Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Optics Communications 253 (2005) 44–55
OPTICS COMMUNICATIONS
www.elsevier.com/locate/optcom

Multiple-target detection by using joint transform correlator with compressed reference images

Joewono Widjaja *, Ubon Suripon

Institute of Science, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand

Received 26 November 2004; received in revised form 29 March 2005; accepted 18 April 2005

Abstract

Effects of JPEG compression of reference image on multiple-target detection by using joint transform correlator are quantitatively studied by using computer simulation. Two types of images with different spatial-frequency contents are used as test scenes in the presence of noise in the input plane and the contrast difference. The results show that in comparison with the use of the compressed reference with high spatial-frequency contents, the multiple-target detection by using the joint transform correlator with the compressed reference with low spatial-frequency contents produces better detection performance in that it is robust to noise and contrast difference for a wide range of compression qualities. While in the presence of noise and contrast difference, the compression of the reference image with high spatial-frequency contents may cause false alarms.

PACS: 42.30.Sy; 42.79.Hp

Keywords: Joint transform correlator; Multiple-target detection; JPEG-image compression; Detection performance

1. Introduction

With a rapid development of technologies for array detector and display, joint transform correlator (JTC) has been found to be useful for implementing real-time pattern recognition and classification [1–5], because it does not need a priori synthesis of matched-filters. In real-time JTC, a target image captured by a charge-couple device (CCD) sensor and a reference image stored in a computer system are displayed side-by-side onto an electrically addressed spatial light modulator (EASLM) placed in a front focal plane of a Fourier transforming lens. By illuminating perpendicularly the EASLM with collimated laser light, the joint power spectrum (JPS) of the input images

* Corresponding author. Tel.: +66 44 224194; fax: +66 44 224185.
E-mail address: widjaja@ccs.sut.ac.th (J. Widjaja).

0030-4018/$ - see front matter
doi:10.1016/j.optcom.2005.04.042

are generated at the back focal plane of the lens. The JPS is captured by the CCD sensor, and then is re-Fourier transformed either optically or digitally in order to obtain the correlation output. By quantifying the correlation output, the degree of similarity between the target and the reference images can be measured. Although the JTC architecture is indeed useful for implementing classification systems, the process of displaying the target and the reference images onto the EASLM introduces a time delay that is caused by a serial nature of signal communication between the computer and the EASLM. This time delay is dependent upon the file size of the images. Furthermore, in real-world automatic target recognition, a number of target images to be detected always increases. Since the JTC is sensitive to rotation and scale changes of the target, a database of the reference images must contain all possible variations of rotation and scale changes of the target. Thus, the application of the JTC to automatic target recognition requires considerable storage capability.

In order to solve these problems, a real-time JTC by using JPEG-compressed reference images has been proposed by Widjaja and Suripon [6]. In our previous work, the recognition performance of JTC under situations where the target is corrupted by noise and has contrast difference with respect to the reference are quantitatively studied by using two types of images with different spatial-frequency contents. In comparison with the use of the compressed reference with high spatial-frequency content, the results show that the JTC by using the 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 order to further verify the feasibility of our proposed method, computer simulation of the multi-target detection by using the JTC with compressed reference images is done by taking into account the presence of noise and contrast difference. In this work, the JPEG algorithm is also employed to compress the reference images. The detection performance is measured by using a ratio of the autocorrelation peak intensity to the intensity of the secondary peak (PSR).

2. Theoretical background

2.1. Multiple-target JTC using compressed reference images

A schematic diagram for implementing the real-time multiple-target detection by using the JTC with compressed reference images is illustrated in Fig. 1. The compressed reference $r_c(x,y)$ and the input scene consisting of N target images $t_i(x,y)$ are displayed side-by-side with a separation of $2x_0$ onto the EASLM. This joint input image can be mathematically written as

$$f(x,y) = r_c(x - x_0, y) + \sum_{i=1}^{N} t_i(x + x_i + x_0, y + y_i), \tag{1}$$

where x_i and y_i correspond to the relative position of the target in the x and the y-directions, respectively. Under a presence of an additive white Gaussian noise $n(x,y)$ at the input scene and a contrast difference between the target and the compressed reference images, the joint input image can be rewritten as

$$f(x,y) = r_c(x - x_0, y) + c_T \sum_{i=1}^{N} t_i(x + x_i + x_0, y + y_i) + n(x + x_0, y), \tag{2}$$

where c_T is the amplitude ratio of the target to the reference images. c_T becomes greater, equal, or smaller than 1 when the contrast of the reference

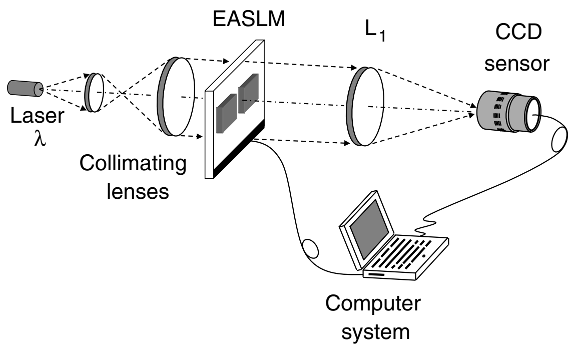

Fig. 1. Schematic diagram of optical setup for implementing real-time multiple-target detection by using the JTC with compressed reference images.

image is lower, equal, or higher than that of the target, respectively. After a Fourier transformation by the lens L_1, the JPS captured by the CCD sensor is found to be

$$
\begin{aligned}
U(u,v) = {} & |R_c(u,v)|^2 + c_T^2 \sum_{i=1}^{N} |T_i(u,v)|^2 + |N(u,v)|^2 + c_T N(u,v) \left[\sum_{i=1}^{N} T_i^*(u,v) \exp[-j2\pi(x_i u + y_i v)] \right] \\
& + c_T N^*(u,v) \left[\sum_{i}^{N} T_i(u,v) \exp[j2\pi(x_i u + y_i v)] \right] \\
& + c_T^2 \left[\sum_{i=1}^{N} \sum_{\substack{k=1 \\ k \neq i}}^{N} T_i^*(u,v) T_k(u,v) \exp\{-j2\pi[(x_i - x_k)u + (y_i - y_k)v]\} \right] \\
& + c_T^2 \left[\sum_{i=1}^{N} \sum_{\substack{k=1 \\ k \neq i}}^{N} T_i(u,v) T_k^*(u,v) \exp\{j2\pi[(x_i - x_k)u + (y_i - y_k)v]\} \right] \\
& + c_T R_c(u,v) \left[\sum_{i=1}^{N} T_i^*(u,v) \exp\{-j2\pi[(x_i + 2x_0)u + y_i v]\} \right] \\
& + c_T R_c^*(u,v) \left[\sum_{i=1}^{N} T_i(u,v) \exp\{j2\pi[(x_i + 2x_0)u + y_i v]\} \right] \\
& + R_c(u,v) N^*(u,v) \exp(-j4\pi x_0 u) + N(u,v) R_C^*(u,v) \exp(j4\pi x_0 u),
\end{aligned}
\tag{3}
$$

where (u,v) are the spatial-frequency coordinates in the horizontal and the vertical directions at the Fourier plane, respectively. They are related to the actual coordinates (x,y) by $x = \lambda f u$ and $y = \lambda f v$, with λ and f stand for the wavelength and the focal length of the lens L_1, respectively. In Eq. (3), $R_c(u,v)$, $T_i(u,v)$ and $N(u,v)$ are the Fourier transforms of the compressed reference, the ith target, and the noise, respectively. The first three terms of Eq. (3) are associated with the autocorrelations of the reference, the input targets, and the noise, respectively. The fourth and fifth terms are cross-correlations between the targets and the noise. The cross-correlations between different targets correspond to the sixth and seventh terms, while the eighth and the ninth are the cross-correlations between the reference and the input targets. The last two terms correspond to the cross-correlation between the reference and the noise. After displaying the captured JPS onto the EASLM, the second Fourier transformation produces the correlation output at the back focal plane of the lens L_1. The correlation signals corresponding to the last four terms are of particular interest, because these terms appear at the same position. They can be mathematically written as:

$$
\begin{aligned}
C(x,y) = {} & r_c(x,y) * c_T \sum_{n=1}^{N} t_i(x,y) \\
& * \delta[x \pm (x_i + 2x_0), y \pm y_i] \\
& + r_c(x,y) * n(x,y) * \delta(x \pm 2x_0, y).
\end{aligned}
\tag{4}
$$

The first term of Eq. (4) corresponds to the desired cross-correlation of the input target scene with the compressed reference images which is scaled by the contrast ratio c_T. The second one is the unwanted correlation of the compressed reference with the noise. However since the input scene consists of multiple targets, besides the contrast difference, the detection of the correct target may be affected by the correlation of the compressed reference with the non-target and the noise.

2.2. JPEG compression algorithm

In the JPEG algorithm, an image compression is achieved by discarding information that is not easily perceived by human eye such as high spatial-frequency components and small variations in color of images. Hence, the JPEG performs well on the continuous-tone images, while the image with many sudden jumps in intensity or color will not be compressed well. The JPEG compression algorithm employs a discrete cosine transform (DCT) to provide information of images in the spatial-frequency domain. By quantizing the spatial-frequencies and rounding its quantized information to the nearest integer, the redundant information is discarded. Further compression is done by encoding the results with a run length encoding and Huffman coding. This process compresses the data content of the source image without degradation of image quality. A rigorous discussion of the JPEG-image compression can be found in [7].

3. Results and discussions

3.1. Compressed reference images

In order to study the effects of the reference image compression on the multiple-target detection by using the JTC, fingerprint and human face images with high and low contrasts were employed as the target and the reference images. The size of each image was 23 KB and consisted of 124×186 pixels with gray scale levels. The input target scene contained two different images with one of them was identical to the original reference image. Fig. 2(a) and (b) show the input scene and the reference for the high-contrast fingerprint and human face, while their low-contrast images are shown in Figs. 2(c) and (d). The original reference images were compressed into the JPEG format by using the ACDsee software (The 2000 ACD systems, Ltd.). In this software, the quality of the compressed images is determined by a parameter called the quality factor (QF) whose value can be varied from 0 to 100. The lower the QF, the better the compression. Fig. 3 shows the compression ratio (CR) of each reference image as a function of the QF. The CR defined as the ratio of the original to the compressed image size describes the amount of storage size saved by the compression. It can be seen from Fig. 3 that the CR of the human face image is higher than that of the fingerprint image, because the human face image contains more low spatial-frequency components. In the quantization process, reduction of the magnitude of these frequency components increases the number zeroes. As a consequence, more redundant information is discarded. Since the low-contrast image also contains more low spatial-frequency components, its CR becomes higher than that of the high-contrast image.

3.2. Simulation results

In the simulation, the input target scene and the compressed reference image were combined to form the joint input image with the separation $2x_0 = 248$ pixels, while the distance between the desired target and the undesired object images was 186 pixels in the y-direction. The Fourier transformation and the noise were calculated by using the 2-D FFT and the IMNOSIE commands of Matlab 6.0 run on a personal computer, respectively. The JPS was computed by taking the modulus square of the generated Fourier spectrum of the joint input image. By calculating the inverse Fourier transform of the JPS and then taking the modulus squared of the result, the correlation output intensity was generated. Finally, the correlation quality of the multiple-target detection output was quantified by measuring the PSR in the correlation plane.

3.2.1. Compressed high-contrast fingerprint as the reference image

The three-dimensional (3-D) plots of the correlation output of the multiple-fingerprint detection by using the JTC with the high-contrast fingerprint reference are shown in Fig. 4. The correlation output shown in Fig. 4(a) consists of two peaks, because there are two targets. The first peak

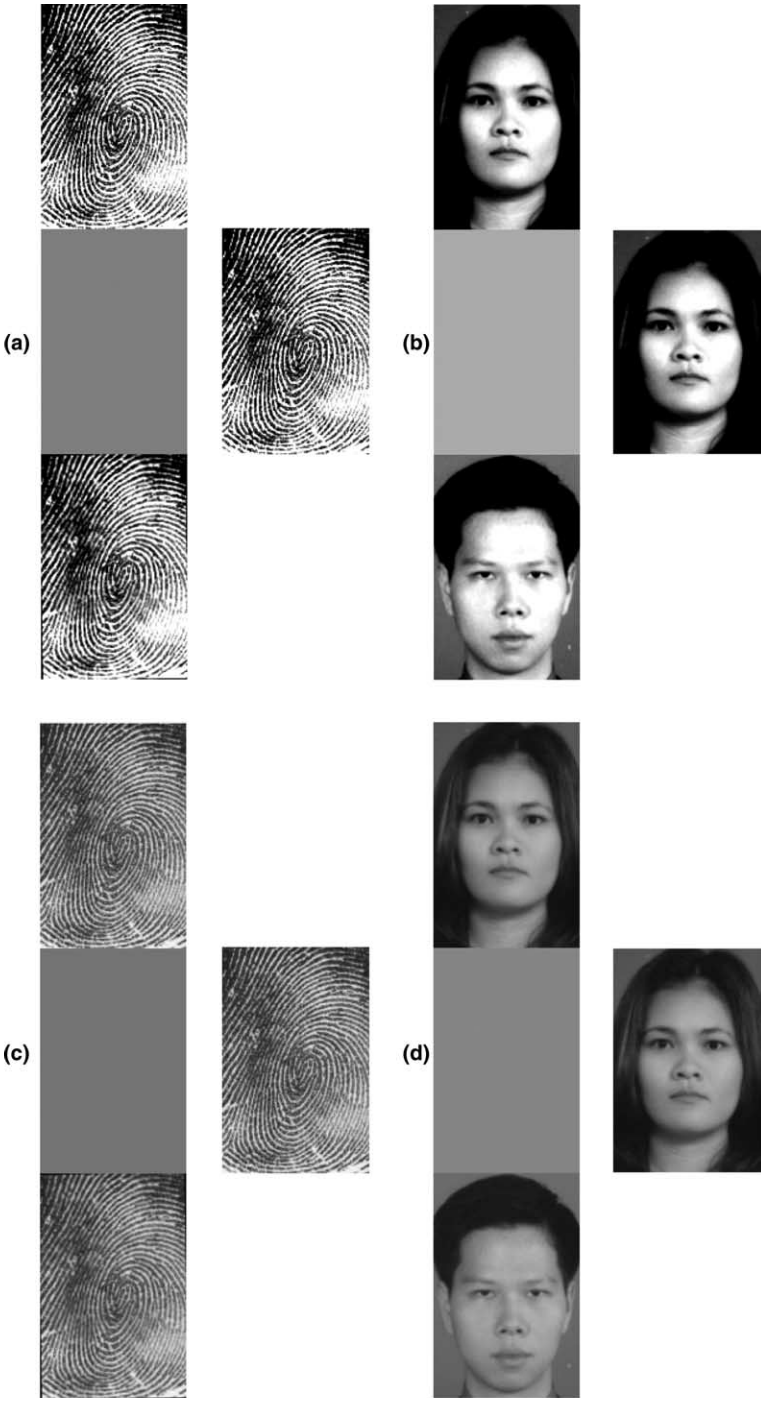

Fig. 2. Multiple-target and reference images: (a) high-contrast fingerprint, (b) high-contrast human face, (c) low-contrast fingerprint, and (d) low-contrast human face.

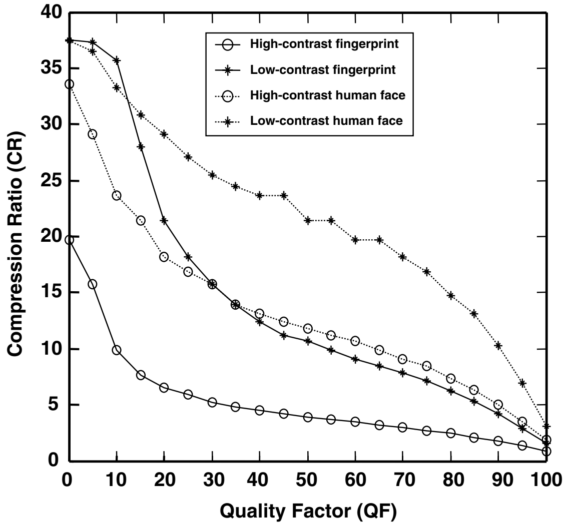

Fig. 3. The CR as function of the QF.

produced by the autocorrelation of the uncompressed high-contrast fingerprint, while the second one is the cross-correlation of the reference with the non-target fingerprint. For this reason, the first peak is higher and sharper than the undesired secondary peak. Fig. 4(b)–(d) shows the correlation outputs of the JTC by using the compressed high-contrast fingerprint at QF = 10 as the reference. In Fig. 4(b), the high-contrast multiple-fingerprint target is free from noise. The resultant correlation peak is slightly lower compared with that of the uncompressed case shown in Fig. 4(a). This is the effect of the compression on the reference image [6,8]. Fig. 4(c) and (d) show the resultant detection of the high- and low-contrast multiple-targets which are corrupted by noise with variance $\sigma^2 = 1$, respectively. In the case of the noisy high-contrast fingerprint target, the correlation plane is slightly noisy and the peak intensity of both desired and undesired correlations are about one order of magnitude lower than the noise-free multiple-target. However, in the case of the noisy low-contrast target, the correlation plane appears very noisy. This makes the determination of the correct correlation peaks become difficult. The reason of this is that the low-contrast image has luminance which is weaker than the noise [6]. This produces noisy correlation plane.

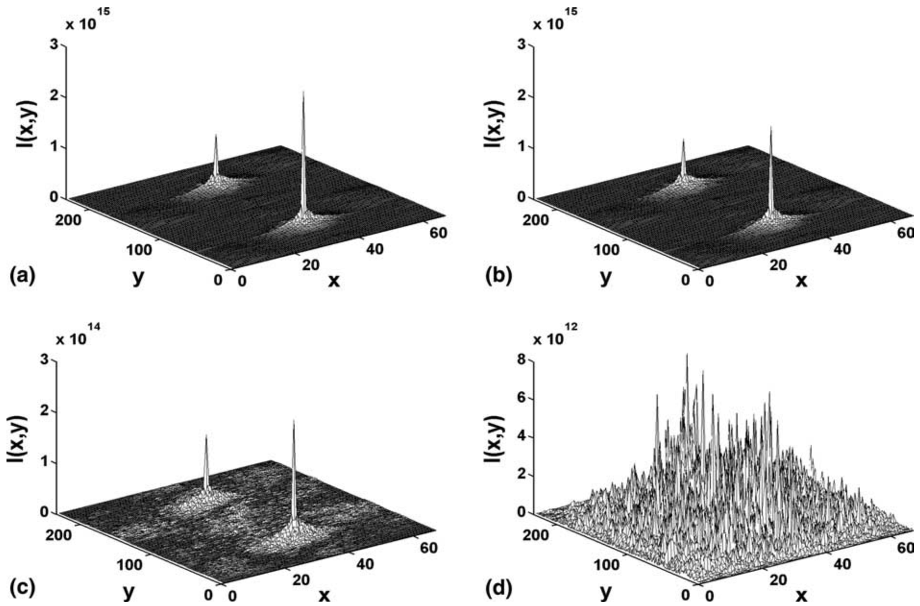

Fig. 4. Simulation results of the multiple-target JTC, (a) autocorrelation of the uncompressed high-contrast fingerprint; and cross-correlation outputs by using the compressed high-contrast fingerprint reference (QF = 10) under a situation that the multiple-target scene is: (b) noise-free high-contrast fingerprint image, (c) noisy high-contrast fingerprint image ($\sigma^2 = 1$) and (d) noisy low-contrast fingerprint image ($\sigma^2 = 1$).

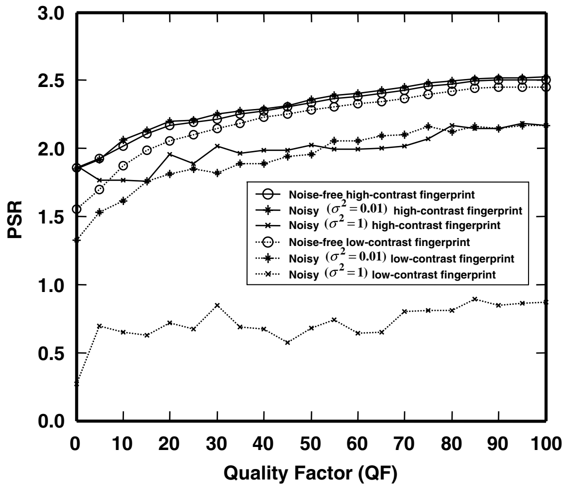

Fig. 5. The PSR-based measurement of the detection performance of the JTC as a function of the QF of the compressed high-contrast fingerprint reference.

Since the target has lower contrast than the reference, the correlation of the compressed reference with the target is reduced by a factor c_T which is less than 1. Because of these two reasons, the effect of noise on the recognition of the low-contrast target is found to be stronger than that of the high-contrast target with same noise level.

Fig. 5 shows the variation of the PSRs as a function of the QF of the compressed high-contrast fingerprint reference for different multiple-target scenes. In general, the PSRs gradually increase as the QF increases, because less information is discarded from the compressed reference image. When the degree of similarity between the target and the compressed reference images improves, the correlation quality becomes better and the PSRs increase. However, the PSRs of the low-contrast fingerprint target are lower than that of the high-contrast target and they decrease drastically when the noise level increases. As mentioned in the preceding paragraph, this is the consequence of scaling of the correlation function by the contrast ratio and also by the sensitivity of the low-contrast image to noise. Therefore, the recognition of the noisy multiple-fingerprint depends on the contrast rather than the compression level of the reference image.

3.2.2. Compressed low-contrast fingerprint as the reference image

Fig. 6(a)–(d) illustrates the 3-D correlation outputs of the multiple-fingerprint detection by using JTC with the low-contrast fingerprint as the reference image. The autocorrelation peak of the uncompressed low-contrast fingerprint image shown in Fig. 6(a) is almost as sharp as the autocorrelation of the high-contrast target shown in Fig. 4(a). However, its peak intensity reduces by three orders of magnitude, because the contrast is low. The secondary peak is also lower and broader than that of the high-contrast fingerprint. Fig. 6(b) illustrates the output correlation of the noise-free low-contrast multiple-fingerprint with the low-contrast reference compressed at QF = 10. Both correlation peaks become broad and their peaks further decrease. This occurs because the compressed low-contrast fingerprint reference image suffers more from loss of high spatial-frequency contents. As the compressed reference image now contains mainly the low spatial-frequency components, the impulse response of the JTC with this compressed reference image becomes broad. Fig. 6(b) also reveals that, because of the compression, the intensity of the desired correlation peak is reduced more than that of the undesired peak, yielding false detection. Fig. 6(c) and (d) shows the detection outputs of the noisy low- and high-contrast multiple-fingerprint target with variance $\sigma^2 = 1$, respectively. Since the luminance of the noisy low-contrast multiple-target is weaker than the noise, the correlation term $r_c(x,y) * n(x,y)$ gives the greatest output. As a result, the correlation peaks are buried in strong noise. However, when the target scene is the noisy high-contrast fingerprint with variance $\sigma^2 = 1$, the target luminance is stronger than the noise and the contrast ratio c_T is greater than unity. The output correlation of the compressed reference with the multiple-target becomes greater than that of the other terms. Therefore, although the correlation plane is noisy, the correlation peaks can be clearly observed.

Fig. 7 shows the PSRs as a function of the QF of the compressed low-contrast fingerprint reference for different multiple-target scenes. It is obvious that as the QF increases, the PSRs increase more

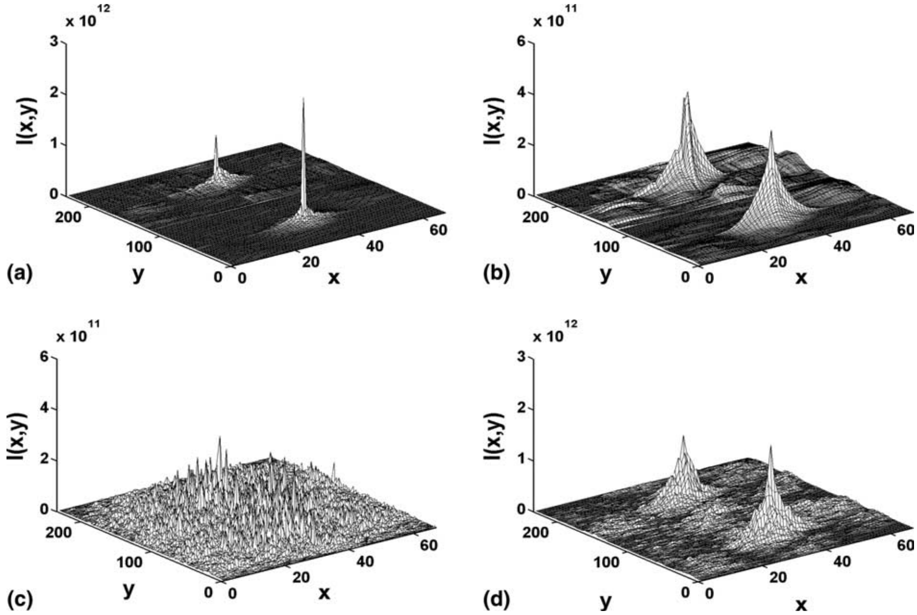

Fig. 6. Simulation results of the multiple-target JTC, (a) autocorrelation of the uncompressed low-contrast fingerprint; and cross-correlation outputs by using the compressed low-contrast fingerprint reference (QF = 10) under a situation that the multiple-target scene is: (b) noise-free low-contrast fingerprint, (c) noisy low-contrast fingerprint ($\sigma^2 = 1$) and (d) noisy high-contrast fingerprint ($\sigma^2 = 1$).

rapidly than that of the compressed high-contrast fingerprint shown in Fig. 5. This indicates that the use of the compressed low-contrast fingerprint reference degrades significantly the performance

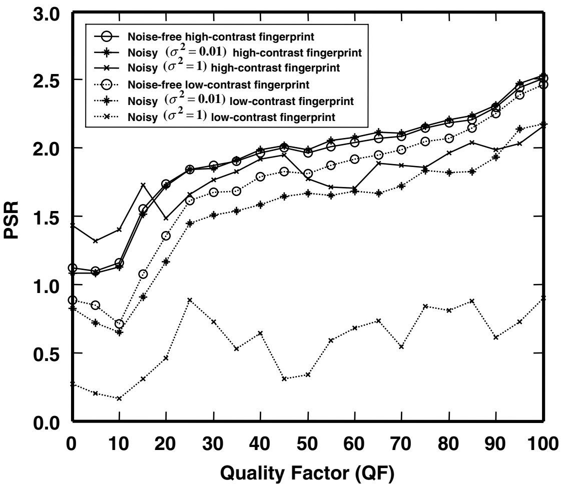

Fig. 7. The PSR-based measurement of the detection performance of the JTC as a function of the QF of the compressed low-contrast fingerprint reference.

of multiple-target detection by the JTC. The severe effect of compression on the detection of the low-contrast multiple-fingerprint can be observed when the QF is less than 20 in which the PSRs become smaller than unity. It is obvious that in comparison with the high-contrast target, the detection of the low-contrast multiple-fingerprint target is dependent more upon the noise. When the variance of the noise is equal to 1, the PSR of the noisy low-contrast multiple-fingerprint never exceeds unity. This is because the low-contrast target image is easily corrupted by the noise. Therefore, besides the compression, the detection of the noisy multiple-target is dependent upon the contrast difference between the target and the reference image. In order to avoid the false detection, the low-contrast fingerprint reference cannot be compressed as small as the high-contrast reference.

3.2.3. Compressed high-contrast human face as the reference image

Fig. 8(a) shows the 3-D plot of the output correlation of the high-contrast multiple-human face

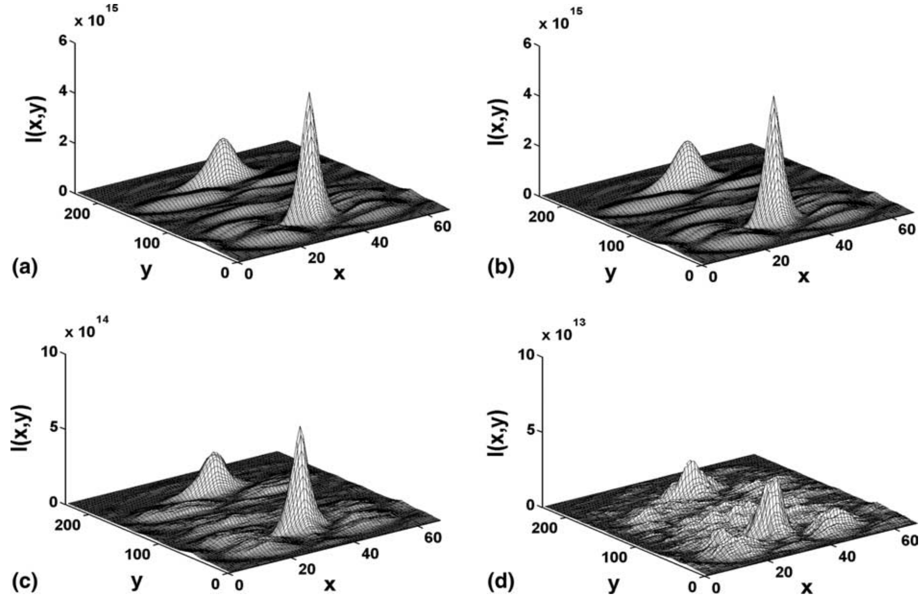

Fig. 8. Simulation results of the multiple-target JTC, (a) autocorrelation of the uncompressed high-contrast human face; and cross-correlation outputs by using the compressed high-contrast human face reference (QF = 10) under a situation that the multiple-target scene is: (b) noise-free high-contrast human face, (c) noisy high-contrast human face ($\sigma^2 = 1$) and (d) noisy low-contrast human face ($\sigma^2 = 1$).

with the uncompressed reference. It is clear that the desired correlation output is sharper and its peak intensity is higher than that of the undesired peak. The broadening of the autocorrelation peak is because the human face reference image contains less high spatial-frequency components. Fig. 8(b)–(d) illustrate the correlation outputs of the JTC by using the compressed high-contrast human face as the reference at QF = 10 for different multiple-target scenes. Fig. 8(b) illustrates the output detection of the noise-free high-contrast multiple-target which is slightly affected by the compression. Fig. 8(c) and (d) show the resultant detections of the high- and the low-contrast multiple-targets which are corrupted by noise with variance $\sigma^2 = 1$, respectively. It can be seen that the correlation outputs depend on the contrast difference. Due to the lower contrast, the peak intensities of Fig. 8(d) are lower and its correlation plane is more noisy than that of Fig. 8(c).

The variation of PSRs as a function of the QF of compressed high-contrast human face reference for the different multiple-target scenes is shown in Fig. 9. It is obvious that all PSRs are almost independent of the compression. Their magnitudes which are always greater than

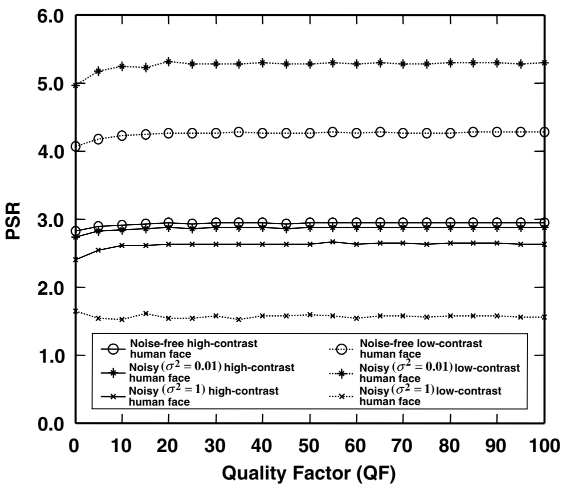

Fig. 9. The PSR-based measurement of the detection performance of the JTC as a function of the QF of the compressed high-contrast human face reference.

unity depend on the contrast difference and noise. This ensures that false alarm will not occur in the correlation plane. In the case of the detection of the multiple high-contrast human face, the PSRs are slightly affected by the noise. However, contrary results are obtained when the contrast of the detected target are low. In the case of the noise-free multiple-target, the correlation of the low-contrast non-target human face with the compressed high-contrast reference produces lower peak intensity than that of the high-contrast non-target, because of the contrast difference. As a result, the detection of the low-contrast target yields a higher PSR than that of the high-contrast target. When the multiple low-contrast human face is corrupted by the noise with variance $\sigma^2 = 0.01$, the peak intensity produced by the correlation of the non-target is further degraded because of the hyper sensitive nature of the low-contrast image to noise. Thus, its PSR increases further. However, when the variance of the noise increases to be 1, the PSR of the low-contrast multiple-target reduces drastically. This is because of the stronger noise signal than the target image, both correlation terms produced by the target and the non-target

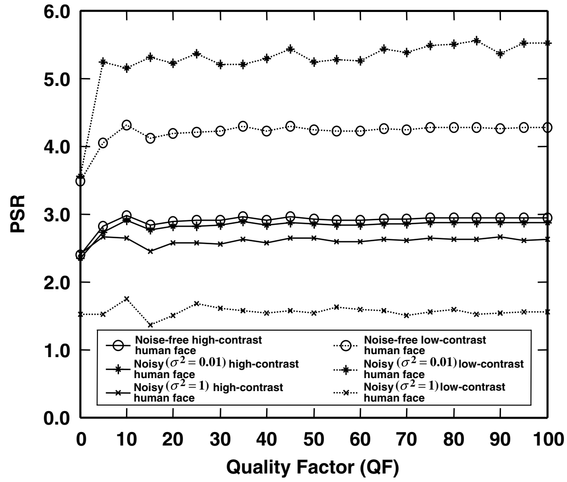

Fig. 11. The PSR-based measurement of the detection performance of the JTC as a function of the QF of the compressed low-contrast human face reference.

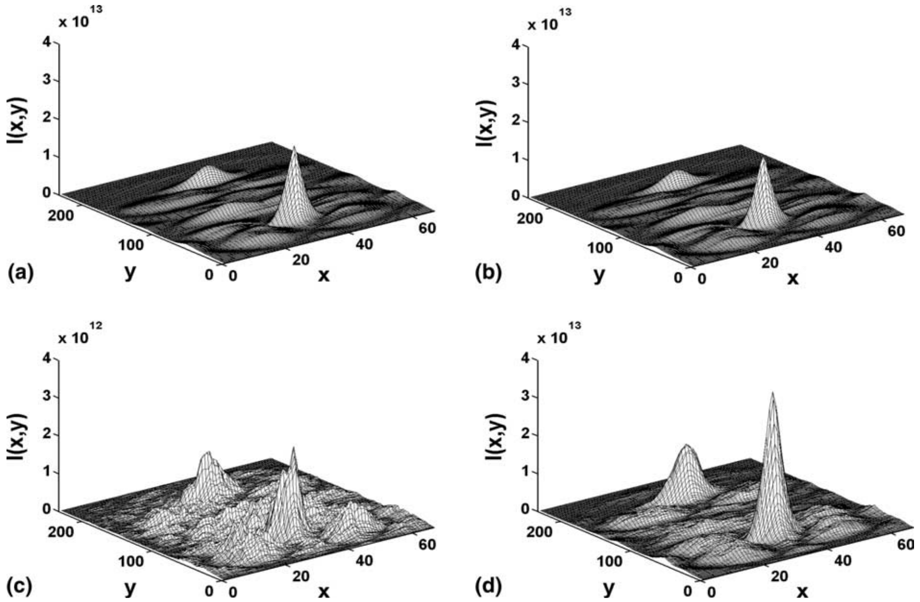

Fig. 10. Simulation results of the multiple-target JTC, (a) autocorrelation of the uncompressed low-contrast human face; and cross-correlation outputs by using the compressed low-contrast human face reference (QF = 10) under a situation that the multiple-target scene is: (b) noise-free low-contrast human face, (c) noisy low-contrast human face ($\sigma^2 = 1$) and (d) noisy high-contrast human face ($\sigma^2 = 1$).

Table 1
Performance degradation of the multiple-target detection by using the JTC with compressed reference images

Compressed reference image	Condition of multiple-target images			
	Noise free		Noisy	
	High contrast	Low contrast	High contrast	Low contrast
High-contrast fingerprint	Gradual	Gradual and significant	Gradual and significant	Severe
Low-contrast fingerprint	Sudden	Sudden and significant	Sudden and significant	Severe
High-contrast human face	No	No	No	No
Low-contrast human face	Very small	Very small	Very small	Very small

are corrupted by the strong correlation of term $r_c(x, y) * n(x, y)$.

3.2.4. Compressed low-contrast human face as the reference image

Fig. 10(a)–(d) show the correlation outputs of the multiple-human face detection by using JTC with the low-contrast human face as the reference image. Fig. 10(a) shows the output detection of the noise-free low-contrast human face by using the uncompressed reference. The correlation peaks are broad and their peak intensities are reduced by about two orders of magnitude compared with that of the high-contrast human face target, because of the lower contrast. When the reference image is compressed at QF = 10, the intensity of both correlation peaks in Fig. 10(b) further decreases, and their correlation profiles broaden. Fig. 10(c) and (d) show the detection outputs of the noisy low- and the high-contrast multiple-targets with variance $\sigma^2 = 1$, respectively. It is clear that when the target is high-contrast human face, the degradation of correlation output caused by the noise is less than that of the low-contrast target. The desired correlation peak is even higher than the case of autocorrelation shown in Fig. 10(a). This is because the image contrast is higher.

Fig. 11 shows the variation of PSRs as a function of the QF of the compressed low-contrast human face reference for different multiple-target scenes. In comparison with the compressed high-contrast human face, the use of the compressed low-contrast gives similar results except that the PSRs decrease abruptly when the compression QF becomes very low. This is caused by the degradation of the reference image by JPEG compression.

4. Conclusions

In order to solve storage problem and improve the response time of the real-time multiple-target detection by using the JTC, the effects of compression of the reference image on the detection performance have been quantitatively studied by using the computer simulation. This study employed the fingerprint and the human face images as test scenes in the presence of noise in the input plane and the contrast difference that may arise from unbalanced illuminations. Table 1 summarizes the performance degradation of the multiple-target detection by using the JTC with compressed reference images. When the reference image has low spatial-frequency contents such as human face, the effects of compression of the reference on the multiple-target detection by using JTC is not significant for all given target scenes regardless of the noise and the contrast difference. This is an agreement with our previous study of the single target detection [6]. This result verifies the feasibility of our proposed method. Our study shows that in contrast with the use of the compressed reference with low spatial-frequency contents, the multiple-target recognition by using the compressed reference with high spatial-frequency contents is not only determined by the contrast, but also the noise and the compression as well. Finally, it is worth mentioning that the compression of the low-contrast reference with high spatial-frequency contents may yield false alarms.

Acknowledgement

This work was supported by the Institute of Research and Development, Suranaree University of

Technology, under Grants SUT 2547 and SUT 2548.

References

[1] F.T.S Yu, X.J. Lu, Opt. Commun. 52 (1984) 10.
[2] F.T.S. Yu, S. Jutamulia, T.W. Lin, D.A. Gregory, Appl. Opt. 26 (1987) 1370.
[3] J. Widjaja, Y. Tomita, J. Mod. Opt. 43 (1996) 1993.
[4] M.S. Alam, M. Akhteruzzaman, Proc. IEEE Natl. Aerospace Electron. Conf. (2000) 434.
[5] A. Alsamman, M.S. Alam, Opt. Eng. 42 (2003) 560.
[6] J. Widjaja, U. Suripon, Opt. Eng. 43 (2004) 1737.
[7] W.B. Pennebaker, J.L. Mitchell, JPEG Still Image Data Compression Standard, Van Nostrand Reinhold, New York, 1993.
[8] J. Widjaja, Opt. Lasers Eng. 39 (2003) 501–506.