Three-dimensional computational holographic imaging and recognition using independent component analysis

BY CUONG MANH DO* AND BAHRAM JAVIDI

Department of Electrical and Computer Engineering, University of Connecticut, 371 Fairfield Road U-2157, Storrs, CT 06269-2157, USA

We present computational holographic three-dimensional imaging and automated object recognition based on independent component analysis (ICA). Three-dimensional sensing of the scene is performed by computational holographic imaging of the objects using phase-shifting digital holography. We used principal components analysis to reduce data dimension and ICA to recognize the three-dimensional objects. In this paper, kurtosis maximization-based algorithm is used. To the best of our knowledge, this paper is the first to report using ICA in three-dimensional imaging technology.

Keywords: three-dimensional image sensing; three-dimensional coherent imaging; holography; independent component analysis; three-dimensional object recognition

1. Introduction

Three-dimensional object recognition is of great interest because of enhanced classification performance compared with two-dimensional imaging. A number of approaches have been suggested for three-dimensional automated object recognition, including digital holography (Javidi & Tajahuerce 2000).

Digital holography is a technique used to record three-dimensional objects using the interference between an object wave and a reference wave captured by an image sensor such as a charged couple device (CCD) array (Goodman 1996; Kreis 2005). The three-dimensional images are reconstructed numerically using digital techniques. The benefit of digital holography is that we can focus on any section of the three-dimensional volume object without mechanical focusing adjustment. Phase-shifting digital holography with multiple exposures is used to remove the DC term and conjugate images from the holograms. Multispectral three-dimensional computational holographic imaging and object reconstruction have been demonstrated using image fusion (Javidi & Okano 2002; Do et al. 2005; Javidi et al. 2005, 2006a; Kwon & Nasrabadi 2005; Frauel et al. 2006; Do & Javidi 2007), and three-dimensional object detection and recognition have been explored (Sadjadi 2000; Javidi 2002; Mahalonobis et al. 2004; Javidi et al. 2006b; Sadjadi & Mahalonobis 2006).

Recently, independent component analysis (ICA) technique has gained attention as an effective method to extract statistically independent components.

* Author for correspondence (cuong.do@uconn.edu).

Received 10 August 2007
Accepted 5 November 2007

This journal is © 2007 The Royal Society
from observed data. ICA has been applied to two-dimensional image processing. However, to the best of our knowledge, ICA has not been investigated in three-dimensional image recognition using computational holographic imaging. In this paper, we examine the benefits of ICA in three-dimensional image recognition using computational holographic imaging.

The structure of this paper is as follows. Section 2 provides fundamental information about the phase-shifting digital holography technique. Section 3 presents an overview of the FastICA algorithm. We discuss three-dimensional object recognition in §4. Section 5 shows our experiment results and we conclude in §6.

2. Phase-shifting digital holography

Phase-shifting digital holograms are recorded using the optical set-up illustrated in figure 1. The two toy car objects are located at a distance of 880 mm from the CCD camera. In figure 1, BS$_1$ and BS$_2$ are the beam splitters, M$_1$ and M$_2$ are the plane mirrors and RP$_1$ is the retardation plate. An on-axis configuration is used for recording. The recorded wavefront at the hologram plane, which is the interference between the reference beam and the object beam, is as follows (Kim & Javidi 2004):

$$H(x, y, \theta) = |O(x, y)|^2 + |R(x, y)|^2 + e^{-i\theta}O(x, y)R^*(x, y) + e^{i\theta}O^*(x, y)R(x, y),$$

(2.1)

where $O(x, y)$ and $R(x, y)$ are object and reference waves, respectively, and $\theta$ is the induced phase shift.

The synthesized hologram is obtained by using two holograms having $\lambda/4$ phase difference; the conjugate image is removed (Kim & Javidi 2004),

$$u(x, y) = O(x, y)R^*(x, y) = \frac{1}{2} \left\{ |H(x, y, 0) - |O(x, y)|^2 + |R(x, y)|^2 | + i[H(x, y, \pi/2) - |O(x, y)|^2 + |R(x, y)|^2] \right\}.$$

(2.2)
With this optical set-up, we can generate a phase-shifting hologram using only two recorded holograms instead of four holograms with four stepped phase differences as introduced by Yamaguchi & Zhang (1997). The optical set-up should be free from vibration in order to obtain precise recorded holograms.

We can reconstruct the object field at any distance from the hologram plane. The amplitude distribution at the reconstructed image plane \((x_0, y_0)\), which is located at a distance \(d\) from the hologram plane \((x, y)\), can be calculated using the following inverse Fresnel transformation (Creath 1985; Goodman 1996; Yamaguchi & Zhang 1997; Zhang & Yamaguchi 1998; Javidi & Tajahuerce 2000; Tajahuerce et al. 2001; Osten et al. 2002; Kim & Javidi 2004; Ferraro et al. 2005; Maycock et al. 2006b; Nomura et al. 2006; Stern & Javidi 2006):

\[
\begin{align*}
    u_0(x', y') &= \exp(\frac{ikd}{id\lambda}) \exp\left(\frac{ik}{2d} (x'^2 + y'^2)\right) \times F \left\{ u_i(x, y) \exp\left(\frac{ik}{2d} (x^2 + y^2)\right) \right\}, \\
    \text{(2.3)}
\end{align*}
\]

where \(\lambda\) is the wavelength of the laser beam and \(F\) is the Fourier transform.

3. The fast fixed-point ICA algorithm based on kurtosis maximization

ICA is used to extract statistically independent components from a set of observations (Bartlett et al. 1998, 2002; Hyvärinen et al. 2001; Maycock et al. 2006a). Among different approaches to estimate the independent components, such as the maximum-likelihood function, minimization of mutual information, maximization of non-Gaussianity, we chose the kurtosis maximization-based ICA algorithm. Kurtosis is a measure of the non-Gaussianity of a variable and the algorithm has a fast convergence property. Principal components analysis (PCA) helps to overcome high-dimension-related problems.

The algorithm is based on the idea of maximizing non-Gaussianity (Hyvärinen et al. 2001). The ICA components maximally deviate from Gaussian distribution. We assume that there are \(k\) observations in the data vector \(x, x = [x_1, x_2, \ldots, x_k]^T\), and \(m\) statistically independent components in \(s, s = [s_1, s_2, \ldots, s_m]^T\), where \(m \leq k\). Assuming that the data vector \(x\) is synthesized by linear combinations of independent components, then

\[
    x = As, \quad \text{(3.1)}
\]

where \(A\) is the mixing matrix. Given \(x\), the algorithm estimates the unmixing matrix, \(W\), such that

\[
    s = Wx. \quad \text{(3.2)}
\]

Kurtosis is a measure of non-Gaussianity of a random variable, which is defined as follows:

\[
    \text{kurt}(y) = E\{y^4\} - 3(E\{y^2\})^2, \quad \text{(3.3)}
\]

where \(E\) denotes the expectation operator, and kurtosis of a Gaussian random variable is zero while non-Gaussian random variables have non-zero kurtosis.

The fast fixed-point algorithm using kurtosis, which is used to find the direction so that \(y = w^Tz\) is maximally non-Gaussian, is defined as follows (Hyvärinen et al. 2001):

\[
    w \leftarrow E\{z(w^Tz)^3\} - 3w, \quad \text{(3.4)}
\]
where $T$ denotes the transpose operation; $w$ is a row of the matrix $W$; and $z$ is the whitened data vector which is obtained by

$$z = D^{-1/2} E^T x,$$

where $D = \text{diag}(d_1, \ldots, d_n)$ is the diagonal matrix of the eigenvalues of the covariance matrix $E(x x^T)$; $E = (e_1, \ldots, e_n)$ is the eigenvector matrix; and $x$ is the data vector.

When the estimates of several independent components are required, it is essential that the vectors $w_i$ are orthogonal, $i=1, \ldots, m$, where $m$ is the number of independent components to estimate. We orthogonalize the vectors $w_i$ using symmetric orthogonalization, where the vectors $w_i$ are estimated in parallel; here the Gram–Schmidt orthogonalization approach is applied (Hyvärinen et al. 2001),

$$w_1 = e_1,$$

$$w_2 = e_1 - \frac{\langle w_1 e_2 \rangle}{\langle w_1 w_1 \rangle} w_1,$$

$$w_n = e_n - \sum_{k=1}^{n-1} \frac{\langle w_k e_n \rangle}{\langle w_k w_k \rangle} w_k,$$

where $w_i$ is the $i$th row of the matrix $W$ and $\langle \cdot \rangle$ is the inner product notion.

4. Three-dimensional object recognition using computational holographic imaging and ICA algorithms

Instead of using directly reconstructed holographic images as input data for the ICA algorithm, we use the PCA coefficients (figure 2). This helps to reduce the amount of processed data by the ICA algorithm so that the computational speed can be increased and memory capacity is saved effectively. We use PCA to reduce the data dimension, which represents an $n$-dimensional vector in an $m$-dimensional domain ($m \leq n$). For matrix $X$, of which columns are $n$-dimensional vectors, the covariance matrix is

$$\sum_{xx} = E\{(x - \mu_x)(x - \mu_x)^T\},$$

where $\mu_x$ is the mean vectors of $X$ and the eigenvalue decomposition is calculated, such that

$$\sum_{xx} E = ED,$$

where $E = [e_1, \ldots, e_n]$ is the eigenvector matrix, whose columns are eigenvectors. $D = \text{diag}([d_1, \ldots, d_n])$ is the eigenvalue matrix of the covariance matrix, of which diagonal components $d_i$ are eigenvalues. The projection along the directions of eigenvectors with the highest eigenvalues retains the highest variability of the original dataset. By choosing the first $m$ eigenvectors, we can decrease the number of dimensions such that the principal component space retains most of the percentage of the variance of the dataset. Here, the first $m$ eigenvectors are chosen and normalized to be PCA transform $W_{PCA}$ ($m \leq n$) and the projected
vector $X_{\text{PCA}}$ in the PCA domain is expressed as (Yeom & Javidi 2004)

$$X_{\text{PCA}} = W_{\text{PCA}}^T X = E_m^T X.$$ (4.3)

PCA coefficients are in columns of $X_{\text{PCA}}$. The representation of $m$ reconstructed holographic images in the ICA domain is given by

$$S = W_{\text{ICA}} X_{\text{PCA}},$$ (4.4)

where $W_{\text{ICA}}$ is the estimated ICA transform. Similarly, we obtain the representation for the testing reconstructed holographic images as

$$S_{\text{test}} = W_{\text{ICA}} W_{\text{PCA}}^T X_{\text{test}},$$ (4.5)

where $X_{\text{test}}$ is the zero-mean matrix of testing reconstructed holographic images, whose column vectors are formed by the testing images (Bartlett et al. 1998, 2002; Shang et al. 2006).

Recognition performance is evaluated using the nearest neighbour algorithm, in which the cosine of the angle between the testing and training projected vectors is calculated as

$$c = \frac{s_{\text{test}} s_{\text{train}}}{\|s_{\text{test}}\| \|s_{\text{train}}\|},$$ (4.6)

where $s_{\text{test}}$, $s_{\text{train}}$ are column vectors of the matrix $S_{\text{test}}$, $S_{\text{train}}$, respectively. Similarity is obtained with a maximal cosine value. If the cosine between the testing and training projected vectors is largest or the angle between them is smallest, they are considered to be in the same class.
5. Experimental results

Experiments have been implemented using an argon laser of 514.5 nm as shown in figure 1. The CCD has a pixel size of $12 \times 12$ $\mu m$ and 2048x2048 pixels. The toy cars that are used as the experimented three-dimensional objects are of the size $25 \times 25 \times 35$ mm. Synthesized holograms of the two car objects located at a distance of 880 mm from the CCD camera are calculated from recorded holograms based on equation (2.2). Instead of reconstructing the three-dimensional volume, the image slices are reconstructed at different locations along the longitudinal direction so that every part of the object can be brought into focus in succession. We reconstructed 300 holographic images at every millimetre from 731 to 1030 mm for each class. By doing this, we can recognize multiple classes at different distances from the CCD camera. Figure 3 shows the reconstructed holographic images of the two cars at a distance of 880 mm. The size of the reconstruction window chosen is $400 \times 400$ pixels.

We use 150 reconstructed holographic images from each class for our training purposes. The rest of the data are used to test the performance. The first and the second 150 column vectors of data matrix $X$ contain the reconstructed holographic images in increasing order of odd reconstruction distance from 731 to 1029 mm for the first and the second classes, respectively. PCA technique is then applied to the matrix $X$. We choose the first 200 eigenvectors as the PCA basis. Figure 4 shows the first four eigenvectors with the highest associated eigenvalues calculated from the covariance matrix of the data matrix $X$. The eigenvalue is decreased from the first to the fourth eigenvector.

We measure the kurtoses of independent components to evaluate the performance of the fast fixed-point algorithm. Figure 5 shows the kurtosis chart of the 200 independent components estimated by maximizing the kurtoses of the $X_{PCA}$ coefficients. The higher the kurtoses are, the sparser the distribution of the independent components become.

Figure 6a–f shows the probability density functions (pdfs) of independent component numbers 20, 40, 60, 80, 100 and 120, respectively, which are located in the matrix $S$. Sixty bins are used to plot the histograms. The pdfs are very...
Figure 4. First four eigenvectors with the highest associated eigenvalues calculated from the covariance matrix: eigenvectors (a) 1, (b) 2, (c) 3 and (d) 4.

Figure 5. Kurtoses of the independent components.
sparse with extremely sharp and narrow peaks at zero. Most of the possible values are about the mean value of zero, so there are only a small number of significant coefficients encoding most of the information about the signal. This is the benefit of using ICA. Compared with ICA, the pdfs of the corresponding PCA components shown in figure 7 are not sparse.

In figure 8, the first four basis images in the columns of matrix $A$ are shown, visualized using $W_{PCA}$.

Figure 9 illustrates the orthogonal property of the transform matrix $W$, which is implemented using the Gram–Schmidt approach. Sines between row vectors 40, 80, 120, 160, 200 and all row vectors in $W$ are plotted using different grey scales. At the row positions of 40, 80, 120, 160 and 200, the sines between the row vectors and themselves are zeros, while the sines between them and other row vectors are around the value of 1 owing to the orthogonality.

We test the method using the training reconstructed holographic images as input (figure 10). $X_{test}$ is made similar to $X$ with 300 reconstructed holographic images in the form of column vectors. $S_{test}$ is similar to $S$, whose columns are representations of corresponding holographic images in $X$. The first class of reconstructed holographic images numbered 50, 100, 150 and the second class ones numbered 200, 250, 300, which correspond to the reconstruction distances of 829, 929 and 1029 mm for each class, respectively, are chosen to plot the result. We observe that the cosines reach the value of 1 at the same positions as the

Figure 6. Pdfs of independent component numbers (a) 20, (b) 40, (c) 60, (d) 80, (e) 100 and (f) 120.
column vectors in $S$, where we have multiplications of the same two column vectors, whereas they are very small and approximately the value of zero at other positions. The vertical grey line is the border between class 1 and class 2.

In the next test, for each class, we use another set of 150 testing images, which are not used in the training step. The first and the second 150 column vectors of data matrix $X_{test}$ contain the reconstructed holographic images in increasing order of even reconstruction distance from 732 to 1030 mm for the first and the second classes, respectively. Figure 11 shows the test result. We choose to plot the results for the reconstructed holographic images numbered 50, 100, 150 from class 1 in the testing set $X_{test}$, which correspond to the reconstruction distances of 830, 930, 1030 mm, and reconstructed holographic images numbered 200, 250, 300 from class 2, which correspond to the same distances as above. After applying $W_{PCA}$ and $W_{ICA}$, the column vectors in $S_{test}$ are the representations of these reconstructed test holographic images. We observe that, for the first class, the cosines reach peak values at the column vectors numbered 50, 100 and 150 in $S$; or, in other words, the test holographic images match with the reference holographic images reconstructed at distances of 829, 929 and 1029 mm. For the second class, the cosines reach the maximum at the column vectors 201, 251 and 300 in $S$; or, we can say that the test holographic images match with the reference holographic images reconstructed at distances of 831, 931 and

\[ (a), (b), (c) \]

\[ (d), (e), (f) \]

Figure 7. Pdfs of PCA component numbers (a) 20, (b) 40, (c) 60, (d) 80, (e) 100 and (f) 120.
1029 mm. The result cosine shows a peak when the testing and the training reconstructed holographic images are closest to each other in terms of reconstruction distance. It decreases dramatically for other distances excluding the closest. Considering that the adjacently reconstructed holographic images are highly correlated, we see that the system is very discriminant.

6. Conclusion

ICA has been known as a powerful algorithm to acquire statistically independent components from mixture data. It has been applied in signal processing and two-dimensional image processing. However, to the best of our knowledge, ICA has not been explored in a three-dimensional environment, and this paper is the first to report using ICA and three-dimensional imaging technology together. In this paper, we have presented three-dimensional object sensing by computational holographic imaging and recognition using kurtosis maximization-based ICA. Phase-shifting holography is applied using two phases to sense the three-dimensional scene. Holograms of two car objects are recorded for classification. Three hundred holographic images are reconstructed at different distances from the sensor for each class, with one-half of them being
Figure 9. Sines between basis vectors in the rows of the transform matrix $W$ (rows 40, 60, 80, 120, 160 and 200).

Figure 10. Cosines of the testing and training projected vectors; training data are used (plotted are the results with reconstructed holographic images numbered (a) 50, (b) 100, (c) 150, (d) 200, (e) 250 and (f) 300).
used for training and the other half for testing purposes. We use PCA technique in order to reduce the data dimension, so that the processing speed can be increased and the memory can be saved effectively without affecting the recognition performance. Kurtosis maximization-based ICA algorithm is used to project holographic images to the ICA domain. Performance evaluation is based on the cosines between the testing and training projected vectors as a similarity measure. The PCA–ICA combination appears to work well in three-dimensional recognition of object classes using reconstructed holographic images at different longitudinal depths.

References


Sadjadi, F. (ed.) 2000 Selected papers on automatic target recognition, SPIE-CDROM.


C. M. Do and B. Javidi


