (b)Right camera image.

Fig.10. Experimental input images.

Fig.11. Acquired 3-D points.


to judge from the put-log holes. In the late
thirteenth century, the owner, the rich wool
merchant, Lawrence of Ludlow, built on to
the tower's south wall a long hall with solar,
15.8 x 9.4 metres (52 x 31 ft) internally and
without supporting aisles. At the south end of

(a) Original distorted image.


to judge from the put-log holes. In the lat
thirteenth century, the owner, the rich woo
merchant, Lawrence of Ludlow, built on t
the tower's south wall a long hall with solar
15.8 x 9.4 metres (52 x 31 ft) internally an

(b) Resultant image.

Fig.12. Result of distortion reduction.

Fig.13. Final image.

Figure 12 shows the result of distortion reduction. Lines of letters were transformed to be horizontally parallel to one another.

Figure 13 shows the final result of the proposed method. The effects of shading and specular reflection were removed. An experiment using OCR also showed the validity of the proposed method by the fact that the recognition rate using the images of our method was much higher than using the original images.

4. Conclusions

Visual information extraction from images plays a very important role for many applications using computers in various fileds. In this paper, we introduced two examples of applications developing in our laboratory. One is a checker-pattern screen chroma key for image composition. The other is a method for digital archiving of book pages using a stereo vision system. Further progress of image processing and computer vision technology will explore much more application fields.

Acknowledgments

The works were partially supported by grant-in-aids for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology.

References

[1] http://www.cvl.iis.u-tokyo.ac.jp/mva/

[2] A. Yamashita, S. Kato and T. Kaneko: "Robust Sensing against Bubble Noises in Aquatic Environments with a Stereo Vision System", Proc. of the 2006 IEEE International Conference on Robotics and Automation (ICRA2006), pp.928-933, May 2006.

[3] Y. Tanaka, A. Yamashita, T. Kaneko and K.T.Miura:"Removal of Adherent Waterdrops from Images Acquired with a Stereo Camera System", IEICE Trans. on Information and Systems, Vol.89-D,No.7,pp.2021-2027, 2006.

[4] T. Porter and T. Duff: “Compositing digital images,” Computer Graphics (Proc. SIGGRAPH84), Vol.18, No.3, pp.253–259, 1984.

[5] H. Agata, A. Yamashita and T. Kaneko: “Chroma Key Using a Checker Pattern Background”, IEICE Trans. Inf. & Syst., Vol.E90–D, No.1 , 2007.

[6] M.Kass, A.Witkin and D. Terzopoulos, “Snakes: Active contour models,” Int.J. Comput. Vis., vol.1, no.4, pp.321–331, 1988.

[7] T. Wada, H. Ukida and T. Matsuyama: “Shape from Shading with Interreflections Under a Proximal Light Source: Distortion-Free Copying of an Unfolded Book,” International Journal of Computer Vision, Vol.24, No.2, pp.125–135, 1997.

[8] T. Sato, A. Iketani, S. Ikeda, M. Kanbara, N. Nakajima and N. Yokoya: “Mobile Video Mosaicing System for Flat and Curved Documents”, Proceedings of the 1st International Work-shop on Mobile Vision,, pp.78-92, 2006

[9] T. Amano, T. Abe, T. Iyoda, O. Nishikawa and Y. Sato: “Development of The Next Generation Document Reader –Eye Scanner–,” Proc. of the SPIE, Vol.4669, pp.250–258, 2002.

[10] A.Yamashita, A. Kawarago, T. Kaneko and K.T. Miura: “Shape Reconstruction and Image Restoration for Non-Flat Surfaces of Documents with a Stereo Vision System”,Proc.the 17th International Conference on Pattern Recognition, Vol.1, pp.482-485, 2004.