Experimental verifications of a joint transform correlator using compressed reference images

Joewono Widjaja and Ubon Suripon

Single-target and multiple-target detections by using a joint transform correlator (JTC) with compressed reference images are experimentally verified. Two high-contrast images with different spatial-frequency content are used as test scenes. Although an effect of the additive noise on detection performance of the proposed correlator is more severe than that of the compression, the experimental results confirm the feasibility of implementing the JTC with compressed reference images. © 2006 Optical Society of America

OCIS codes: 070.4550, 070.5010.

1. Introduction

With the development of spatial light modulators (SLMs), a joint transform correlator (JTC) has been found advantageous for target detection systems[1–5] because the synthesis of complex matched filters in advance is eliminated and reference images can be updated in real time. In a real-time JTC, the correlation between an input target detected by an image sensor and a reference image stored in a computer system is accomplished by displaying the two images side-by-side onto a SLM. By capturing its optical Fourier spectrum with a CCD sensor, a joint power spectrum (JPS) of the images is recorded. The captured JPS is then redisplayed on the same SLM in order to be optically Fourier transformed. This yields a correlation output. For additional discussions of various aspects of the JTC, the reader may wish to consult Ref. 6.

However, the use of an electrically addressed SLM (EASLM) in the JTC has several drawbacks. For one example, the process of displaying the target and the reference images onto the EASLM causes a bottleneck because of the serial nature of the signal communication between the computer and the SLM. This time delay is dependent on the file size of the images. Owing to the inherent rotation and scale variances of the JTC, the stored reference images must contain all possible variations of the rotation and scale changes of the target. Thus the application of a JTC to automatic target recognitions requires considerable storage.

To obviate this problem, JPEG compression[7] of the reference images has been proposed and theoretically studied by Widjaja and Suripon.[8,9] In their research the effects of the JPEG compression of reference images on single- and multiple-target detection by using a JTC are quantitatively studied by using computer simulation. The studies are conducted using two types of image with different spatial-frequency content in a situation in which overlapping noise is present in the input and there is a contrast difference between the target and the reference images. In contrast with the use of the compressed reference with high spatial-frequency content, the results of our studies show that the JTC using a compressed reference image with low spatial-frequency content offers a better recognition performance in that it is robust to noise and contrast difference for a wide range of compression levels.

In this work the optical implementation of single- and multiple-target detection using a JTC with JPEG compressed reference images is experimentally verified. For the experiments, the same set of images used in our previous works is employed as test scenes. The performance of the single-target detection is quantified by using a ratio of the correlation peak intensity to the standard deviation of the correlation output or peak-to-correlation deviation (PCD), while that of the multiple-target detection is a ratio of the primary to the secondary peaks or peak-to-

The authors are with the Institute of Science, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand. J. Widjaja's e-mail address is widjaja@sut.ac.th.

Received 6 April 2006; revised 5 June 2006; accepted 16 June 2006; posted 19 June 2006 (Doc. ID 69596).

0003-6935/06/318074-09$15.00/0

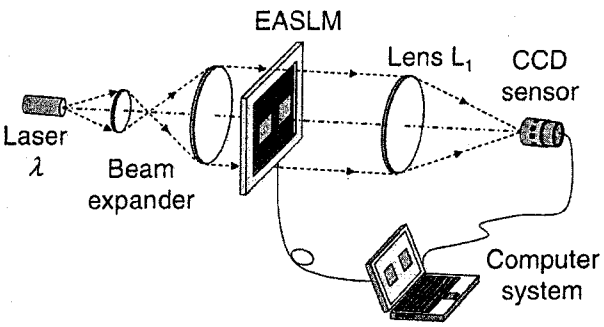

Fig. 1. Schematic of an optical setup for performing experimental verifications of the JTC with compressed reference images.

secondary ratio (PSR). Section 2 introduces the theoretical background of the optical implementation of the JTC using compressed reference images and reviews the JPEG compression algorithm. Experimental results are discussed in Section 3. Our conclusions are discussed in Section 4.

2. Theoretical Background

A. Joint Transform Correlator with Compressed Reference Images

Figure 1 shows a schematic of an optical setup for performing experimental verifications. For the sake of simplicity of the discussion, single-target detection is considered. The input target image $t(x, y)$ captured by the CCD image sensor and the compressed reference image $r_c(x, y)$ stored in a computer system are assumed to have the same size of $N_x \times N_y$ pixels. Let us consider the EASLM placed at the front focal plane of lens L_1, comprised of a matrix of rectangular light-modulating elements with a resolution of $M_x \times M_y$ pixels, a pixel size of $L_x \times L_y$, and a pixel pitch P_x and P_x in the x and y directions, respectively. The amplitude transmittance of the EASLM can be written as

$$t_{\mathrm{EASLM}}(x, y) = \sum_{n=-M_y}^{+M_y} \sum_{m=-M_x}^{+M_x} \delta(x - mP_x, y - nP_y) \otimes \mathrm{rect}\left(\frac{x}{L_x}\right)\mathrm{rect}\left(\frac{y}{L_y}\right), \quad (1)$$

where $\otimes$ denotes the convolution operation. If the horizontal separation between the target and the compressed reference images displayed on the EASLM is MP_x, the joint input image can be expressed as

$$f_{\mathrm{input}}(x, y) = t\left(x + \frac{MP_x}{2}, y\right)\mathrm{rect}\left(\frac{x + MP_x/2}{N_xP_x}\right) \times \mathrm{rect}\left(\frac{y}{N_yP_y}\right) + r_c\left(x - \frac{MP_x}{2}, y\right) \times \mathrm{rect}\left(\frac{x - MP_x/2}{N_xP_x}\right)\mathrm{rect}\left(\frac{y}{N_yP_y}\right). \quad (2)$$

The amplitude transmittance of the joint input image is given by

$$f_{\mathrm{JTC}}(x, y) = t_{\mathrm{EASLM}}(x, y) f_{\mathrm{input}}(x, y) = \sum_{n=-M_y}^{+M_y} \sum_{m=-M_x}^{+M_x} \delta(x - mP_x, y - nP_y) \otimes \mathrm{rect}\left(\frac{x}{L_x}\right)\mathrm{rect}\left(\frac{y}{L_y}\right)\left\{t\left(x + \frac{MP_x}{2}, y\right) \times \mathrm{rect}\left(\frac{x + MP_x/2}{N_xP_x}\right)\mathrm{rect}\left(\frac{y}{N_yP_y}\right) + r_c\left(x - \frac{MP_x}{2}, y\right)\mathrm{rect}\left(\frac{x - MP_x/2}{N_xP_x}\right) \times \mathrm{rect}\left(\frac{y}{N_yP_y}\right)\right\}. \quad (3)$$

The joint Fourier spectrum at the back focal plane of lens L_1,

$$F_{\mathrm{JTC}}(x', y') = \frac{1}{\lambda f}\frac{L_xL_y}{P_xP_y}\left[\sum_{n=-M_y}^{M_y} \sum_{m=-M_x}^{M_x} \times \delta\left(\frac{x'}{\lambda f} - \frac{m}{P_x}, \frac{y'}{\lambda f} - \frac{n}{P_y}\right)\mathrm{sinc}\left(\frac{L_xx'}{\lambda f}\right) \times \mathrm{sinc}\left(\frac{L_yy'}{\lambda f}\right) \otimes F_{\mathrm{input}}(x', y')\right], \quad (4)$$

is recorded by the CCD sensor and is then transferred to the computer. Here, λ and f are the wavelength of coherent light and the focal length of lens L_1, respectively. The Fourier spectrum of the joint input image is equal to

$$F_{\mathrm{input}}(x', y') = T'\left(\frac{x'}{\lambda f}, \frac{y'}{\lambda f}\right)\exp\left(j2\pi\frac{MP_xx'}{2\lambda f}\right) + R_c'\left(\frac{x'}{\lambda f}, \frac{y'}{\lambda f}\right)\exp\left(-j2\pi\frac{MP_xx'}{2\lambda f}\right), \quad (5)$$

where

$$T'\left(\frac{x'}{\lambda f}, \frac{y'}{\lambda f}\right) = T\left(\frac{x'}{\lambda f}, \frac{y'}{\lambda f}\right) \otimes N_xN_yP_xP_y\,\mathrm{sinc}\left(\frac{N_xP_xx'}{\lambda f}\right) \times \mathrm{sinc}\left(\frac{N_yP_yy'}{\lambda f}\right), \quad (6a)$$

$$R_c'\left(\frac{x'}{\lambda f}, \frac{y'}{\lambda f}\right) = R_c\left(\frac{x'}{\lambda f}, \frac{y'}{\lambda f}\right) \otimes N_xN_yP_xP_y\,\mathrm{sinc}\left(\frac{N_xP_xx'}{\lambda f}\right) \times \mathrm{sinc}\left(\frac{N_yP_yy'}{\lambda f}\right). \quad (6b)$$

The recorded JPS of the images can be expressed as

$$
\begin{aligned}
|F_{\mathrm{JTC}}(x',\,y')|^2 = \frac{1}{\lambda^2 f^2}\left(\frac{L_x L_y}{P_x P_y}\right)^2 \sum_{n=-M_y}^{+M_y}\sum_{m=-M_x}^{+M_x} \mathrm{sinc}^2\left(\frac{mL_x}{P_x}\right) \\
\times \mathrm{sinc}^2\left(\frac{nL_y}{P_y}\right)\left\{\left|R_c'\left(\frac{x'}{\lambda f}-\frac{m}{P_x},\,\frac{y'}{\lambda f}-\frac{n}{P_y}\right)\right|^2\right. \\
+\left|T'\left(\frac{x'}{\lambda f}-\frac{m}{P_x},\,\frac{y'}{\lambda f}-\frac{n}{P_y}\right)\right|^2 \\
+R_c'\left(\frac{x'}{\lambda f}-\frac{m}{P_x},\,\frac{y'}{\lambda f}-\frac{n}{P_y}\right) \\
\times T'\left(\frac{x'}{\lambda f}-\frac{m}{P_x},\,\frac{y'}{\lambda f}-\frac{n}{P_y}\right) \\
\times \exp\left[-j2\pi\left(\frac{x'}{\lambda f}-\frac{m}{P_x}\right)\right] \\
+R_c'\left(\frac{x'}{\lambda f}-\frac{m}{P_x},\,\frac{y'}{\lambda f}-\frac{n}{P_y}\right) \\
\times T'^{*}\left(\frac{x'}{\lambda f}-\frac{m}{P_x},\,\frac{y'}{\lambda f}-\frac{n}{P_y}\right) \\
\left.\times \exp\left[j2\pi\left(\frac{x'}{\lambda f}-\frac{m}{P_x}\right)\right]\right\}. \qquad (7)
\end{aligned}
$$

Equation (7) shows that the pixel structure of the EASLM produces the multiple power spectrum of the joint input image at intervals of $\lambda f/P_x$ and $\lambda f/P_y$ in the horizontal and vertical directions, respectively. Since they are modulated by a broad 2D sinc function, only the zero order ($m = 0$, $n = 0$) of the JPS has the highest intensity, while the higher orders are attenuated. For this reason, the recorded zero-order spectrum will be used for the next computation. By expressing the complex field distribution $R_c'(x', y')$ and $T'(x', y')$ into their amplitude and phase distributions

$$
R_c'\left(\frac{x'}{\lambda f},\,\frac{y'}{\lambda f}\right) = \left|R_c\left(\frac{x'}{\lambda f},\,\frac{y'}{\lambda f}\right)\right| \exp\left[j\phi_R\left(\frac{x'}{\lambda f},\,\frac{y'}{\lambda f}\right)\right],
$$

$$
T'\left(\frac{x'}{\lambda f},\,\frac{y'}{\lambda f}\right) = \left|T'\left(\frac{x'}{\lambda f},\,\frac{y'}{\lambda f}\right)\right| \exp\left[j\phi_T\left(\frac{x'}{\lambda f},\,\frac{y'}{\lambda f}\right)\right],
$$

the zeroth-order JPS can be expressed as

$$
\begin{aligned}
|F_{\mathrm{JTC}_0}(x',\,y')|^2 = \left|R_c'\left(\frac{x'}{\lambda f},\,\frac{y'}{\lambda f}\right)\right|^2 + \left|T'\left(\frac{x'}{\lambda f},\,\frac{y'}{\lambda f}\right)\right|^2 \\
+2\left|R_c'\left(\frac{x'}{\lambda f},\,\frac{y'}{\lambda f}\right)\right|\left|T'\left(\frac{x'}{\lambda f},\,\frac{y'}{\lambda f}\right)\right| \\
\times \cos\left[\frac{2\pi M P_x x'}{\lambda f}+\phi_R\left(\frac{x'}{\lambda f},\,\frac{y'}{\lambda f}\right)\right. \\
\left.-\phi_T\left(\frac{x'}{\lambda f},\,\frac{y'}{\lambda f}\right)\right]. \qquad (8)
\end{aligned}
$$

The third term in Eq. (8) is in particular interest because it provides the desired correlation output. This term contains the desired product of the Fourier spectra of the target and the compressed reference images that are sampled by the sinusoidal fringes. Since the CCD sensor inherently has finite resolution, the sensor must record the fringes faithfully. If the CCD sensor used has a spatial resolution of $2M_{\mathrm{CCD}x} \times 2M_{\mathrm{CCD}y}$ with pitches of $P_{\mathrm{CCD}x}$ and $P_{\mathrm{CCD}y}$, the digitized zeroth-order JPS can be mathematically written as

$$
\begin{aligned}
|F_{\mathrm{JTC}_0}(x',\,y')|^2 = \left\{\left|R_c'\left(\frac{x'}{\lambda f},\,\frac{y'}{\lambda f}\right)\right|^2 + \left|T'\left(\frac{x'}{\lambda f},\,\frac{y'}{\lambda f}\right)\right|^2\right. \\
+2\left|R_c'\left(\frac{x'}{\lambda f},\,\frac{y'}{\lambda f}\right)\right|\left|T'\left(\frac{x'}{\lambda f},\,\frac{y'}{\lambda f}\right)\right| \\
\times \cos\left[\frac{2\pi M P_x x'}{\lambda f}+\phi_R\left(\frac{x'}{\lambda f},\,\frac{y'}{\lambda f}\right)\right. \\
\left.\left.-\phi_T\left(\frac{x'}{\lambda f},\,\frac{y'}{\lambda f}\right)\right]\right\} \sum_{n=-M_{\mathrm{CCD}y}}^{+M_{\mathrm{CCD}y}}\sum_{m=-M_{\mathrm{CCD}x}}^{+M_{\mathrm{CCD}x}} \\
\times \delta(x'-mP_{\mathrm{CCD}x},\,y'-nP_{\mathrm{CCD}y}) \\
\otimes \mathrm{rect}\left(\frac{x'}{L_{\mathrm{CCD}x}}\right)\mathrm{rect}\left(\frac{y'}{L_{\mathrm{CCD}y}}\right), \qquad (9)
\end{aligned}
$$

where $L_{\mathrm{CCD}x} \times L_{\mathrm{CCD}y}$ represents the size of the light-detecting rectangular element of the CCD. When the active area of the sensor is greater than the size of the zeroth-order spectrum, a faithful recording is achieved provided that the sampling theorem[10] is satisfied. Therefore the sampling frequency of the sensor must be at least two times higher than the frequency of sinusoidal fringes $f_{\mathrm{CCD}} \geq 2 f_{\mathrm{Fringes}}$ or

$$
\frac{\lambda f}{2MP_x} \geq P_{\mathrm{CCD}x}. \qquad (10)
$$

Inequality (10) indicates that the spatial separation of the target and the compressed reference images in the input plane are determined by the focal length of the Fourier transform lens, the operating wavelength of the coherent light, and also the spatial resolution of the CCD sensor.

Next, the digitized JPS is redisplayed on the EASLM in order to produce the correlation output. Since the pixel pitch of the CCD sensor is different from that of the EASLM, the spatial extension of the redisplayed joint power spectrum is scaled by the ratio of the pixel pitch given by

$$
r_{px} = \frac{P_{\mathrm{CCD}x}}{P_x}, \qquad r_{py} = \frac{P_{\mathrm{CCD}y}}{P_y} \qquad (11)
$$

in the horizontal and the vertical directions, respectively. By taking this ratio into account, the am-

plitude transmittance of the redisplayed JPS is proportionally equal to

$$|F_{\mathrm{JTC}_0}(x', y')|^2 = \left\{\left|R_c'\left(\frac{r_{px}x'}{\lambda f}, \frac{r_{py}y'}{\lambda f}\right)\right|^2 + \left|T'\left(\frac{r_{px}x'}{\lambda f}, \frac{r_{py}y'}{\lambda f}\right)\right|^2 + 2\left|R_c'\left(\frac{r_{px}x'}{\lambda f}, \frac{r_{py}y'}{\lambda f}\right)\right|\left|T'\left(\frac{r_{px}x'}{\lambda f}, \frac{r_{py}y'}{\lambda f}\right)\right| \times \cos\left[\frac{2\pi r_{px}MP_x x'}{\lambda f} + \phi_R\left(\frac{r_{px}x'}{\lambda f}, \frac{r_{py}y'}{\lambda f}\right) - \phi_T\left(\frac{r_{px}x'}{\lambda f}, \frac{r_{py}y'}{\lambda f}\right)\right]\right\} \sum_{n=-M_{\mathrm{CCDy}}}^{+M_{\mathrm{CCDy}}} \sum_{m=-M_{\mathrm{CCDx}}}^{+M_{\mathrm{CCDx}}} \times \delta(r_{px}x' - mP_{\mathrm{CCDx}}, r_{py}y' - nP_{\mathrm{CCDy}}) \otimes \mathrm{rect}\left(\frac{x'}{L_x}\right)\mathrm{rect}\left(\frac{y'}{L_y}\right). \quad (12)$$

The subsequent optical Fourier transform of the redisplayed JPS using lens L_1 produces the correlation output intensity of the form

$$I_c(x, y) \propto \left|c(x, y) \otimes \left[\frac{L_xL_y}{P_xP_y}\mathrm{sinc}\left(\frac{L_xx}{\lambda f}\right)\mathrm{sinc}\left(\frac{L_xy}{\lambda f}\right) \times \sum_{n=-M_y}^{M_y}\sum_{m=-M_x}^{m=-M_x}\delta\left(x - \frac{m\lambda f}{P_x}, y - \frac{m\lambda f}{P_y}\right)\right]\right|^2, \quad (13)$$

where $c(x, y)$ is the desired correlation signal given by

$$c(x, y) = r_c'\left(\frac{x}{r_{px}}, \frac{y}{r_{py}}\right) \otimes r_c'\left(\frac{-x}{r_{px}}, \frac{-y}{r_{py}}\right) + t'\left(\frac{x}{r_{px}}, \frac{y}{r_{py}}\right) \otimes t'\left(\frac{-x}{r_{px}}, \frac{-y}{r_{py}}\right) + r_c'\left(\frac{x}{r_{px}}, \frac{y}{r_{py}}\right) \otimes t'\left(\frac{-x}{r_{px}}, \frac{-y}{r_{py}}\right) \otimes \delta(x - MP_{\mathrm{CCDx}}, y) + r_c'\left(\frac{-x}{r_{px}}, \frac{-y}{r_{py}}\right) \otimes t'\left(\frac{x}{r_{px}}, \frac{y}{r_{py}}\right) \otimes \delta(x + MP_{\mathrm{CCDx}}, y), \quad (14)$$

with

$$r_c'\left(\frac{x}{r_{px}}, \frac{y}{r_{py}}\right) = r_c\left(\frac{x}{r_{px}}, \frac{y}{r_{py}}\right)\mathrm{rect}\left(\frac{x}{N_xP_{\mathrm{CCDx}}}\right) \times \mathrm{rect}\left(\frac{y}{N_yP_{\mathrm{CCDy}}}\right),$$

$$t'\left(\frac{-x}{r_{px}}, \frac{-y}{r_{py}}\right) = t\left(\frac{x}{r_{px}}, \frac{y}{r_{py}}\right)\mathrm{rect}\left(\frac{x}{N_xP_{\mathrm{CCDx}}}\right) \times \mathrm{rect}\left(\frac{y}{N_yP_{\mathrm{CCDy}}}\right). \quad (15)$$

Expression (13) shows that, owing to the pixel structure of the EASLM, the correlation output $c(x, y)$ of Eq. (14) is periodically replicated with the separations $\lambda f/P_x$ and $\lambda f/P_y$ in the horizontal and the vertical directions, respectively. In Eq. (14), the first two terms correspond to the autocorrelations of the compressed reference and of the target images that have appeared at the same position, while the last two terms are the desired cross correlation signals separated from the first two terms by distance $\pm MP_{\mathrm{CCDx}}$. The maximum width of each correlation signal is confined in the area $2N_xP_{\mathrm{CCDx}} \times 2N_yP_{\mathrm{CCDy}}$. The cross correlation and the autocorrelation signals will not overlap provided $M \geq 2N_x$. Figure 2 illustrates the cross-sectional scan of the replicated correlation output along $y = 0$, where $c_0(x, y)$ and $c_1(x, y)$ stand for the zeroth-order and the first-order correlation signals, respectively. For the sake of clarity in the figure, the replicated correlation signals are drawn by not taking the modulating sinc functions into account. Each replicated correlation signal consists of three patterns; the center pattern is the strongest because it corresponds to the first two terms of Eq. (14), and the other two patterns represent the last two terms.

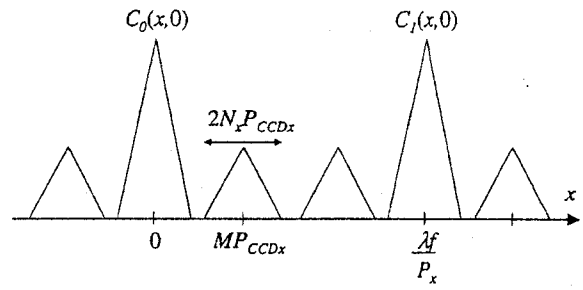

Fig. 2. Locations of the zero-order and the first-order correlation signals.

To prevent overlapping of the adjacent replicas, the spatial separation between replicas must be chosen sufficiently wide. It can be seen from Fig. 2 that the separation is determined by the following condition:

$$\frac{\lambda f}{P_x} \geq 2MP_{\mathrm{CCDx}} + 2N_xP_{\mathrm{CCDx}}, \quad (16)$$

which is dependent on the characteristic of the EASLM, the CCD sensor, the focal length of the Fourier transform lens, and the wavelength of the laser light. When $M = 2N_x$, the focal length of lens L_1 must satisfy

$$f \geq \frac{6N_xP_{\mathrm{CCDx}}P_x}{\lambda}. \quad (17)$$

B. JPEG Compression

JPEG compression is one of the digital compression standards designed for still images.[7] To compress still images, the JPEG algorithm divides picture elements of the input image into 8×8 pixel blocks. The spatial frequency of each block is independently calculated by using the 2D discrete cosine transform (DCT). The DCT generates 8×8 spatial frequencies consisting of 1 dc and 63 ac frequency coefficients. Since as a general rule the pixel values of the image vary slowly, the value of the dc coefficient is larger than that of the ac

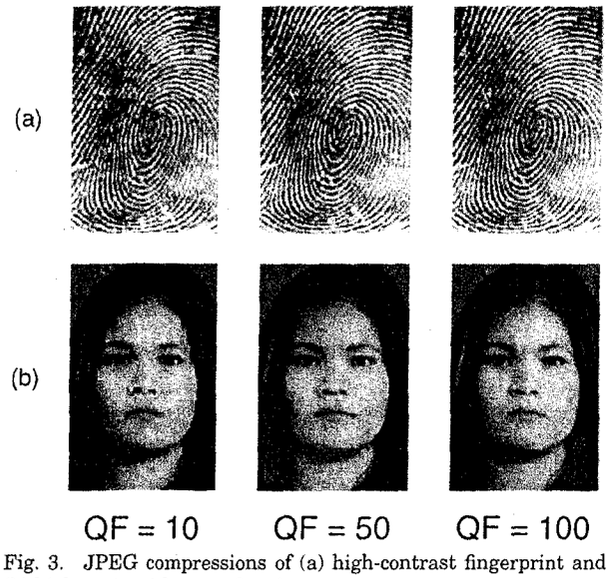

Fig. 3. JPEG compressions of (a) high-contrast fingerprint and (b) high-contrast human face images.

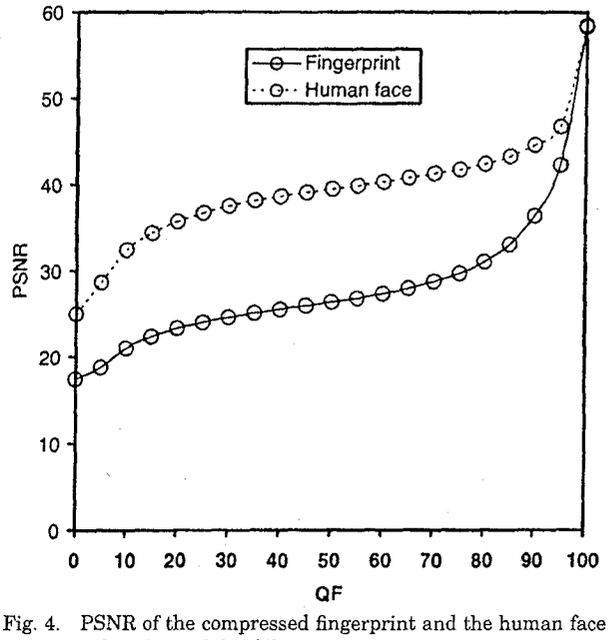

Fig. 4. PSNR of the compressed fingerprint and the human face images as functions of the QF.

coefficient. Next, the resultant DCT of each block is quantized by dividing each value of the DCT coefficient by a corresponding number from a quantization table. To discard high-spatial-frequency content, the quantization values for the ac coefficients are set to be larger than those of the dc coefficients. Each quantized DCT coefficient is then rounded to the nearest integer. As a result, most ac coefficients become zero, which causes irretrievable loss. To obtain higher compression, the dc coefficient is then processed by storing the difference between the dc coefficients of consecutive blocks. As for the ac coefficients, which consist of zero values, further compression is achieved by using run length encoding and Huffman coding.

3. Experimental Verifications

To perform experimental verifications, high-contrast fingerprint and human face images were prepared and duplicated as the target and reference images. Unlike in our previous work,[8,9] a low-contrast image was not employed, because the limited contrast ratio of the EASLM used prevented an efficient display of images with small variations of gray-level pixels. The images consisted of 124 × 186 pixels with 8-bit gray scale levels, and their size was 23 kbyte. The reference images were compressed into JPEG format by using the ACDsee software (ACD Systems, Ltd.) with a different quality factor (QF) whose value could be varied from 100 to 0. A high value QF discards less information than one with a small value. Thus the higher the value of the QF, the better the image quality and the bigger the file size of the compressed image. Figures 3(a) and 3(b) show the compressed fingerprint and human face images with QF = 10, 50, and 100, respectively. It can be seen from Fig. 3(b) that as the QF is lowered the quantization applied independently on 8 × 8 pixel blocks generates more visible gray scale discontinuities along the block boundaries called the blocking artifacts.[7] This is because the human face image contains more pixels with slowly varying values compared to the fingerprint image. Thus its blocking artifacts appear on a larger area.

The quality of the compressed image was objectively measured by using a peak signal-to-noise ratio[11] (PSNR) given by

$$\mathrm{PSNR} = 10 \log_{10}\left\{\frac{255^2}{\frac{1}{N_x \times N_y}\sum_{j=1}^{N_y}\sum_{i=1}^{N_x}[f(i,j) - f_c(i,j)]^2}\right\}. \tag{18}$$

Here $f(i, j)$ and $f_c(i, j)$ are the original and the compressed images with dimensions of $N_x \times N_y$ pixels, respectively. $f(i, j)$ takes integer values between 0 and 255 for an 8-bit gray scale image. The large PSNR means that the degree of similarity between the original and the compressed images is high. Figure 4 shows the PSNR of each compressed reference image as a function of the QF. It can be seen from the figure that the PSNRs become higher as the QF increases. However, the PSNR of a compressed human face is higher than that of the fingerprint. This is because the fingerprint image contains more high spatial-frequency components than the human face image does. Since the high spatial-frequency determines the fine details of the image, discarding high spatial-frequency components of the fingerprint degrades the fine detail information more than that of the human face image. Therefore owing to severe degradation the PSNR of the compressed high-contrast fingerprint is lower.

In the experiments the single-target and the multiple-target detections were performed by combining the target and the compressed reference images as the joint input image $f_{input}(x, y)$ with the horizontal separation of $M = 248$ pixels. In the case of multiple-target detection, the input scene contained two different images with one of them identical to the original reference image. Since as a practical matter the input targets captured by CCD sensors are far from ideal, an additive Gaussian noise generated by using the IMNOISE command of MATLAB 6.0 was added to the target images. The joint input image was then displayed on a Jenoptik SLM-M/460 EASLM with a resolution of 832 × 624 pixels with a pixel size of 27 μm × 23 μm and a pixel pitch of 32 μm × 32 μm. By perpendicularly illuminating the EASLM with collimated coherent light generated from a He–Ne laser source operating at a wavelength of 632.8 nm, the JPS was generated at the back focal plane of lens L_1. An 8-bit Pulnix TM-2016-8 CCD sensor with a pixel resolution of 1920 × 1080 and a pixel size and pitch of 7.4 μm × 7.4 μm was used to capture the generated JPS and the correlation output intensity. To reduce the excessive light intensities of the JPS and the correlation output, a neutral density (ND) filter was installed in front of the sensor. However, because the amplitude and the bandwidth of the Fourier components of the human face image are smaller than those of the fingerprint,[8] the sinusoidal fringes of the JPS generated during the human face detection is weaker. Therefore to faithfully record the fringes, the ND filter was not employed during the recording of the JPS of the human face. As for Fourier transforming lens L_1, a lens with the focal length of $f = 300$ mm was used. This is because substitutions of the related parameters into Eq. (15) give the focal length of $f \geq 278$ mm.

A. Single-Target Detections

1. Compressed High-Contrast Fingerprint as Reference

Figures 5(a), 5(b), and 5(c) show the experimental output correlations of the JTC. The autocorrelation peak of the uncompressed fingerprint shown in Fig. 5(a) is almost two times higher than the cross correlation generated by using the compressed fingerprint with a QF of 10 shown in Fig. 5(b). As reported in our previous work,[8] this is caused by the degradation of the reference image by compression. Figure 5(c) shows the output of the correlation of the noisy fingerprint target with a variance of $\sigma^2 = 1$ and the compressed fingerprint reference with a QF of 10. It can be seen from the figure that the correlation peak disappears. This is because, besides the additive noise, the sinusoidal fringes are corrupted by speckle noise that is inherently generated by light scattering and multiple internal reflections from the EASLM and the surface of the sensor. Therefore the correlation peak cannot be detected.

Figure 6 shows the normalized PCDs as a function of the QF of the compressed fingerprint for different

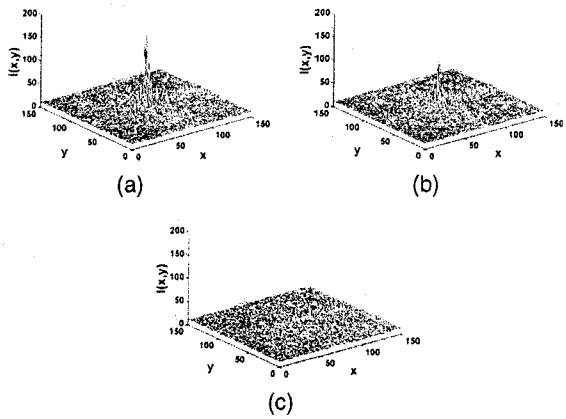

Fig. 5. Experimental results of the single-fingerprint detection using the proposed JTC. (a) Autocorrelation output of the uncompressed fingerprint and cross correlation outputs using the compressed fingerprint reference (QF = 10) in a situation in which the target is (b) noise-free and (c) noisy ($\sigma^2 = 1$).

target scenes. It is clear that, when the fingerprint target is free from noise or is corrupted by noise with variance $\sigma^2 = 0.01$, the PCDs increase gradually as the QF increases. However, when the fingerprint target is corrupted by strong noise, the PCD reduces drastically regardless of the compression quality. In comparison with the simulation, the detection of the single-fingerprint target using the proposed JTC is more dependent on noise level than on compression.

2. Compressed High-Contrast Human Face as Reference

The experimental output correlations generated by the proposed JTC are shown in Figs. 7(a)–7(c). The output correlations are not as broad as those obtained by the computer simulations,[8] because the correla-

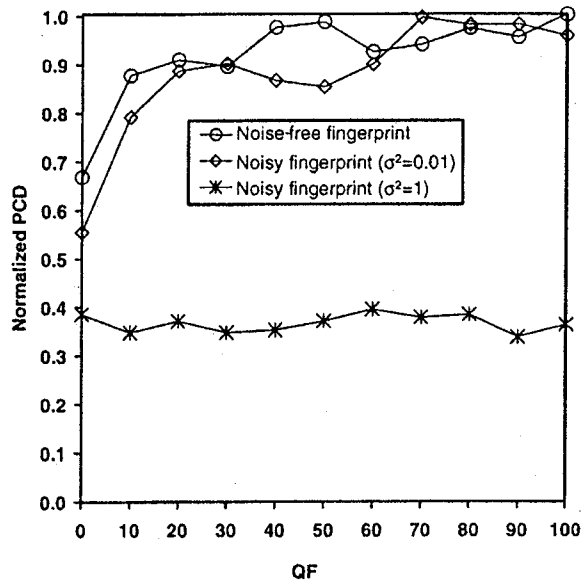

Fig. 6. PCD of single-fingerprint detection as a function of the QF of the compressed fingerprint reference.

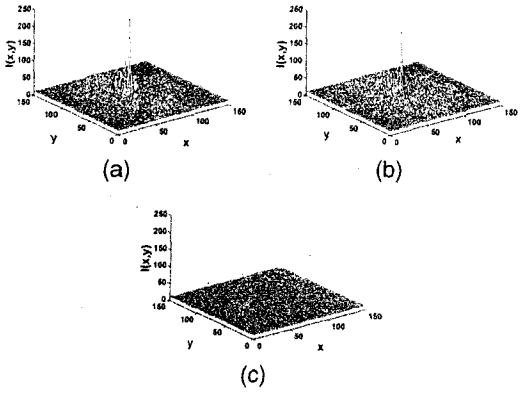

Fig. 7. Experimental results of the single-human-face detection using the proposed JTC. (a) Autocorrelation output of the uncompressed human face and cross correlation outputs using the compressed human face reference (QF = 10) in a situation in which the target is (b) noise-free and (c) noisy ($\sigma^2 = 1$).

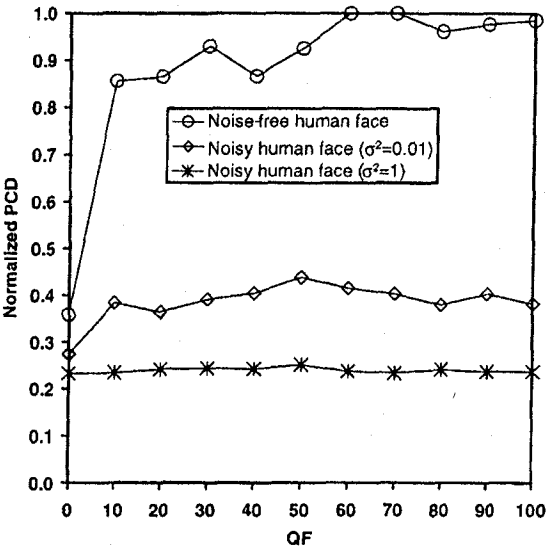

Fig. 8. PCD of a single human face detection as a function of the QF of the compressed human face reference.

tion peak intensities are reduced by the ND filter. As a consequence, the peak height and width of the correlation signal reduce greatly. When the noise-free human face target is detected by the reference image compressed at a QF of 10, the correlation output shown in Fig. 7(b) decreases slightly. However, its peak is almost as sharp as the autocorrelation output shown in Fig. 7(a). In comparison with the outputs of the fingerprint detections shown in Fig. 5, the correlation planes of Figs. 7(a) and 7(b) appear to be noisier. This is because the recording of the JPS without the ND filter causes the low-frequency components of the generated JPS to become clipped; however, more fringes in the higher frequencies are recorded. As the clipped JPS is subsequently redisplayed onto the EASLM, not only is more light diffracted but also more speckle noise is generated to the correlation output. Therefore although the correlation peak height is high, the speckle noise increases. Moreover, as the blocking artifact increases, more low-frequency components of the JPS are clipped. This in turn further increases the speckle noise. The correlation output of the noise-corrupted human face target with the variance $\sigma^2 = 1$ is depicted in Fig. 7(c). For the same reason discussed previously, the correlation peak cannot be observed because it is buried in the noise.

Figure 8 shows the normalized PCDs as a function of the QF of a compressed human face for different target scenes. When the target is a noise-free human face, the PCD decreases gradually as the QF becomes smaller. However, when the QF is less than 10, the PCD reduces rapidly. As mentioned in the discussion of Fig. 5(c), owing to the increase of the blocking artifact the speckle noise in the correlation plane increases. In the presence of additive Gaussian noise with variance $\sigma^2 = 0.01$ in the input target, the PCD drops below 0.5 because the fringe structure has been corrupted by the speckle noise. As the additive noise exists, the correlation peak degenerates. When the additive noise becomes stronger, the correlation peak reduces rapidly. Subsequently, the PCD further decreases. Therefore owing to the presence of the speckle noise, the detection performance of the proposed JTC using a compressed human face depends on both the compression and the additive noise. Note that several methods for enhancing the fringes of the JPS by dc blocking, binarizing, or manipulating the fringes have been reported.[12–14]

B. Multiple-Target Detections

1. *Compressed High-Contrast Fingerprint as Reference*

Figures 9(a)–9(c) show the experimental outputs obtained from the detections of the multiple fingerprint by using the proposed JTC. The correlation output shown in Fig. 9(a) consists of two distinctive peaks

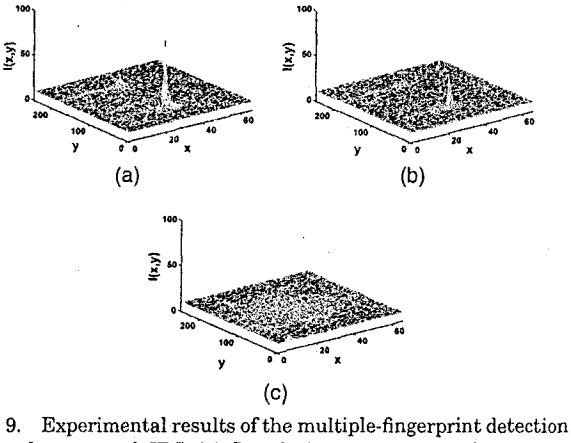

Fig. 9. Experimental results of the multiple-fingerprint detection using the proposed JTC. (a) Correlation output using the uncompressed fingerprint reference and correlation outputs using the compressed fingerprint reference (QF = 10) in a situation in which the target is (b) noise-free and (c) noisy ($\sigma^2 = 1$).

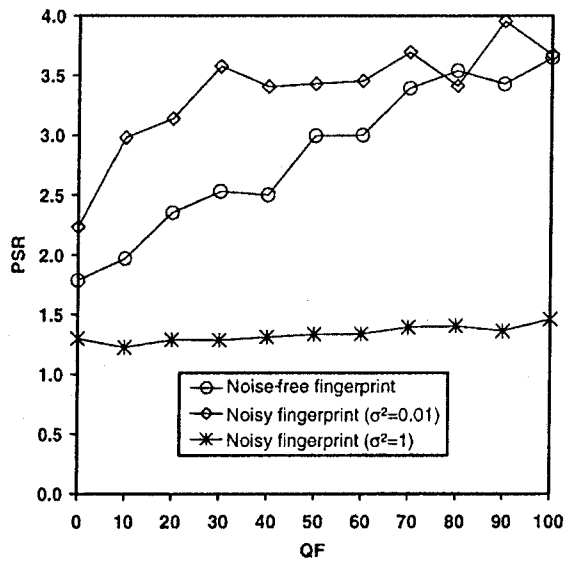

Fig. 10. PSR of multiple fingerprint detection as a function of the QF of the compressed fingerprint reference.

with different heights. The peak with the greater height is generated by the autocorrelation of the uncompressed high-contrast fingerprint, while the low peak is the cross correlation output of the reference and the nontarget fingerprint. Here the correlation peaks obtained from the detection of the target and the nontarget are defined as the primary and the secondary peaks, respectively. The decrease of both correlation peaks resulted from the detection of the noise-free multiple fingerprint by the compressed reference with a QF of 10 can be seen from Fig. 9(b). Figure 9(c) shows the detection of the noisy multiple-fingerprint target with a variance of $\sigma^2 = 1$ by the JTC using the reference compressed at a QF of 10. It is clear that no apparent correlation peaks are detected. In particular, the secondary correlation peak is minimized.

Figure 10 shows the PSRs as a function of the QF of the compressed fingerprint reference for different multiple-target scenes. In the detection of the noise-free target, the PSR increases gradually as the QF becomes higher. When the target is the noise corrupted multiple fingerprint with a variance of $\sigma^2 = 0.01$, its resultant PSR is higher than that of the noise-free target. As shown in Fig. 9(b) when the target is corrupted by noise, the decrease of the secondary peak caused by compression is more significant than that of the primary peak. Since the value of the denominator reduces faster than that of the numerator, the PSR becomes higher. Furthermore, as the variance of the noise increases to unity, the JPS is strongly corrupted by the noise. Since the primary correlation peak is further degraded, its PSR reduces sharply to less than 1.5. Although multiple-target detection depends on the compression of the reference images, the experimental results confirm that the proposed JTC can distinguish the desired target from the nontarget.

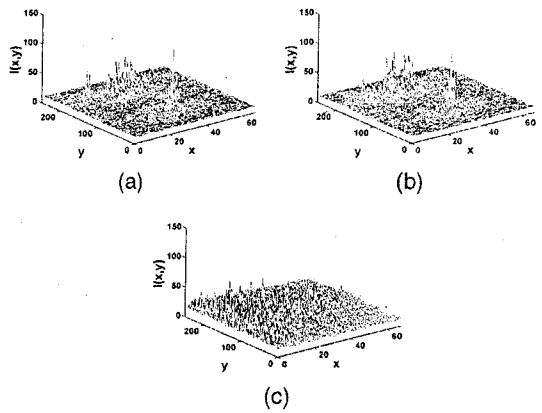

Fig. 11. Experimental results of the multiple human face detection using the proposed JTC. (a) Correlation output using the uncompressed human face reference and correlation outputs using the compressed human face reference (QF of 10) in a situation in which the target is (b) noise-free and (c) noisy ($\sigma^2 = 1$).

2. *Compressed High-Contrast Human Face as the Reference*

Figure 11(a) is a 3D plot of the correlation output of the multiple-human-face detection using the JTC with an uncompressed human face reference, and Figs. 11(b) and 11(c) illustrate the correlation outputs of the JTC using the compressed reference with a QF of 10. It can be seen from Fig. 11(a) that the primary correlation peak intensity is higher and sharper than that of the secondary peak. When the reference is compressed, the degradation of the primary peak intensity shown in Fig. 11(b) is insignificant compared with the secondary peak. However, when the multiple target is strongly corrupted by noise, there are no observable correlation peaks that can be seen from Fig. 11(c).

Figure 12 show the variation of the PSRs as a function of the QF of the compressed human face

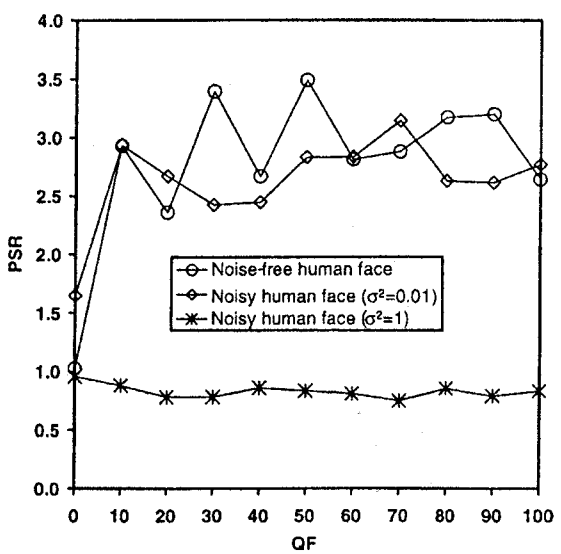

Fig. 12. PSR of multiple human face detection as a function of the QF of the compressed human face reference.

reference for different multiple-target scenes obtained from the experiments. The PSRs of the noise-free and the noisy high-contrast human face with variance $\sigma^2 = 0.01$ vary by approximately 3 except when the QF is smaller than 10. In the presence of strong noise in the input scene, the PSR reduces to approximately 1 because the fringes are degenerated. Although the speckle noise is present, the experimental results show that the multiple-human-face detection using the proposed JTC can be done for a wide range of compressions. This is in agreement with the simulation result.[9]

4. Conclusions

We have verified experimentally the JTC by using JPEG-compressed reference images, where high-contrast fingerprint and human face images were used as the test scenes. The experimental results confirm the feasibility of implementing the JTC with compressed reference images. However, unlike the simulation results,[8,9] the effect of additive noise on the detection performance of the proposed JTC is more severe than that of compression. This may be caused by the speckle noise that is inherently generated as a result of light scattering and multiple internal reflections that come from the EASLM and the CCD sensor.

This work was supported by the Institute of Research and Development, Suranaree University of Technology, under grant SUT 2548.

References

1. F. T. S. Yu, S. Jutamulia, T. W. Lin, and D. A. Gregory, "Adaptive real-time pattern recognition using liquid crystal TV based joint transform correlator," Appl. Opt. **26,** 1370–1372 (1988).
2. J. Widjaja, "Automatic holographic particle sizing using wavelet-based joint transform correlator," Optik **107,** 132–134 (1998).
3. S. Jutamulia, G. Mu, H. Zhai, and F. Song, "Use of laser diode in joint transform correlator," Opt. Eng. **43,** 1751–1758 (2004).
4. C. H. Loo and M. S. Alam, "Invariant object tracking using fringe-adjusted joint transform correlator," Opt. Eng. **43,** 2175–2183 (2004).
5. M. S. Alam and E. H. Horache, "Optoelectronic implementation of fringe-adjusted joint transform correlation," Opt. Commun. **236,** 59–67 (2004).
6. M. S. Alam and B. J. Thompson, eds., *Selected Papers on Optical Pattern Recognition Using Joint Transform Correlation* (SPIE Press, 1999).
7. W. B. Pennebaker and J. L. Mitchell, *JPEG Still Image Data Compression Standard* (Van Nostrand–Reinhold, 1993).
8. J. Widjaja and U. Suripon, "Real-time joint transform correlator using compressed reference images," Opt. Eng. **43,** 1737–1745 (2004).
9. J. Widjaja and U. Suripon, "Multiple-target detection by using joint transform correlator with compressed reference images," Opt. Commun. **253,** 44–55 (2005).
10. J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, 1996).
11. S. Yang, X. Zhang, and S. Mitra, "Performance of lossy compression algorithms from statistical and perceptual metrics," in *Proceedings of 12th IEEE Symposium on Computer-Based Medical Systems* (IEEE, 1999), pp. 242–247.
12. A. Tanone, C. M. Muang, F. T. S. Yu, E. C. Tam, and D. A. Gregory, "Effects of thresholding in joint-transform correlation," Appl. Opt. **31,** 4816–4822 (1992).
13. M. S. Alam and M. A. Karim, "Fringe-adjusted joint transform correlation," Appl. Opt. **32,** 4344–4350 (1993).
14. C. T. Li, S. Yin, and F. T. S. Yu, "Nonzero-order joint transform correlator," Opt. Eng. **37,** 58–65 (1998).