AR | Image Analysis Technology | Image Overlay Technology

Technology Reports

Special Articles on Mobile AR Technology

Mobile AR Using Image Analysis and Sensor Technologies —Providing New Expressiveness and Operational Feel to Mobile Terminals—

Communication Device Development Department Yasuo Morinaga
Manabu Ota

In the past several years, initiatives to apply AR technology to smartphones and other mobile environments have become quite active. However, they have not yet spread within the general public, and currently only a segment of interested users and companies are experimenting with ways to use the technology to provide services. Mobile terminals are equipped with all of the elements necessary to implement AR in a mobile environment; they have a camera, they can determine their own location and orientation, and they are readily available. For these reasons, we have taken the initiative to demonstrate new expressiveness and new types of operational feel with mobile terminals using image analysis and sensor technologies.

1. Introduction

Augmented Reality (AR) technology is a UI technology that overlays virtual information (content) over real space, presenting content that is not physically present at the location to the user. Recently, the number of mobile terminals equipped with high resolution cameras, accelerometers[*1], geomagnetic sensors[*2] and other types of sensors like GPS function is increasing. With these advances in mobile terminals, they have reached a level adequate for implementing AR technology in a mobile environment, and various companies have started to bring AR approaches to the market. NTT DOCOMO has advanced applications of AR in mobile terminals and has conducted tests for presenting new scenarios for using mobile terminals to users. We hope to present completely new use cases by giving mobile terminals new expressiveness and operational feel.

In this article, we introduce the basic processes in AR technology for mobile terminals (hereinafter referred to as "mobile AR") along with some service examples, and describe some future initiatives with mobile AR technology.

2. Categories of AR Technology

Categories of AR technology are shown in **Table 1**. AR technologies

*1 **Accelerometer**: A sensor that measures changes in speed. Equipping a mobile terminal with an accelerometer allows it to sense orientation and movements.

*2 **Geomagnetic sensor**: An electronic compass which can detect the geomagnetic field to determine orientation. Able to detect the orientation of a mobile terminal.

can be broadly divided into two categories; those that use image analysis to overlay content, and those that use various sensors to obtain the current location and orientation, and use it to determine what content to overlay and where to display it. The former can be divided into methods that use so-called "markers," which are like 2D barcodes, and "markerless" methods that extract feature points from images in ordinary space and use them to place overlay content. Display accuracy and other aspects could be improved further by implementing AR that combines these technologies.

3. AR Using Image Analysis

3.1 Image Analysis Technologies

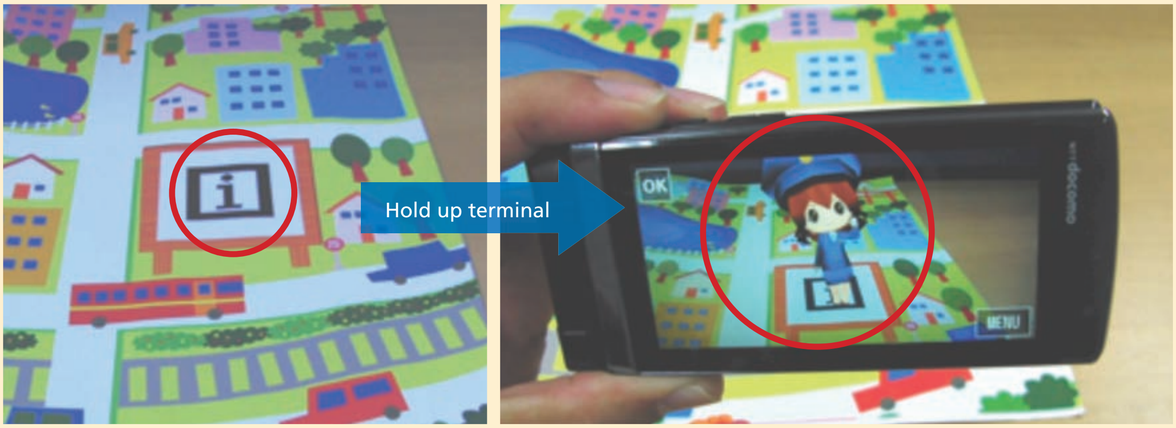
As an example of AR using image analysis, a type of AR using markers is shown in **Figure 1**. When a marker is detected by the mobile terminal camera, corresponding content is overlaid on the image. By detecting the size and any deformation in the marker, an appropriate size and orientation of the overlay content can be decided. The general process for detecting markers to overlay content over real space is shown in **Figure 2**. In this analysis, each frame captured by the camera is processed so that the position and angle can be specified in real time.

A 2D guide map application is one possible example of a service using this mechanism. Markers are placed on the map, and when they are viewed through the mobile terminal, 3D content can be displayed at the marker location, adding rich expressiveness to the guide map.

We have also conducted tests focusing on the generality of the markers and on detecting multiple markers simultaneously. We hypothesized scenarios where multiple markers would need to be detected simultaneously, including the guide map mentioned earlier and a furniture placement simulation, and enabled a mobile terminal to process them simultaneously. The processing load increases with the number of markers, but we are devising a way to improve processing speed by setting marker-detection parameters and detecting them in stages.

Table 1 Categories of AR technology

Category		Required technology
Image analysis	Markers	Pattern matching for marker shapes, detection of deformation, size
	Markerless	Extraction of feature points in acquired image
Using sensors		Obtaining position data from GPS and orientation from accelerometers and geomagnetic sensors



A 3D character is displayed when the mobile terminal camera is held up to the marker

Figure 1 Example of AR using markers

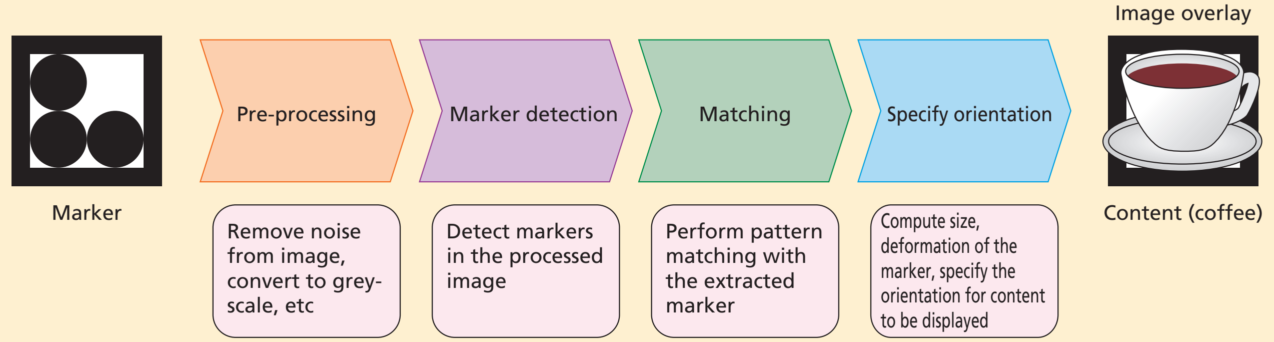


Figure 2 General image-overlay process when using markers

A very important issue with AR is determining the correct position and orientation for overlay content. Fluctuations in lighting conditions or the markers themselves can cause changes in the captured images, and especially with mobile terminals, hand-shake can cause the overlay content to flicker or be deformed.

To resolve this issue in our system, we applied a smoothing process to reduce the effects of lighting, fluctuations and hand-shake on the overlay content. In particular, we also applied a smoothing process to the fine animation of content at the same time, also improving the expressiveness of the content itself.

The smoothing parameters can also be adjusted according to the AR application, so emphasis can be placed on factors like tracking or display smoothness.

3.2 Display Control Technologies

The process in Section 3.1 can be applied to multiple markers simultaneously, but when multiple markers are detected with this approach, we must also determine the display order for content. We analyze the distance from the view point (here meaning the mobile terminal camera) to the marker using information about the size and deformation of marker images obtained by the camera and use it to decide the layering of content that will overlay each marker.

We also define the camera depth[*3] of view and use it to optimize the overlay content. We construct an area equivalent to the range of view from the viewpoint and use it to determine what content should be displayed. This reduces the amount of processing on unnecessary content and optimizes the number of items displayed.

4. Sensor Technologies

The orientation of a mobile terminal is expressed in terms of pitch angle[*4] (X-axis), roll angle[*5] (Y-axis) and yaw angle[*6] (Z-axis), and can be determined by the terminal's internal three-axis geomagnetic sensor and three-axis accelerometer (**Figure 3**). The orientation and movement of the mobile terminal are detected, and the positions of the overlay content can be specified by computing how much it is tilted away from a standard direction.

Our current approach implements a type of spatial-layout AR that allows the user to "peak" through the viewfinder of the mobile terminal into a virtual space positioned in space but separate from reality, so that it appears the other reality spreads out on the other side of the camera. Concrete screen examples are shown in **Figure 4**.

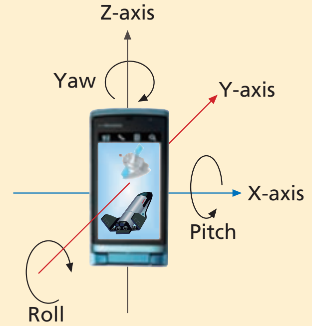


Figure 3 Orientation of a mobile terminal

*3 **Depth**: In this article, distance from the camera position.

*4 **Pitch angle**: A rotational angle in 3D space using the X-axis as the axis of rotation.

*5 **Roll angle**: A rotational angle in 3D space using the Y-axis as the axis of rotation.

*6 **Yaw angle**: A rotational angle in 3D space using the Z-axis as the axis of rotation.

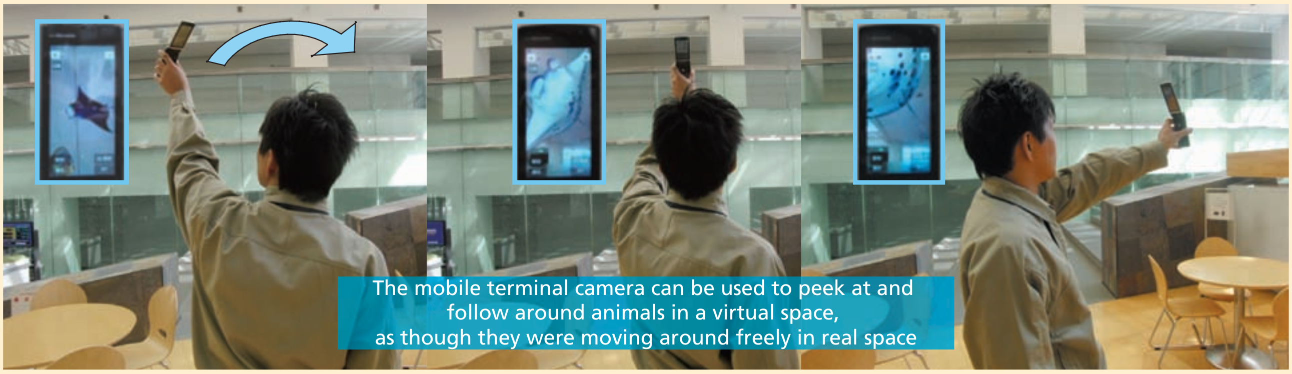


Figure 4 Example of spatial-layout AR

In order to unify these real and virtual spaces, the positions of the coordinate systems of the real and virtual spaces maintained by the mobile terminal must coincide. With spatial-layout AR, at any point in time the mobile terminal detects its orientation, and aligns the virtual coordinate system with a relative coordinate system that has the mobile terminal as its origin.

Our approach uses a geomagnetic sensor to determine the orientation of the mobile terminal, but geomagnetic sensors are easily affected by surrounding magnetic fields. For example, there are cases when the geomagnetic sensor itself cannot function properly and the correct orientation cannot be obtained in real time, such as when many nearby electronic devices create magnetic fields, or in a room surrounded by iron reinforcing bar. For mobile AR services using sensor technology, it will be important to tackle these technical issues, while looking carefully at service application areas that can accommodate them.

5. Conclusion

In this article, we have described the basic processes of AR technology, given some service examples, and described some future initiatives. In the approach described here, we focused our attention on image analysis and sensor technologies in particular, but as next steps, we are also considering scenarios for applying AR to mobile environments, linking to location data[1], and tying it to the network through a browser or Virtual Reality (VR)*7. We have also presented these at CEATEC 2010 in October, 2010. We are also testing markerless technology, examining image analysis accuracy in city environments and evaluating methods for controlling and displaying content combined with location data.

We will continue to explore mobile AR applications, including this sort of fundamental assessment, and in the future, will propose new functionality needed on mobile terminals as well as exciting new services.

REFERENCE

[1] K. Ochiai et al.: "Information Search and Presentation Methods for AR Services using Location Data," NTT DOCOMO Technical Journal, Vol.12, No.4, pp.4-10, Mar. 2011.

*7 **VR**: Technology that creates a virtual space using computer graphics on a PC or other computer.