Optical Engineering 44(12), 127001 (December 2005)

Compression of ray information in three-dimensional integral imaging

Ju-Seog Jang
Seokwon Yeom
Bahram Javidi
University of Connecticut
Electrical and Computer Engineering Department
371 Fairfield Road, U-2157
Storrs, Connecticut 06269-2157
E-mail: bahram@engr.uconn.edu

Abstract. We present a method to compress ray information in three-dimensional (3-D) integral imaging (II) using the Karhunen-Loeve transform (KLT). Elemental images in II are highly correlated because they are picked-up by numerous lenslets, thus, the KLT can compress an integral image more effectively than 2-D ordinary images. In the hybrid coding scheme, the KLT is distinctly applied to disjoint subsets which are partitioned by a vector quantization (VQ). In the optical experiments, we show that 3-D image quality degradation is negligible if approximately 10% of eigenvectors with the largest eigenvalues are used. We also evaluate the quality of uncompressed images by signal-to-noise ratio (SNR) and peak-to-peak signal-to-noise ratio (PSNR). The presented coding scheme is shown to be better than JPEG (Joint Photographic Experts Group) for the lower bit rates in the II compression. Optical and numerical experiments are presented. [DOI: 10.1117/1.2148947]

Subject terms: three-dimensional imaging; image compression; image reconstruction-restoration; three-dimensional image processing; integral imaging; Karhunen-Loeve transform; vector quantization.

Paper 040252RR received May 6, 2004; revised manuscript received Apr. 8, 2005; accepted for publication May 11, 2005; published online Jan. 5, 2006.

1 Introduction

To record and produce true three-dimensional (3-D) images in space using two-dimensional (2-D) display devices and incoherent light sources, integral imaging (II), or real-time integral photography (IP), has been investigated.[1–5] In II, 3-D images are formed by the intersection of discrete rays coming from an integral image (a set of 2-D elemental images), which is a 2-D representation of ray information that is sampled spatially at various lenslet locations using a lenslet array (Fig. 1). As in holography, the display of II can provide observers with true 3-D images with full parallax and continuous viewing points.

However, to produce high-resolution large-scale 3-D integral images, an array of a large number of lenslets is required.[5] In addition, each elemental image should have enough resolution.[5] Therefore, transmission and storage of an integral image may require substantial bandwidth and memory space.[6,7] Therefore, it is obvious that compression of ray information is important. There were prior efforts in this research direction. To compress line images, which are needed for a display of 3-D images with only horizontal parallax using a lenticular lens sheet, a discrete cosine transform was used.[8] Recently, considering elemental images as consecutive frames, an MPEG (Moving Picture Experts Group) scheme was applied to a small number of full parallax elemental images.[9]

In this paper, we present a method to compress integral images using a hybrid Karhunen-Loeve transform (KLT) coding. Lenslets have a small pitch and they are closely positioned, thus, the ray information in an integral image is highly correlated. The KLT can be more effective in compression of the ray information because it is dependent of data statistics.[10–13]

In the hybrid coding scheme, the KLT is applied distinctly to disjoint subsets in one integral image. The partition of the integral image is performed by a vector quantization (VQ). The hybrid scheme has been researched to show better result than the compression of either[12,13] VQ or KLT. At the final stage, rounding-off is followed to allocate proper size of bits to reduced data. Figure 2 shows a block diagram of the hybrid KLT coding scheme.

Using 3-D display experiments, we show that the KLT can compress ray information in II. We evaluate the quality degradation of the reconstructed image by signal-to-noise ratio (SNR) and peak-to-peak signal-to-noise ratio (PSNR). We also investigate the effects of the size of image blocks and the number of subsets in the partition. Finally, we compare the presented technique with the Joint Photographic Experts Group[14] (JPEG).

Although the hybrid KLT has been widely used for many other applications including data compression,[10–13] to the best of our knowledge, this is the first report on application of the hybrid KLT to 3-D II compression.

2 Benefits of Elemental Image Compression in II

For 3-D image pickup in II, direction and intensity information of the rays coming from a 3-D object is spatially sampled by use of a lenslet array and a 2-D image sensor, as depicted in Fig. 1. The size of a lenslet element is usually[2] 1 to 2 mm. The ray information sampled by each lenslet appears as a small 2-D image with its own perspective. So it is referred to as an elemental image. Reconstruction of the 3-D image of the object from 2-D elemental

0091-3286/2005/$22.00

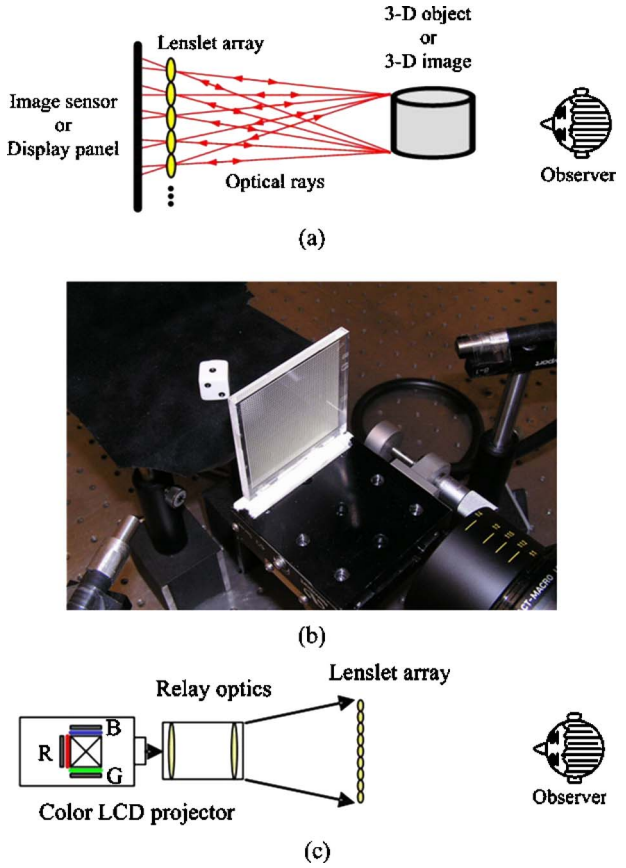


Fig. 1 Pickup of 3-D images in II: (a) schematic diagram, (b) optical setup, (c) 3-D object reconstruction using color LCD projector.

images is a reverse of the pickup process. The recorded 2-D elemental images are displayed on a 2-D display panel, such as a liquid crystal display (LCD) panel, and then rays coming from the elemental images are redirected to form a real image where multiple rays cross.

The product of depth-of-focus and lateral resolution square (PDLRS) of 3-D images in II is limited by $1/\lambda$, where λ is the illumination wavelength.[15] This means that if we display a 3-D image with large depth, we have to sacrifice the resolution, and vice versa. It is difficult to obtain 3-D image resolution that is 1 or 2 lines/mm, when the 3-D image depth is around 2 or 4 m in the visible wavelength range. Although the absolute image resolution is low, the viewing resolution can be improved by increasing the distance between observers and the 3-D image. In this case, an important parameter that should be considered for image quality is the total number of voxels in 3-D images to be displayed, which is approximately the number of 2-D image pixels ($N_x \times N_y$) in lateral dimensions multiplied by the number of depth levels (N_z).

To produce high-quality 3-D images with a large depth, elemental images with a large number of pixels are required. Suppose we display a 3-D image with a total number of voxels $N_x \times N_y \times N_z \approx 10^3 \times 10^3 \times 10^3$. Then, it is obvious that we need a display panel with more than $m_r \times 10^9$ pixels, where m_r is the ray multiplicity factor to form a voxel. This is because each voxel in II is determined by a crossing point of multiple rays (extension lines of multiple

Fig. 2 Block diagram of the hybrid KLT coding scheme.

rays for virtual image). Each ray is originated from a pixel of a distinct elemental image in the display panel through the corresponding lenslet. In II, at least 4 pixels from distinct elemental images comprise one voxel to show full-parallax view.[10] Because a voxel is a small virtual space where multiple rays cross in the 3-D space, we cannot define the shape of the voxel. The factor m_r for a given voxel is determined by both the longitudinal distance between the voxel and the lenslet array and $f/\#$ of the lenslet.[5]

The question to be addressed is how many lenslets (elemental images) are needed for a visually pleasing quality of display? When the gap between the lenslet array and the display panel is close to the focal length of the lenslets (we call this[15] depth-priority II), the lateral size of a point image equals the lenslet size. Assume we display the elemental images on a 1 m $\times$ 1 m LCD, which is located at 1.5 m in front of the viewer. Then, the pitch of a lenslet becomes about 1 mm to fulfill the minimum resolution for eyes.[17,18] Therefore, 3-D images are displayed with $N_x \times N_y = 10^3 \times 10^3$ pixels, and we require 10^6 lenslets (elemental images) for the depth-priority II. Because the number of depth levels (N_z) is linearly proportional to the number of lenslets in one lateral direction (N_x or N_y), approximately 10^9 voxels can be displayed.

To realize supermultiview, there should be more than 40 to 60 views between two eyes.[17,18] If we assume there are

Table 1 LBG-VQ procedures

	Procedures
Step 0	Set Δ (splitting vector) and S (desired number of subsets), and let $s=1$.
	Compute a centroid (codebook vector) of I_0. $\mathbf{v}_s = 1/N \Sigma_{i=1}^{N} \mathbf{r}_i$.
Step 1	Splitting each codebook vector to two initial centroids for $j=1,\ldots,s$: $\mathbf{v}_j = \mathbf{v}_j + \Delta$, $\mathbf{v}_{s+j} = \mathbf{v}_j - \Delta$, then set $s=2s$.
Step 2	1. Set $n_j=0$.
	2. For all vectors (image blocks) in I_0, find a subset that minimizes Euclidean distance from the vector to the codebook vector: $\hat{j} = \arg\min_j \|\mathbf{r}_i - \mathbf{v}_j\|$, $j=1,\ldots,s$, and update the subset: $\Omega_{\hat{j}} = \{\Omega_{\hat{j}} \mathbf{r}_i\}$, and set $n_j = n_j + 1$.
	3. For all s subsets, update codebook vectors: $\mathbf{v}_j = 1/n_j \Sigma_{i=1}^{n_j} \mathbf{r}_i^j$, where $\Omega_j = \{\mathbf{r}_1^j, \ldots, \mathbf{r}_{n_j}^j\}$.
	4. If there is any change of codebook vectors after substeps 2 and 3, go to substep 1, otherwise, go to step 1 until $s=S$.

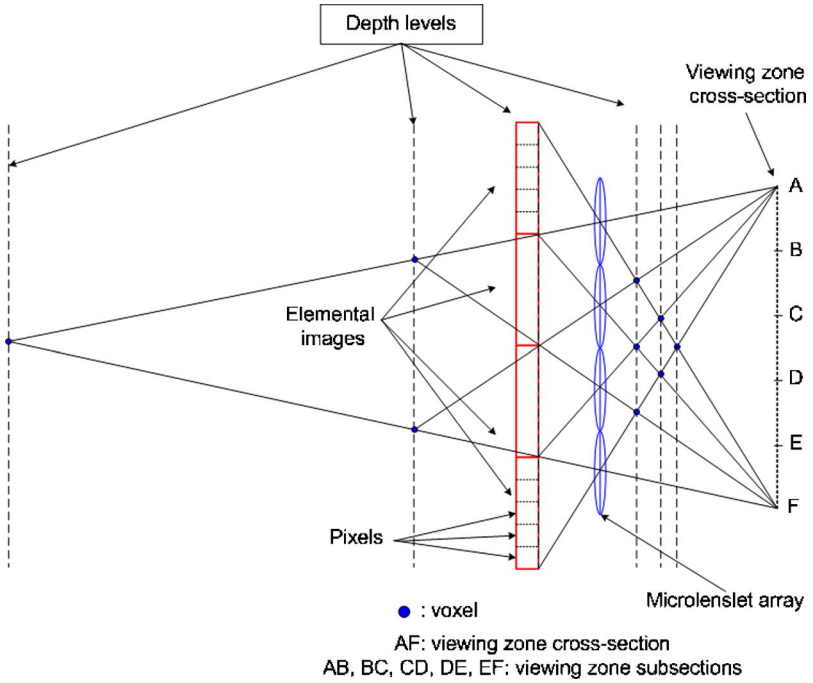


Fig. 3 Schematic diagram in two dimensions for voxels and viewing zone.

20 viewing zone subsections between two eyes of which distance is 6.5 cm, we require about 300×300 viewing zone subsections for the full-parallax view ($300\approx1$ m $\times20$ viewing zone subsections/6.5 cm). Because the number of viewing zone subsections corresponds to the number of pixels in one elemental image, we need a total of 9×10^{10} pixels ($300\times300\times10^3\times10^3$) to produce 10^9 voxels, which means $m_r=90$ on average.

Figure 3 shows a 2-D schematic diagram for voxels and viewing zone;[16,19] 4 lenslets comprise six depth levels (N_z) including the display panel and the number of viewing zone subsections is the same with the number of pixels in one elemental image; circles represent voxels; dashed lines represent depth levels; and a dotted line represents a viewing zone cross section. More detailed discussion about voxel and viewing zone is found in Refs. 16 and 19.

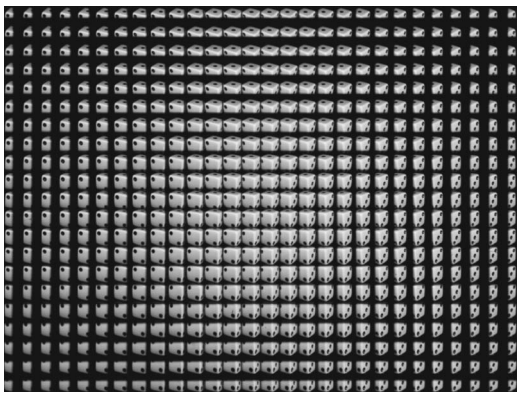


Fig. 4 Elemental images obtained from direct camera pickup of a die.

In depth-priority II, the total number of elemental images is much larger than the number of pixels in each elemental image, and their product (i.e., the total number of pixels in the entire set of elemental images) is a very large number. As a consequence, storage and transmission of images can be significant problems, therefore, proper compression scheme should be considered to reduce the large amount of II data. Because elemental images are highly

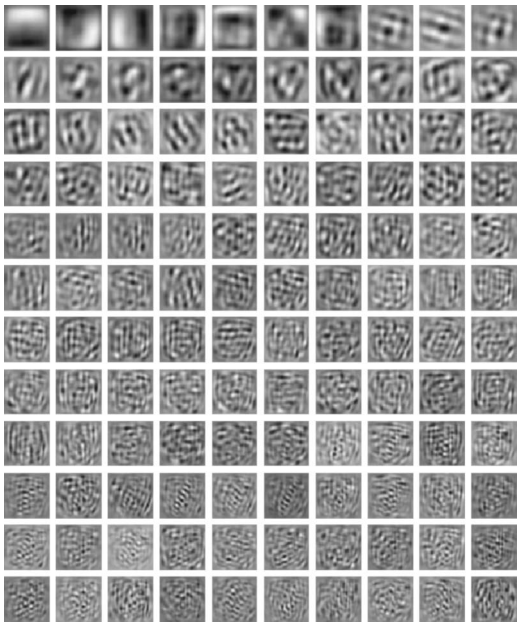


Fig. 5 First 120 eigenvectors of an elemental image represented by gray-scale images in a sequence of left to right, and then top to bottom.

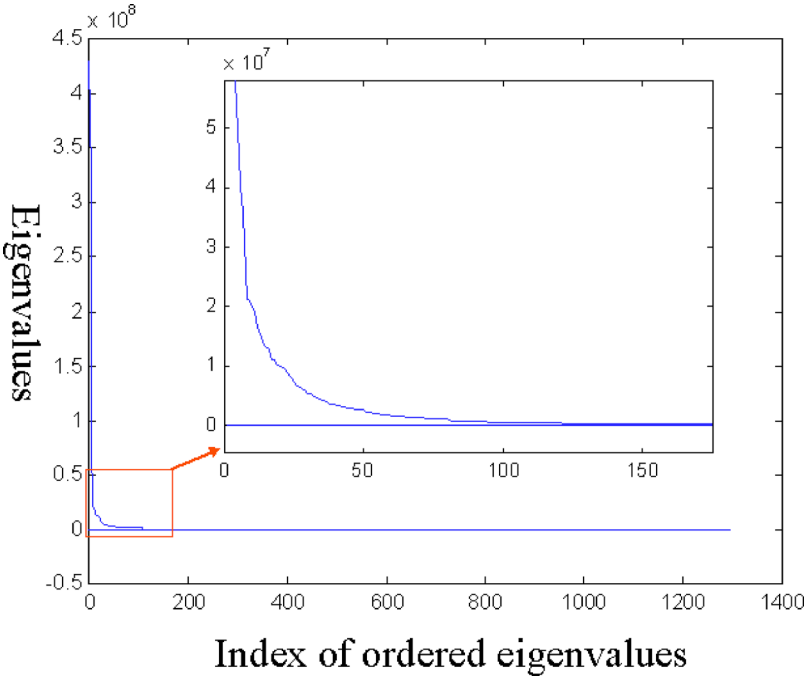


Fig. 6 Distribution of eigenvalues.

correlated, the KLT, which has been widely used for projecting a set of vectors or images onto a low dimensional subspace can be used efficiently to reduce the ray information in II. We present the hybrid compression using the KLT in the following sections.

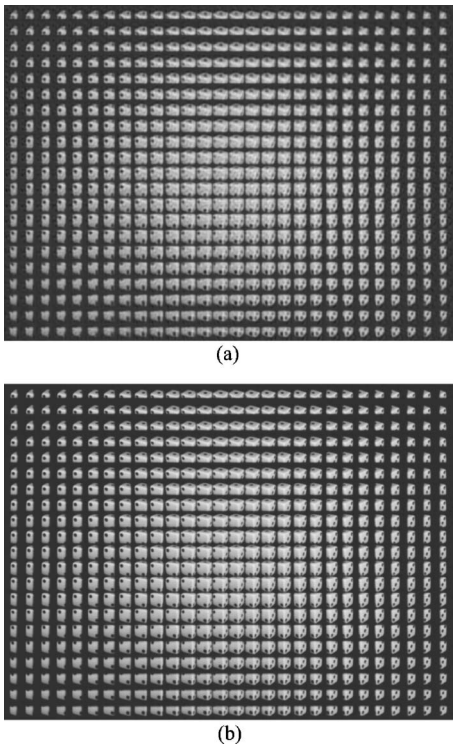


Fig. 7 Reconstructed elemental images after KLT compression without rounding-off (S=1 and d=37): (a) k_j=30 and (b) k_j=180.

3 Hybrid Compression Using KLT

In this section, we present an overview of the compression algorithms applied to II.

3.1 *VQ*

An integral image I_o is assumed to be composed of N block images of size $d \times d$ pixels. The n'th block image is considered a $M(=d \times d)$-dimensional vector $\mathbf{r}_n$, that is, $\mathbf{r}_n = [P_1^n \cdots P_M^n]^t$, where P_i^n is the i'th pixel value in the n'th block image; and the superscript t denotes matrix transpose. Let us assume that I_o is partitioned by S disjoint subsets Ω_j, $j=1, \cdots, S$ and $\Omega_j = \{\mathbf{r}_1^j, \cdots, \mathbf{r}_{n_j}^j\}$, where n_j is the number of block images in the subset Ω_j.

VQ generates a set of codebook vectors to satisfy Lloyd's optimality condition.[20] Each codebook vector is a representative of vectors in each disjoint subset. Lloyd's optimality condition based on Euclidean distance is the following[12,20]:

1. All vectors in the subset Ω_j have a smaller Euclidean distance to the codebook vector $\mathbf{v}_j$ than any other codebook vector:

$$\Omega_j = \{\mathbf{r} | \|\mathbf{r} - \mathbf{v}_j\| < \|\mathbf{r} - \mathbf{v}_i\|, \forall_j \neq i\}, \tag{1}$$

   where $\|\cdot\|$ stands for Euclidean distance.
2. The codebook vector $\mathbf{v}_j$ for the subset Ω_j is a centroid of the subset that is the sample mean vector:

$$\mathbf{v}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} \mathbf{r}_i^j. \tag{2}$$

We adopt the Linde-Buzo-Gray (LBG) algorithm to solve the preceding two optimality conditions.[21] In the

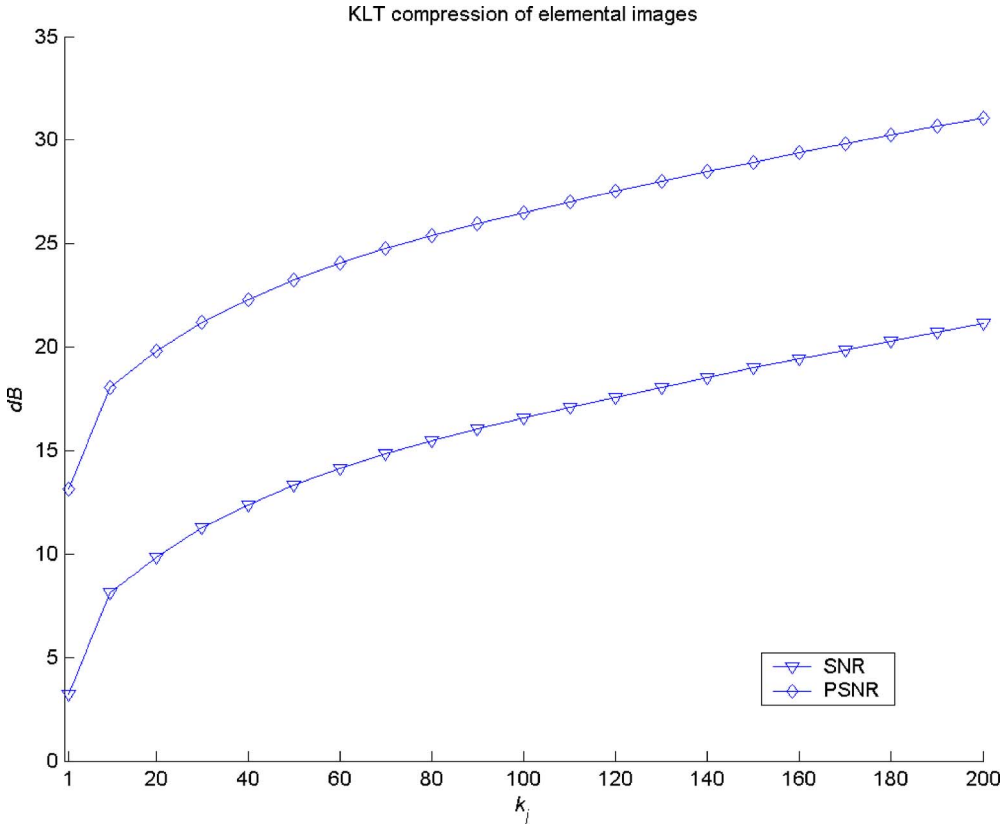


Fig. 8 KLT compression of elemental images without rounding-off (S=1 and d=37).

LBG-VQ, we set an initial codebook vector as the sample mean of entire data, and then split it into two initial codebook vectors. An iterative algorithm generates final partition and codebook vectors. Then, each final codebook vector is split to two and then new partition is performed again by the iterative algorithm to generate new final codebook vectors. We repeat this process until the desired number of codebook vectors is obtained. Table 1 shows the LBG-VQ procedures.

3.2 *KLT*

After M block images are divided into S disjoint subsets during the VQ process, we distinctly apply the KLT to each subset. The sample mean vector of the subset Ω_j is equivalent to the codebook vector $\mathbf{v}_j$. The covariance matrix of the subset Ω_j is obtained as:

$$C_j = \frac{1}{n_j - 1} \sum_{i=1}^{n_j} (\mathbf{r}_i^j - \mathbf{v}_j)(\mathbf{r}_i^j - \mathbf{v}_j)^t. \tag{3}$$

The KLT matrix for the subset Ω_j is defined as: $\Gamma_j = [\mathbf{e}_1^j, \ldots, \mathbf{e}_{k_j}^j]$, where k_j is the number of eigenvectors; and $\mathbf{e}_1^j, \ldots, \mathbf{e}_{k_j}^j$ are the k_j eigenvectors of the C_j. They are corresponding to the k_j largest eigenvalues of the C_j. The KLT achieves dimensionality reduction by projecting an M-dimensional vector $\mathbf{r}_i^j$ onto the k_j-dimensional subspace. The projected vector is:

$$\mathbf{w}_i^j = \Gamma_j^t(\mathbf{r}_i^j - \mathbf{v}_j), \; i = 1, \ldots, n_j. \tag{4}$$

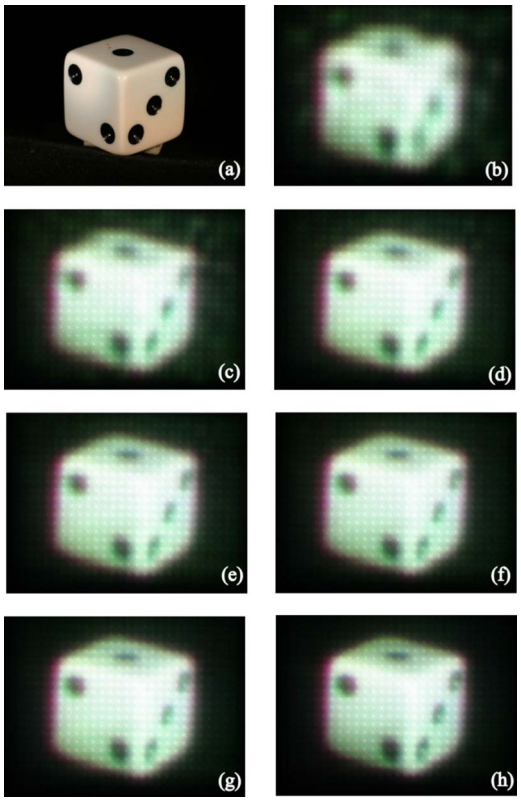


Fig. 9 Optically reconstructed 3-D images using the first k_j eigenvectors without rounding-off (S=1 and d=37): (a) original object, (b) k_j=30, (c) k_j=60, (d) k_j=90, (e) k_j=120, (f) k_j=150, (g) k_j=180, and (h) when original elemental images are used.

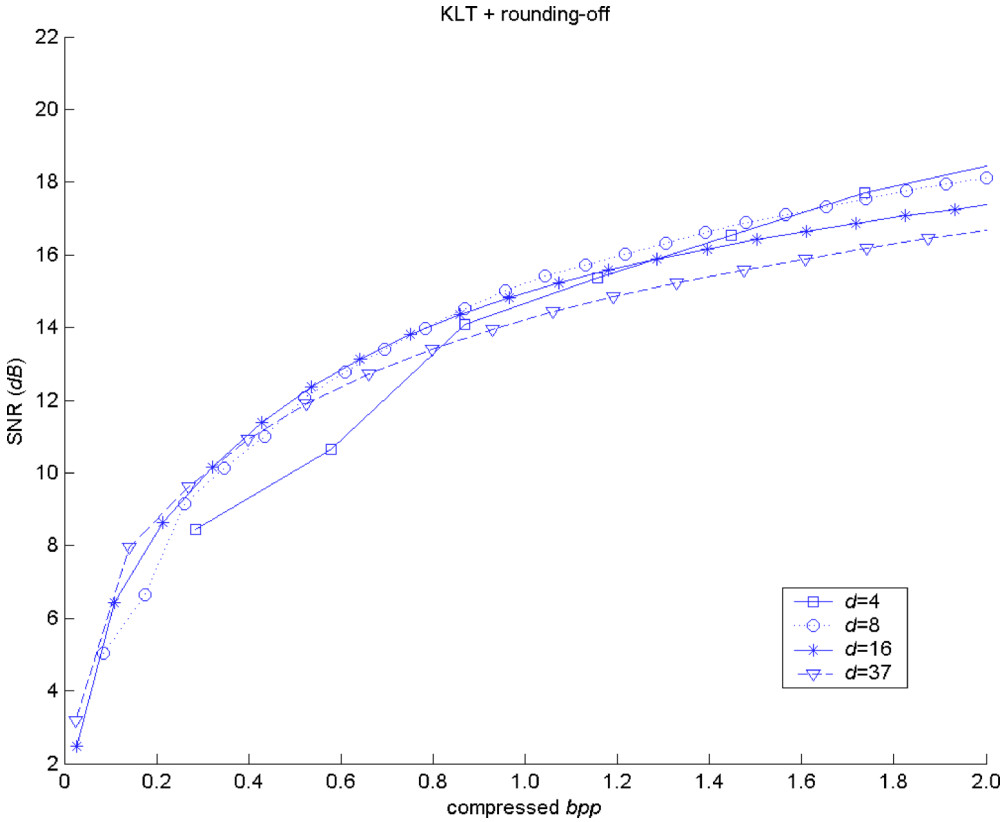


Fig. 10 KLT compression with rounding-off (S=1 and d=4, 8, 16, and 37).

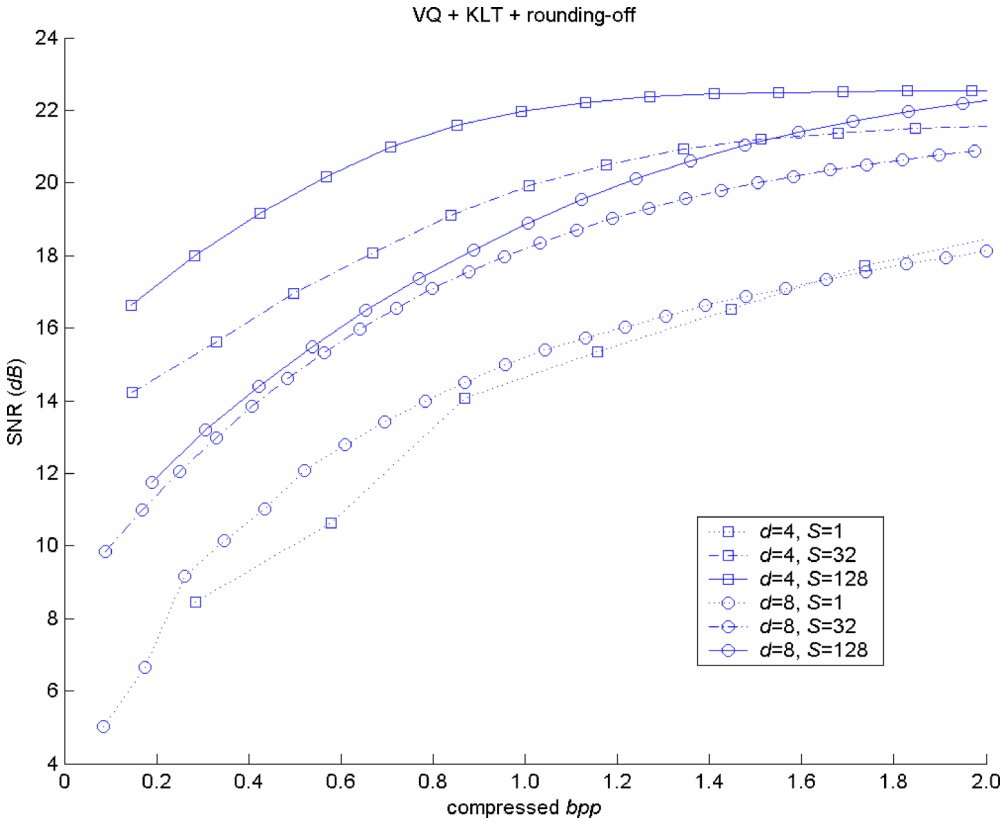


Fig. 11 Hybrid KLT compression (VQ+KLT+rounding-off) for d=4 and 8, and S=1, 32, and 128.

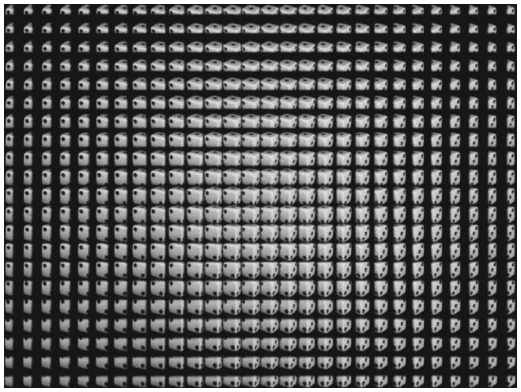


Fig. 12 Uncompressed image when $d=4$, $S=128$, and $T_w=0.1$, resulting in SNR=16.63 dB, PSNR=26.56 dB, and bit rate =0.145 bpp.

The mean-squared error (MSE) is the summation of the MSE for the subset Ω_j:

$$E=\sum_{j=1}^{S}E_j=\sum_{j=1}^{S}\frac{1}{n_j}\sum_{i=1}^{n_j}\|\mathbf{r}_i^j-\hat{\mathbf{r}}_i^j\|^2, \tag{5}$$

where E is total MSE and E_j is the MSE of the subset Ω_j. It is well known that the KLT is a linear transformation which minimizes E_j, when $\hat{\mathbf{r}}_i^j$ is defined as

$$\hat{\mathbf{r}}_i^j=\Gamma_j\mathbf{w}_i^j+\mathbf{v}_j=\Gamma_j\Gamma_j^t(\mathbf{r}_i^j-\mathbf{v}_j)+\mathbf{v}_j,\ i=1,\ldots,n_j. \tag{6}$$

3.3 *Rounding-Off*

At the final stage, we perform "rounding-off" to quantize reduced vectors, KLT matrices, and codebook vectors. By rounding-off the data, we can allocate proper size of bits to each component of reduced ray information. In 2-D discrete cosine transform (DCT) compression such as JPEG, a scalar quantizer is used with predetermined codebook tables. However, adequate quantization tables for the KLT are hard to synthesize because of its dependency on image data. In this paper, we round-off the information to the nearest integer after multiplying proper scale factors.

Therefore, the reconstructed (uncompressed) vector (after rounding-off) is obtained as:

$$\tilde{\mathbf{r}}_i^j=\frac{\langle T_\Gamma\Gamma_j\rangle\langle T_w\mathbf{w}_i^j\rangle}{T_\Gamma T_w}+\frac{\langle T_v\mathbf{v}_j\rangle}{T_v}, \tag{7}$$

where $\langle\cdot\rangle$ is an operator which stands for the nearest integer; and T_Γ, T_w, and T_v are scale factors for the KLT matrixes, the projected coefficient vectors, and the codebook vectors, respectively. The corresponding MSE with rounding-off is

$$\tilde{E}=\sum_{j=1}^{S}\frac{1}{n_j}\sum_{i=1}^{n_j}\|\mathbf{r}_i^j-\tilde{\mathbf{r}}_i^j\|^2, \tag{8}$$

where $\tilde{E}$ is the MSE with rounding-off. Uncompressed integral image I_u is composed of reconstructed vectors $\tilde{\mathbf{r}}_i^j$, $i=1,\ldots,n_j$, and $j=1,\ldots,S$.

4 Performance Evaluation

We can use the k_j eigenvectors for the n_j block images in the subject Ω_j, instead of using the original $M\times n_j$ pixels. In this case, the KLT matrix $(M\times k_j)$ and the codebook vector $(M\times 1)$ as well as the n_j projected vectors $(k_j\times 1)$ are required to reconstruct ray information for the subset Ω_j. As a result, we reduce the information from NM pixels to $M(\Sigma_{j=1}^{S}k_j+1)+\Sigma_{j=1}^{S}n_jk_j$ coefficients. After rounding them off, the number of bits B, which are allocated to the reduced data becomes:

$$B=M\sum_{j=1}^{S}[k_jQ_\Gamma(j)+Q_v(j)]+\sum_{j=1}^{S}n_jk_jQ_w(j), \tag{9}$$

where $Q_\Gamma(j)$, ,$Q_w(j)$, and $Q_v(j)$ are defined as:

$$Q_\Gamma(j)=\log_2\lfloor\max(T_\Gamma\Gamma_j)-\min(T_\Gamma\Gamma_j)\rfloor, \tag{10}$$

$$Q_w(j)=\log_2\lfloor\max(T_w[\mathbf{w}_1\cdots\mathbf{w}_{n_j}])-\min(T_w[\mathbf{w}_1\cdots\mathbf{w}_{n_j}])\rfloor, \tag{11}$$

$$Q_v(j)=\log_2\lfloor\max(T_v\mathbf{v}_j)-\min(T_v\mathbf{v}_j)\rfloor, \tag{12}$$

where $\max(\cdot)$ and $\min(\cdot)$ stand for a maximum and minimum value in the parenthesis, respectively. Therefore, the compressed bit rate in (bits per pixel) (bpp) can be defined as

$$\text{bit rate}=\frac{M\sum_{j=1}^{S}[k_jQ_\Gamma(j)+Q_v(j)]+\sum_{j=1}^{S}n_jk_jQ_r(j)}{MN}(\text{bpp}). \tag{13}$$

To evaluate image quality, we compute two numerical values, SNR and PSNR:

$$\text{SNR}(I_o,I_u)=10\log_{10}\left[\frac{\text{VAR}(I_o)}{\text{MSE}(I_o,I_u)}\right](\text{dB}), \tag{14}$$

where $\text{VAR}(I_o)$ is the variance of the original integral image I_o:

$$\text{VAR}(I_o)=\frac{1}{MN}\sum_{n=1}^{MN}\|p_n-\bar{p}\|^2. \tag{15}$$

In Eq. (15), p_n is the n'th pixel; $\bar{p}$ is the sample mean of the pixels in I_o; $\text{MSE}(I_o,I_u)$ is the MSE between I_o and I_u:

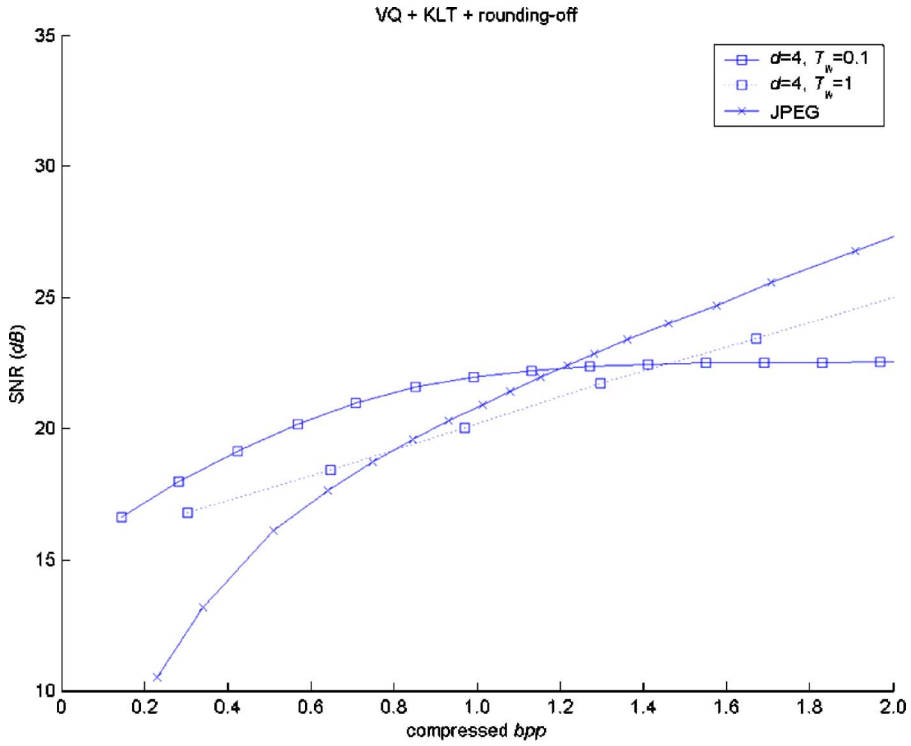
$$\text{MSE}(I_o,I_u)=\frac{1}{MN}\sum_{j=1}^{S}\sum_{i=1}^{n_j}\|\mathbf{r}_i^j-\mathbf{r}_i^j\|^2=\frac{1}{MN}\sum_{n=1}^{MN}\|p_n-\tilde{p}_n\|^2, \tag{16}$$

where $\tilde{p}_n$ is the reconstructed pixel corresponding to p_n.

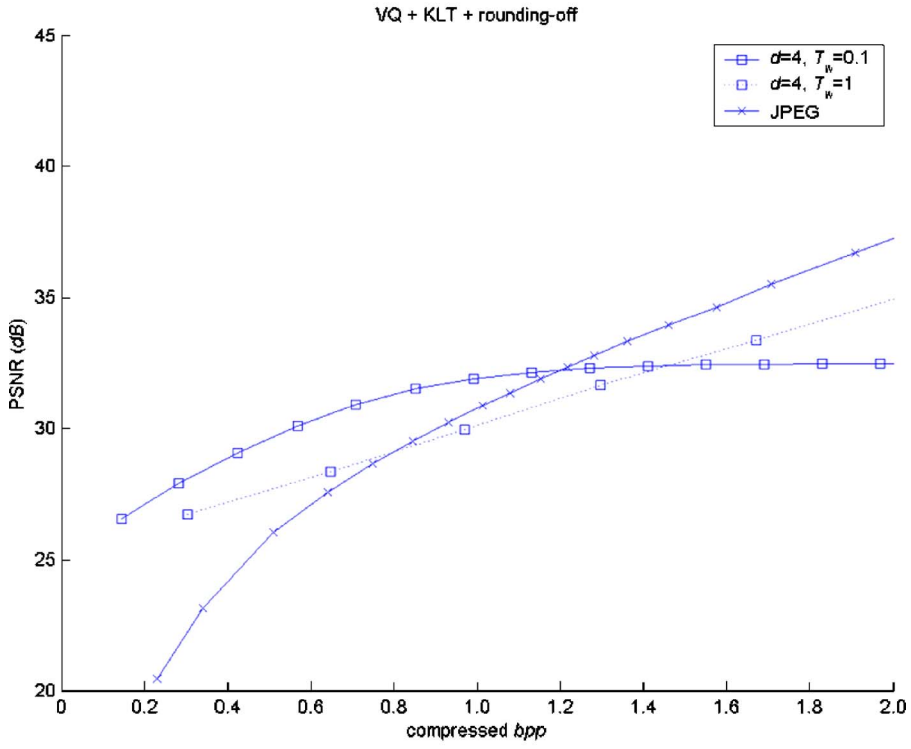
PSNR is defined as:

$$\text{PSNR}(I_o,I_u)=10\log_{10}\left[\frac{p_{\max}^2}{\text{MSE}(I_o,I_u)}\right](\text{dB}), \tag{17}$$

where $p_{\max}$ is the maximum value of the pixel in I_o.



(a)



(b)

Fig. 13 Hybrid KLT compression (d=4, S=128, and T_r=0.1 and 1) and JPEG compression: (a) SNR and (b) PSNR.

Since the quality evaluation of 3-D imaging is deeply related with subjective and perceptual concerns, we use two common metrics for 2-D image compression, assuming the quality of elemental images is proportional to 3-D display performance. In the experiments, we also present 3-D display figures to help evaluate 3-D display performance in terms of SNR and PSNR.

5 Experimental Results

The object to be imaged is a die with a side length of 1.5 cm. We obtain elemental images using a lenslet array with 53×53 planoconvex lenslets. Each lenslet element is square-shaped and has a uniform base size of 1.09 $\times1.09$ mm, with less than 7.6 μm separating the lenslet elements. The focal length of the lenslets is approximately 3 mm. The elemental images of the object were obtained by direct camera pickup [Fig. 1(b)], and then they are converted by rotating every elemental image around its own center in a computer for orthoscopic 3-D image reconstruction. Resulting elemental images are shown in Fig 4. A total of 28×21 elemental images are used in our experiments ($N=588$). Each elemental image has 37×37 pixels. The ray information is recorded in TIFF (tagged image file format) by allocating 6 bit to one pixel, thus p_{max} in Eq. (17) is 63.

In the first experiment, we present compression results for the KLT of elemental images without rounding-off, i.e., each image block corresponds to one elemental image ($S=1$, $d=37$). Figure 5 shows the first 120 eigenvectors calculated from the 588 elemental images, in which they are represented by gray-scale images with suitable dc bias to avoid negative values. Eigenvectors with very small eigenvalues appear as faint random noise patterns and they are not shown here. The distribution of eigenvalues is shown in Fig. 6. The number of eigenvectors (k_j) that have significant eigenvalues is about 150. For comparison, the reconstructed (uncompressed) elemental images when $k_j=30$ and $k_j=180$ are illustrated in Fig. 7. SNR and PSNR for $k_j=30$ are 11.28 and 21.21 dB, respectively, 18.99 and 28.93 dB for $k_j=150$, respectively, and 20.31 and 30.24 dB for $k_j=180$, respectively. Figure 8 shows the SNR and PSNR for the KLT compression varying numbers of eigenvectors ($k_j=1,10,20,\ldots,200$).

In Fig. 9, 3-D images are reconstructed optically using the uncompressed elemental images for $k_j=30,60,\ldots,180$. A 2-D color LCD projector that has 3 (RGB) panels was used to display the elemental images [Fig. 1(c)]. We used a projection type of II system, in which elemental images are directly projected onto the lenslet array.[5] Each panel has 1024×768 square pixels with a pixel pitch of 18 μm. One can see that the quality of optically reconstructed 3-D images using the 150 eigenvectors [Fig. 9(f)] is close to that of using the original elemental images [Fig. 9(h)]. The results of the optical reconstruction are reasonable given the quality of commercially available lenslets and LCD panels. The poor quality of image reconstruction in Fig. 9(b) is due to the substantial image compression applied to the elemental images. The other reconstructions in Fig. 9 with lower compression ratios produce reasonable quality images.

In the second experiment, we validate the performance of the KLT compression with the hybrid scheme. Figure 10 shows the compression results for the KLT without VQ ($S=1$), and different sizes of image blocks ($d=4$, 8, 16, and 37). It shows almost similar performance for different d. The scale factors T_Γ, T_w, and T_v are set at 100, 0.1, and 10, respectively. The scale factors are chosen heuristically when better results are produced. We produce different bpp by changing $k_j=1,2,\ldots,7$ for $d=4$; $k_j=1,2,\ldots,24$ for $d=16$; $k_j=1,4,8,\ldots,76$ for $d=16$; $k_j=1,10,20,\ldots,140$ for $d=37$; and k_j is set at the same value for all subsets.

Figure 11 shows SNR for the hybrid KLT scheme for different block size ($d=4$ and 8) and different number of subsets ($S=1$, 32, and 128). The quality of the reconstructed image is much higher when the number of subsets is increased, however, no benefit of partition is found when d is 16 or 37. Figure 12 shows the uncompressed image when $d=4$, $k_j=1$, $S=128$, and $T_w=0.1$, resulting in SNR is 16.63 dB; PSNR is 26.56 dB, and bit rate is 0.145 bpp.

In Fig. 13, we show the compression results of the different scale factor ($T_w=0.1$ and 1) and compare them with the JPEG. The baseline JPEG compression is performed by a built-in function in MATLAB R12.1. We change the "Quality" factor from 1, 5, 10, 80 in converting TIFF to JPEG format. The compressed bit rate for the JPEG is computed by dividing the image size in JPEG format by the total number of pixels in the original image. The hybrid KLT scheme shows better performance for the lower bit rate, even without entropy coding. Therefore, it is clear that the KLT can compress the ray information more effectively.

6 Conclusion and Future Plan

We showed that the KLT can be efficiently used for the compression of ray information in 3-D integral images. Optical experiments were presented to demonstrate ray compression for 3-D display. Numerical criteria were also presented to evaluate the performance. As the image block size became smaller and the number of subsets increased, the hybrid KLT coding showed better results. By adjusting the scale factors, we improved the quality of the image significantly.

The computational complexity of the encoding process using the VQ and KLT may be prohibitive for real-time application, thus, the presented coding scheme is more suitable when only real-time decoding is required such as an II movie or 3-D image retrieval and storage. However, there have been more investigations to reduce the computational complexity[22,23] to make the encoding task closer to practical usage.

Some tasks remain for this research for future study; although the hybrid coding scheme is optimal in the VQ and the KLT, it did not aim at minimizing SNR or PSNR by which the compression performance is evaluated. We may develop an adaptive scheme of varying parameters to solve this problem.

References

1. G. Lippmann, "La photographie integrale," *Comptes-Rendus Acad. Sci.* **146**, 446–451 (1908).
2. C. B. Burckhardt, "Optimum parameters and resolution limitation of integral photography," *J. Opt. Soc. Am.* **58**, 71–76 (1968).
3. N. Davies, M. McCormick, and L. Yang, "Three-dimensional imaging systems: a new development," *Appl. Opt.* **27**, 4520–4528 (1988).
4. F. Okano, H. Hoshino, J. Arai, and I. Yuyama, "Real-time pickup method for a three-dimensional image based on integral photography," *Appl. Opt.* **36**, 1598–1603 (1997).

5. J.-S. Jang, Y.-S. Oh, and B. Javidi, "Spatiotemporally multiplexed integral imaging projector for large-scale high-resolution three-dimensional display," *Opt. Express* **12**, 557–563 (2004).
6. T. Okoshi, *Three Dimensional Imaging Techniques*, Academic Press, London (1976).
7. T. Motoki, H. Isono, and I. Yuyama, "Present status of three-dimensional television research," *Proc. IEEE* **83**, 1009-1021 (1995).
8. R. Zaharia, A. Aggoun, and M. McCormick, "Adaptive 3D-DCT compression algorithm for continuous parallax 3D integral imaging," *Signal Process. Image Commun.* **17**, 231–242 (2002).
9. S. Yeom, A. Stern, and B. Javidi, "Compression of 3D color integral images," *Opt. Express* **12**, 1632–1642 (2004).
10. K. Fukushima, *Introduction to Statistical Pattern Recognition*, Academic Press, New York (1990).
11. M. Rabbani, Ed., *Selected Papers on Image Coding and Compression*, Milestone Series MS48, SPIE Press, Belingham, WA (1992).
12. N. Kambhatla and T. K. Leen, "Dimension reduction by local principal component analysis," *Neural Comput.* **9**, 1493–1516 (1997).
13. R. D. Dony and S. Haykin, "Optimally adaptive transform coding," *IEEE Trans. Image Process.* **4**, 1358–1370 (1995).
14. V. Bhaskaran and K. Konstantinides, *Image and Video Compression Standards, Algorithms and Architectures*, 2nd ed., Kluwer Academic Publishers, Boston (1997).
15. J.-S. Jang, F. Jin, and B. Javidi, "Three-dimensional integral imaging with large depth of focus using real and virtual image fields," *Opt. Lett.* **28**, 1421–1423 (2003).
16. J.-Y. Son, "Display requirements for 3 dimensional images," *Int. in Proc. Workshop on Optical Display and Information Processing*, pp. 16–30, Gyeongju, Korea (2002).
17. T. Izumi, Supervising Ed., *Fundamental of 3-D Image Technique*, NHK Science and Technology Lab., Ohmsa, Tokyo (1995).
18. J.-Y. Son, B. Javidi, and V. V. Saveljev, "Synthesizing 3D images based on voxels," *Proc. SPIE* **5202**, 1–11 (2003).
19. J.-Y. Son, V. V. Saveljev, J.-S. Kim, S.-S. Kim, and B. Javidi, "Viewing zones in three-dimensional imaging systems based on lenticular, parallax-barrier, and microlens-array plates," *Appl. Opt.* **43**, 4985–4992 (2004).
20. S. P. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inf. Theory* **IT-28**, 129–137 (1982).
21. Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Trans. Commun.* **COM-28**, 84–95 (1980).
22. A. Levy and M. Lindenbaum, "Sequential Karhunen-Loeve basis extraction and its application to images," *IEEE Trans. Image Process.* **9**, 1371–1374 (2000).
23. A. Weingessel and K. Hornik, "Local PCA algorithms," *IEEE Trans. Neural Netw.* **11**, 1242–1250 (2000).

Ju-Seog Jang received his BS degree in electromechanical engineering from Pusan National University, Korea, in 1984 and his MS and PhD degrees, both in electrical engineering, from the Korea Advanced Institute of Science and Technology, Seoul, in 1986 and 1989, respectively. From 1989 to 1991, he was a research associate at the Joint Institute for Laboratory Astrophysics at the University of Colorado, Boulder. From 1991 to 1993, he was a senior researcher at the Electronics and Telecommunications Research Institute, Taejon, Korea. In 1993, he joined the Division of Electronics, Computers, and Telecommunications at Pukyong National University. In 1994 and 1995, he was a visiting faculty member in computation and neural systems at the California Institute of Technology, Pasadena. In 2001 and 2002, he was a visiting faculty member in the Department of Electrical and Computer Engineering at the University of Connecticut, Storrs. He passed away in June 2004.

Seokwon Yeom received his BS degree in physics from Inha University, Inchon, Korea in 1995. He received his BS and MS degrees from the Department of Electronics and Computer Engineering, Korea University, Seoul, Korea, in 1997 and 1999, respectively. He also received his MS degree from the Department of Electrical and Computer Engineering, University of Connecticut, Storrs, Connecticut, in 2003. He is currently a PhD candidate at the same university. His research interests are pattern recognition, optical information processing, image segmentation, image compression, and target tracking.

Bahram Javidi is the Board of Trustees Distinguished Professor at the University of Connecticut. He is fellow of the Institute of Electrical and Electronics Engineers (IEEE), fellow of the Optical Society of America (OSA), and fellow of the International Society for Optical Engineering (SPIE). Recently, he was the winner of the SPIE Dennis Gabor award. In 1990, the National Science Foundation named him a Presidential Young Investigator. He has received the IEEE Lasers and Electro-optics Society Distinguished Lecturer Award twice in 2003 and 2004. He was the recipient of the IEEE Best Journal Paper Award from IEEE Transactions on Vehicular Technology in 2002. He has been awarded the University of Connecticut Board of Trustees Distinguished Professor Award, the School Of Engineering Distinguished Professor Award, University of Connecticut Alumni Association Excellence in Research Award, the Chancellor's Research Excellence Award, and the first Electrical and Computer Engineering Department Outstanding Research Award. He has completed several books including *Smart Imaging Systems* published by SPIE Press in 2001 and *Optical Pattern Recognition* published by SPIE Press in 1994. Dr. Javidi has published over 200 technical articles in major journals. He has published over 250 conference proceedings, including over 90 keynote and invited conference papers, and 60 invited presentations.