Paper

Factorization-based Automatic 3D Face Modeling from Turntable Image Sequence using Monocular Camera

Yoshiko SUGAYA†*, Shingo ANDO†**, Isao MIYAGAWA†, Akira SUZUKI† (Member), Hideki KOIKE† (Member)
† NTT Cyber Space Laboratories, NTT Corporation

Summary Creating a 3D model of a human face from an image sequence remains an important problem in the fields of computer vision and computer graphics. We propose a method based on a factorization algorithm; it automatically constructs a 3D face model from an image sequence that contains turntable motion of the face as captured by a monocular camera. Our proposal uses multiple Active Appearance Models for feature tracking to avoid manual operations and to recover the 3D shape of the whole face. In contrast to previous methods, our approach requires no 3D generic model and can obtain the user’s individual 3D face shape. Experiments conducted on real image sequences demonstrate the effectiveness of the proposed method.

Keywords: 3D Face Modeling, Factorization Method, Active Appearance Model, Turntable Motion, Monocular Camera

1. Introduction

3D face models have become widely used in a variety of applications, such as face identification and movie making[1],2). With the spread of cell phones and other camera-equipped appliances, applications that use a 3D face model as a 3D avatar in these small appliances are receiving a lot of attention. Most current 3D face modeling systems and methods, however, still require large expensive measuring sets that force users to remain stationary for some time to obtain 3D shape information. In recent years, a number of methods have been proposed that use generic models for 3D face modeling to ease the user’s burden in the measurement process. The generic-model-based approach has the advantage that a model can be created from just a few images and large measurement sets are not required.

The 3D morphable model proposed by Blanz et al.1) can recover face shape and texture from single image by utilizing a morphable model that is created from a large quantity of 3D laser scanned face shape data in a preliminary step. 3D face shape and texture can be obtained through the iterative adjustment of the generic morphable model to the input image. Facial feature points used in the iterative matching process, however, must be manually labeled in the input image. Xin et al.3) achieved automatic 3D face modeling based on the morphable model. They introduced an automatic initialization process that estimates head pose in a video sequence as well as the parameters of the morphable model. Motion Portrait4) is a 3D face modeling application that adopts the generic-model-based approach. While the texture of the input frontal face image is automatically mapped to the face model, all 3D models have the same facial structure.

On the other hand, the structure-from-motion-based approach can acquire 3D shape information of a human face from an image sequence captured
by a monocular camera. A popular way of addressing the structure-from-motion problem is factorization methods\(\textsuperscript{5,6}\). Bregler et al.\(\textsuperscript{7}\) reconstructed 3D face shape from an image sequence exhibiting facial expressions and head pose changes by extending factorization method to three step decomposition. Although factorization algorithms can simultaneously recover structure and motion from one image sequence without any a-priori model or information about the scene, they usually require that all feature points be well tracked. Compared with the generic-model-based methods, numerous 3D modeling methods belong to the structure-from-motion-based approach utilize turntable systems to reconstruct a whole model and have demonstrated that the utilization of rotational motion geometry as an additional constraint to the factorization method improves their reconstruction performance\(\textsuperscript{8,9}\).

We propose an automatic 3D face model reconstruction method based on factorization from a turntable motion image sequence. In order to reconstruct the 3D shape of the entire face, we use an image sequence that contains horizontal rotation of the human face as shown in Fig. 1. We consider that the input sequence is captured in the situation where the user sits in front of a fixed camera and turns his or her head to the right (or left). Feature tracking, however, often fails under this rotation due to occlusion and frame-departure of features. Our solution is to divide the input image sequence into subsequences and employ different Active Appearance Models (AAM)\(\textsuperscript{10}\) for each sequence so as to prevent the dropping of feature points. Then we integrate the results from the shorter subsequences as well as the complete original sequence. Unlike the generic-model-based approach, our method can recover the individual 3D face shape since it employs factorization method. Furthermore, utilizing AAMs allows the proposed method to create a 3D face model in a fully automatic manner.

This paper is organized as follows: we start Section 2 by overviewing our proposed method, and then describe AAM-based feature tracking and factorization method with turntable motion constraint. Section 3 provides our experiments on turntable image sequences and a discussion about the noise robustness of 3D reconstruction and the individual differences in 3D faces. These are followed by the conclusion and future work in Section 4.

2. Proposed Method

2.1 Overview

Fig. 2 overviews the proposed method and shows that it consists of three steps; 1) Initialization, 2) 3D reconstruction, and 3) Post-processing. In the initialization step, the input sequence is divided into two subsequences and feature points are then automatically extracted by AAM from three sequences; i.e. the two subsequences and the entire original sequence. Local 3D face shapes are reconstructed by applying our factorization-based 3D reconstruction method to each subsequence and the whole input sequence in the 3D reconstruction step. The post-processing step generates one complete 3D face model by merging the local 3D face shapes obtained from the three sequences. The merged face then undergoes smoothing.
2.2 Initialization

The rotational motion present in the input images, shown in Fig. 1, is expected to cause the failure of conventional feature tracking because of self-occlusion and the departure of features from the frame. To avoid this loss of features, we divide the input image sequence into two short subsequences and then merge the 3D points derived from all sequences. Fig. 3 shows the usage of the three sequences: The whole sequence is used to yield the center part of the face shape. Subsequence#1 and subsequence#2 yield the right and left profiles of the face shape, respectively.

The initialization step also solves the problem with regards to sequence division that image paucity may make feature tracking difficult. We employ AAM as a feature tracker to overcome this problem. Because AAM requires no time contiguous information, it can extract features even if adjacent frames have quite different appearances.

2.2.1 Input Sequence Creation

As the captured video sequence usually consists of a large number of frames, which is one of the main causes of high computation cost, we first subsample the input sequence. First, face poses are estimated in all frames by applying face detection and the pose estimation algorithm to the captured sequence and the detected faces. After the frontal face image is identified in the captured video sequence, one image every few frames from the frontal face image is extracted to create the input image sequence (the head is assumed to rotate around the vertical plane). Next, the image sequence is divided into two subsequences; one runs from the first frame to the frontal face, and the other runs from the frontal face to the last frame.

2.2.2 Feature Tracking

AAM well suits deformable objects like human faces but its fitting performance strongly depends on the training set used to build the model. AAM can fit any face direction in any light condition if the training set covers all face directions and light conditions. In the proposed method, three AAMs for every sequence are precomposed in consideration of this characteristic. One is for the whole sequence to extract some feature points of facial parts, the others, for subsequence#1 and subsequence#2, are intended to extract feature points in facial parts and facial contour in right (or left) profile. This utilization of multiple AAMs allows us to deal with many more features in the subsequences than is possible if the whole sequence is used. As a feature tracker, AAM is applied to each frame in each image sequence.

Fig. 4 shows examples of the AAM training images. Fig. 4 (a) is one of the training samples for the entire sequence; Fig. 4 (b) and 4 (c) are for subsequence#1 and subsequence#2, respectively. Each training sample contains labeled points as shown in Fig. 4. These points are extracted in the feature tracking process.

2.3 3D Reconstruction

As, in the structure-from-motion problem, high accuracy is generally achieved when the camera motion is large enough, small camera motion caused from the sequence division in the initialization step may degrade 3D reconstruction accuracy. To prevent this problem, we introduce a 3D reconstruction method.
that specifically targets turntable motion and that yields high accuracy from a small number of images.

We can assume without loss of the generality that the human face motion in the input sequence shown in Fig. 1, which represents only horizontal rotation with no vertical movement, can be taken as the motion of the object on turntable. This assumption allows us to solve factorization decomposition under the perspective camera model. We describe below the factorization method using the turntable motion constraint.

2.3.1 Factorization Method with Turntable Motion Constraint

In order to recover the 3D shape of the object on a turntable, we set the world coordinates XYZ on the turntable as shown in Fig. 5 (a). For convenience, we place the world origin at the center of mass of the object and fix the camera at position (0,0,R). We also assume that the object on the turntable rotates about the Y axis (Fig. 5 (b)). Without loss of generality, camera motion is considered as rotation of the object on turntable. We denote that the 3D point lying on the object \( P_j = (X_j, Y_j, Z_j) \) rotated by \( \theta_i \) is \( P_{ji} = (X_{ji}, Y_{ji}, Z_{ji}) \) and that \( (x_{ij}, y_{ij}) \) is the observed 2D projection point of \( P_j \).

Under the pinhole camera model, the 2D projection point \( (x_{ij}, y_{ij}) \) with focal length \( f \) can be described as follows:

\[
\begin{bmatrix}
    x_{ij} \\
    y_{ij}
\end{bmatrix} = \begin{bmatrix}
    fX_{ji} / (Z_{ji} - R) \\
    fY_{ji} / (Z_{ji} - R)
\end{bmatrix}, \tag{1}
\]

\[
\begin{bmatrix}
    X_j \\
    Y_j \\
    Z_j
\end{bmatrix} = \begin{bmatrix}
    \cos(\theta_i) & 0 & -\sin(\theta_i) \\
    0 & 1 & 0 \\
    \sin(\theta_i) & 0 & \cos(\theta_i)
\end{bmatrix} \begin{bmatrix}
    X_j \\
    Y_j \\
    Z_j
\end{bmatrix}. \tag{2}
\]

Combining equation (1) and (2) yields

\[
\lambda_{ij} \begin{bmatrix}
    x_{ij} \\
    y_{ij}
\end{bmatrix} = \begin{bmatrix}
    \cos(\theta_i) & 0 & -\sin(\theta_i) \\
    0 & 1 & 0 \\
    \sin(\theta_i) & 0 & \cos(\theta_i)
\end{bmatrix} \begin{bmatrix}
    X_j/R \\
    Y_j/R \\
    Z_j/R
\end{bmatrix}, \tag{3}
\]

where

\[
\lambda_{ij} = \frac{1}{f} \left( \frac{X_j}{R} \sin(\theta_i) + \frac{Z_j}{R} \cos(\theta_i) - 1 \right). \tag{4}
\]

When \( N \) features are being tracked in \( F \) image frames, equation (3) for all points and frames can now be combined into the next measurement matrix.

\[
W = \begin{bmatrix}
    \lambda_{11} \cdot \lambda_{1N} & \cdots & \lambda_{1N} \cdot \lambda_{1N} \\
    \lambda_{21} \cdot \lambda_{2N} & \cdots & \lambda_{2N} \cdot \lambda_{2N} \\
    \vdots & \ddots & \vdots \\
    \lambda_{F1} \cdot \lambda_{F1} & \cdots & \lambda_{F1} \cdot \lambda_{F1} \\
    \lambda_{F2} \cdot \lambda_{F2} & \cdots & \lambda_{F2} \cdot \lambda_{F2} \\
    \vdots & \ddots & \vdots \\
    \lambda_{FN} \cdot \lambda_{FN} & \cdots & \lambda_{FN} \cdot \lambda_{FN}
\end{bmatrix} = MS. \tag{5}
\]

\[
M = \begin{bmatrix}
    \cos(\theta_1) & 0 & -\sin(\theta_1) \\
    \cos(\theta_2) & 0 & -\sin(\theta_2) \\
    \vdots & \vdots & \vdots \\
    \cos(\theta_F) & 0 & -\sin(\theta_F) \\
    0 & 1 & 0 \\
    0 & 1 & 0 \\
    \vdots & \vdots & \vdots \\
    0 & 1 & 0
\end{bmatrix} = \begin{bmatrix}
    m_1^T \\
    m_2^T \\
    \vdots \\
    m_F^T \\
    n_1^T \\
    n_2^T \\
    \vdots \\
    n_F^T
\end{bmatrix}, \tag{6}
\]

\[
S = \begin{bmatrix}
    X_1/R & X_2/R & \cdots & X_N/R \\
    Y_1/R & Y_2/R & \cdots & Y_N/R \\
    Z_1/R & Z_2/R & \cdots & Z_N/R
\end{bmatrix} = \begin{bmatrix}
    s_1 & s_2 & \cdots & s_N
\end{bmatrix}. \tag{7}
\]

These three equations indicate that the rank of \( W \) is at most 3 because \( W \) is decomposed into \( M (2F \times 3) \) and \( S (3 \times N) \). If coefficients \( \lambda_{ij} \) are known, factorization decomposition based on equations (5), (6) and
ear transformation

appropriate

This decomposition of (5) is unique, we have to find $S$ as follows:

$$\begin{bmatrix} X_j \\ Y_j \\ Z_j \end{bmatrix} = R \begin{bmatrix} s_{j1} \\ s_{j2} \\ s_{j3} \end{bmatrix}. \quad (8)$$

The decomposition of (5) is determined by just linear transformation $Q$ ($3 \times 3$). In order to ensure that this decomposition of (5) is unique, we have to find appropriate $Q$ that transforms $M$ and $S$ into $\hat{M}$ and $\hat{S}$ as follows:

$$\hat{M} = MQ, \quad \hat{S} = SQ^{-1}. \quad (9)$$

An appropriate $Q$ is determined by using the following metric constraints based on the turntable motion which is described by equation (6).

$$m_i^T Q^2 m_i = 1, \quad (10)$$
$$n_i^T Q^2 n_i = 1, \quad (11)$$
$$m_i^T Q^2 n_i = 0. \quad (12)$$

As equations (10), (11) and (12) give $3F$ equations, we can obtain $Q$. After finding $Q$, we consider the desired camera motion $M$ as $\hat{M}$ and the desired 3D shape $S$ as $\hat{S}$.

Note that we assumed that the coefficients $\lambda_{ij}$ are known in the above discussion. Our proposed method iteratively updates the coefficients $\lambda_{ij}$ by applying an iterative stabilization algorithm\(^{13}\). Although the iterative stabilization algorithm was originally proposed for modest deviation from planar motion, its principle is also applicable to the turntable motion discussed here.

### 2.3.2 Iterative Stabilization Algorithm

In order to apply the iterative stabilization algorithm to rotational motion, the initial value of $\lambda_{ij}$ is set to $-1$, ($i = 1, 2, \cdots, F; j = 1, 2, \cdots, N$). $M$ and $S$ are then calculated by factorization decomposition based on equations (10), (11) and (12). When the elements in matrix $M$ are represented as $m_i = (m_{i1}, m_{i2}, m_{i3})$, rotation angle $\theta_i$ in the $i$-th frame is obtained as follows:

$$\theta_i = \tan^{-1} \left( \frac{m_{i3}}{m_{i1}} \right). \quad (13)$$

Next $\lambda_{ij}$ is calculated by substituting $X_j/R$, $Y_j/R$, and $\theta_i$ into equation (4). The measurement matrix is recalculated as shown by equation (5) when $\lambda_{ij}$ is updated. After that, renewed $M$ and $S$ are computed by decomposing the updated measurement matrix as mentioned in Section 2.3.1. We repeat this iterative decomposition until the rank of $W$ becomes 3.

This factorization method based on the iterative stabilization algorithm is summarized as follows:

**step1** Initialize $\lambda_{ij} = -1$.

**step2** Decompose the measurement matrix by using singular value decomposition (SVD).

**step3** Compute the $3 \times 3$ matrix, $Q$, that satisfies the proposed metric constraints, then uniquely determine $M$ and $S$ from (9).

**step4** Calculate rotational angle $\theta_i$ using (13).

**step5** Update $\lambda_{ij}$ according to (4).

**step6** If the rank 4 value of the measurement matrix is sufficiently small, then stop, otherwise update each element of the measurement matrix and return to step2.

**step7** Recover the object shape in Euclidean space with (8).

### 2.4 Post-processing

The reconstructed three face shapes have different local coordinates. In order to acquire one final shape by merging the partial shapes, we unify the different local coordinates into one common coordinate by calculating conversion matrices from the correspondence of feature points using unit quaternions\(^{14}\). After conversion of all partial shapes in each local coordinate, the final 3D face shape is created by smoothing the overlapping feature points.

**Fig. 6** (a) is the recovered center part of the face shape from the whole image sequence and **Fig. 6** (b)
is the merged shape. It is readily understand that the shape recovered from the whole sequence (Fig. 6 (a)) is only a part of the entire face. This is largely because only a few feature points are visible in all frames. As it can be seen from Fig. 6 (b), the merged 3D shape covers a wide area of the face.

The number of vertices in the merged final shape may not be enough to express a human face because the AAMs used in feature tracking have sparse feature points. To make the 3D face model smooth, we apply the loop subdivision scheme after merging the 3D shapes. Finally, the frontal face image texture is mapped to the smoothed 3D shape.

3. Experiments and Results

3.1 3D Reconstruction Results

To evaluate the performance of the 3D reconstruction method described in Section 2.3, we conducted experiments using real image sequences of a 100 mm cube which was placed on a turntable system. We performed experiments to examine the two camera setups shown in Fig. 7. In the first experiment, the object sat on the center of the turntable (Fig. 7 (a)). In the other experiment, the object sat on the edge of the turntable (Fig. 7 (b)). Several test images are shown in Fig. 8. Both test sequences consist of 11 frames with roughly 5 degree increments and 45 feature points were manually tracked in every frame. The test image sequences were captured by a Point Grey GRAS-50S5C-C camera (IEEE1394b) with a resolution of 1,600 × 1,200 pixels. The distance of the turntable center to the camera $R$ was 460 mm. We estimated the intrinsic parameters by using Zhang’s method.

We compared the proposal to the iterative factorization method proposed by Christy and Horaud (C-H method). This method achieves perspective factorization by iteratively modifying a projection model to yield an affine model by minimizing the 2D reprojection errors.

Fig. 9 plots the recovered object rotation angle for both methods. It can be seen that both methods recovered the roughly 5 degree increments turntable motion. Table 1 lists the average 3D reconstruc-

![Frame #01](image1)
![Frame #06](image2)
![Frame #10](image3)

Fig. 8 Samples of test image sequences

![Frame #01](image4)
![Frame #06](image5)
![Frame #10](image6)

Fig. 9 Recovered turntable angle

<table>
<thead>
<tr>
<th>Table 1 3D reconstruction error [Unit: mm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-H method</td>
</tr>
<tr>
<td>------------</td>
</tr>
<tr>
<td>Average</td>
</tr>
<tr>
<td>X axis</td>
</tr>
<tr>
<td>Y axis</td>
</tr>
<tr>
<td>Z axis</td>
</tr>
<tr>
<td>(a)</td>
</tr>
<tr>
<td>Average</td>
</tr>
<tr>
<td>X axis</td>
</tr>
<tr>
<td>Y axis</td>
</tr>
<tr>
<td>Z axis</td>
</tr>
<tr>
<td>(b)</td>
</tr>
</tbody>
</table>
tion error (spatial deviation) of the proposed and existing methods. As shown in Table 1, our method has smaller 3D reconstruction error than the existing method. This shows that the turntable motion constraints significantly improve the reconstruction accuracy.

To illustrate the behavior of the proposed iteration, the rank4 value of the measurement matrix is shown in Fig. 10. It is readily understand from Fig. 10 that the rank4 value converges in just a few iterations.

3.2 Modeling Results

To validate our modeling method, an experiment was conducted on human face image sequences. In this experiment, 13 images of a human face from different orientations in the HOIP Database were used as the input image sequence. Fig. 1 shows some of the input images. These face images represent horizontal rotation in 5 degree increments with no vertical movement. The camera parameters and capture environment of the database are as follows: focal length $f$ is 38 mm, the distance of the turntable axis to the camera $R$ is 1500 mm. The input sequence was divided into two subsequences where the number of frames in both subsequences was set to 7 with 1 image overlap. The 7th frame is detected as the frontal face image by the face detection method.

That is, the first subsequence consisted of the initial frame to the 7th frame, and the second subsequence consisted of the 7th frame to the last frame. AAMs for each sequence were built using 210 face images, 30 people captured from 7 directions, as the training set. Each AAM consisted of 69 feature points. We applied the proposed method to the 234 image sequences. Fig. 11 shows different views of the created 3D model. This model consists of 766 vertices and 1485 triangular patches. Fig. 12 shows the 3D face model reconstruction results for 4 people. From the left column to right column: Fig. 12 (a) is the side view of the model shape, 12 (b) is the front view of the model shape, 12 (c) is the front view of the 3D shape with texture and 12 (d) is the frontal face image in the input sequence. As it can be observed from Fig. 12 (a) and 12 (b), the proposed method can recover individual 3D face shape.

3.3 Discussion

Conventional structure-from-motion based methods have difficulty in recovering the tip of the nose and the curvature of the cheeks, whereas our method can reconstruct nose shape and cheek curvature as can be seen in Fig. 11. This is largely because we utilizes two characteristics of AAM in the feature tracking process: 1) AAM fitting performance strongly depends on the training images used to build the model. 2) AAM fitting ensures that feature points are never lost because every feature point in AAM contains weak relationships with adjacent feature points to construct triangular patches. To recover nose and cheek shape, feature points must be set on the tip of nose, cheeks and facial contour in the training images when precomposing multiple AAMs as shown in Fig. 4.

Even though the feature points in nose and cheek areas are extracted in every frame, it is inevitable that the obtained tracking results include much more noise.
than those of other features. To achieve noise robustness, we utilized the factorization method with the turntable motion constraint which includes a noise reduction process. As the 3D reconstruction method described in Section 2.3 is based on factorization method, feature tracking noise is reduced through the iteration of SVD and rank 4 value removal. Fig. 10 indicates that the tracking noise, which appears in the rank 4 value of the measurement matrix, is decreased by a few iteration. Despite the fact that the experiments conducted on the cube image sequences include tracking noise caused by manually labeled feature points, the 3D shape recovery error is less than 4.9 mm as shown in Table 1.

In the experiments in Section 3.2, 156 face models were successfully created by the proposed method while 78 modeling attempts failed due to failure of the feature tracking process. The main reasons for these failures fall into 2 categories; one is occlusion by glasses, beards or hair over the forehead, and the other is facial motion such as eye blink or facial expression changes. It would appear that feature
tracking performance declined because the training sets used for building AAMs in the experiments did not contain face images with glasses, beards, various frontal hair and facial expression changes. As our method assumes that human face in the input image sequence is a rigid object, the 3D shape reconstruction method described in Section 2.3 is applicable to face modeling. We can improve the face modeling system performance by introducing eye blink detection to remove images that include some eye blinks and utilizing AAM training samples that include a variety of facial expression changes, frontal hair, glasses, and beards etc.

4. Conclusion

In this paper, we proposed an automatic 3D face model reconstruction method based on factorization from a turntable image sequence. Our technique requires no generic model and no manual operations in the face modeling process. Utilizing multiple AAMs, we fully automated the feature tracking process and can extract a lot of feature points on the human face, which allows our method to recover the whole face area.

We implemented the proposed method and conducted experiments using real image sequences. The experiments showed that applying the turntable motion constraints to factorization method is effective for improving reconstruction performance and that our modeling method can recover the whole face area. Additionally, each 3D face model well expresses the structural individuality of the individual.

To further investigate the possibilities of the proposed approach, we aim to implement it with a USB camera and employ the created 3D face models for face identification.

References

4) Motion Portrait (http://www.motionportrait.com/)
18) HOIP Face Database (http://www.softopia.or.jp/rd/facedb.html)

(Received July 16, 2010)
(Revised Oct. 27, 2010)

Yoshiko Sugaya She received the B.E. and the M.E. degrees in mechanical engineering from Sophia University, Tokyo, Japan, in 2005 and 2007, respectively. She joined NTT Corporation in 2007. She has been engaged in research on computer vision and projector-camera systems from 2007 to 2010. She is now with Nippon Telegraph and Telephone East Corporation. She is a member of IEICE.
Shingo Ando  He received the B.E. degree in electrical engineering and the Ph.D. degree in engineering from Keio University, Kanagawa, Japan, in 1998 and 2003, respectively. He joined NTT Corporation in 2003. He has been engaged in research and practical application development in the fields of image processing and pattern recognition. He is now with Nippon Telegraph and Telephone West Corporation. He is a member of IEICE.

Isao Miyagawa  He received the B.E. degree in electronics engineering from Fukui University, Fukui, Japan, in 1991, and the Ph.D. degree in engineering from the Graduate School of Engineering of Fukui University, Fukui, Japan, in 2005. He joined NTT Corporation in 1991 and had developed communication technology for color documents and G3 facsimile devices from 1991 to 1996. He has studied computer vision since 1997. He is currently a Research Engineer of NTT Cyber Space Laboratories. He is a member of IEEE and IEICE.

Akira Suzuki  (Member) He received the B.E., M.E., and Dr.Eng. degrees in engineering from Shizuoka University, Shizuoka, Japan, in 1983, 1985, and 1998, respectively. He joined NTT Corporation in 1985. He has been engaged in research and practical application development in the fields of document retrieval, character recognition, knowledge processing, and image processing. He is currently a Senior Research Engineer of NTT Cyber Space Laboratories. He is a member of IIEEJ, IEICE, and ISSJ.

Hideki Koike  (Member) He received the M.S. degree in mathematics from Tohoku University, Miyagi, Japan, in 1985. He joined NTT Corporation in 1985 and engaged in research on image processing technologies. He moved to NTT Comware Corporation in 2001 and engaged in developing RFID systems. He is currently a Senior Research Engineer, Supervisor, and Group Leader of the Visual Media Communications Project of NTT Cyber Space Laboratories. He is a member of IEEE and IEICE.