AUTOMATIC TARGET RECOGNITION BY USING

JOINT TRANSFORM CORRELATOR WITH

COMPRESSED REFERENCE IMAGES

Ubon Suripon

A Thesis Submitted in Partial Fulfillment of the Requirements for

the Degree of Doctor of Philosophy in Laser Technology

Suranaree University of Technology

Academic Year 2005

การรู้จำเป้าหมายอัตโนมัติด้วยจอยท์ทรานส์ฟอร์มคอร์รีเลเตอร์
โดยใช้ภาพอ้างอิงแบบบีบอัด

นางสาวอุบล สุริพล

AUTOMATIC TARGET RECOGNITION BY USING

JOINT TRANSFORM CORRELATOR WITH

COMPRESSED REFERENCE IMAGES

Suranaree University of Technology has approved this thesis submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy.

Thesis Examining Committee

_____
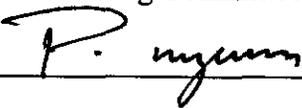
(Assoc. Prof. Dr. Prapun Manyum)

Chairperson

_____
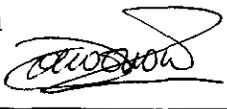
(Assoc. Prof. Dr. Joewono Widjaja)

Member (Thesis Advisor)

_____
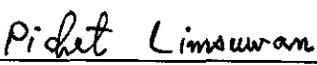
(Assoc. Prof. Dr. Pichet Limsuwan)

Member

_____
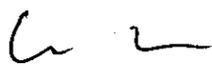
(Asst. Prof. Dr. Eckart Schulz)

Member

_____
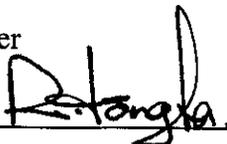
(Dr. Rangsan Tongta)

Member

_____   _____
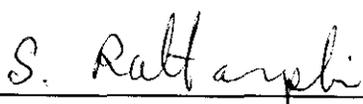
(Assoc. Prof. Dr. Saowanee Rattanaphani)   (Assoc. Prof. Dr. Sompong Thammathawom)

Vice Rector for Academic Affairs   Dean of Institute of Science

อุบล สุริพล : การรู้จำเป้าหมายอัตโนมัติด้วยจอยท์ทรานส์ฟอร์มคอร์รีเลเตอร์โดยใช้ภาพ
อ้างอิงแบบบีบอัด (AUTOMATIC TARGET RECOGNITION BY USING
JOINT TRANSFORM CORRELATOR WITH COMPRESSED
REFERENCE IMAGES) อาจารย์ที่ปรึกษา : รองศาสตราจารย์ ดร.ยูโวโน วิดจายา,
117 หน้า. ISBN 974-533-490-1

งานวิทยานิพนธ์นี้นำเสนอจอยท์ทรานส์ฟอร์มคอร์รีเลเตอร์แบบทันทีโดยใช้ภาพอ้างอิงที่
ถูกบีบอัดแบบ JPEG เพื่อใช้แก้ปัญหาการจัดเก็บและปรับปรุงการตอบสนองทางเวลาของระบบ
การรู้จำอัตโนมัติ โดยศึกษาผลกระทบของการบีบอัดภาพอ้างอิงต่อประสิทธิภาพการรู้จำรูปแบบ
ของจอยท์ทรานส์ฟอร์มคอร์รีเลเตอร์ ทั้งการตรวจหาเป้าหมายแบบเป้าหมายเดี่ยวและแบบหลาย
เป้าหมายภายใต้สภาวะที่ภาพเป้าหมายถูกรบกวนด้วยสัญญาณรบกวนและมีความแตกต่างของ
ความเปรียบต่างเมื่อเทียบกับภาพอ้างอิง ซึ่งภาพที่ใช้ในการศึกษามีทั้งภาพที่มีรายละเอียดของภาพ
สูงและภาพที่มีรายละเอียดของภาพต่ำ ผลการศึกษาจากการจำลองสถานการณ์ด้วยคอมพิวเตอร์
พบว่าประสิทธิภาพการรู้จำรูปแบบของจอยท์ทรานส์ฟอร์มคอร์รีเลเตอร์โดยใช้ภาพอ้างอิงแบบบีบ
อัดที่มีรายละเอียดต่ำดีกว่าภาพอ้างอิงที่มีรายละเอียดสูง กล่าวคือสามารถทนต่อการรบกวนของ
สัญญาณรบกวนและความแตกต่างของความเปรียบต่างที่ระดับการบีบอัดที่กว้าง อย่างไรก็ตาม
เนื่องจากข้อจำกัดของ EASLM ในการแสดงภาพที่มีความเปรียบต่างต่ำ และค่าไดนามิกเรนจ์ของ
อุปกรณ์รับรู้ CCD ที่ค่อนข้างต่ำทำให้ไม่สามารถทำการทดลองเพื่อยืนยันผลจากการจำลอง
สถานการณ์ด้วยคอมพิวเตอร์ได้อย่างครบถ้วนสมบูรณ์

สาขาวิชาเทคโนโลยีเลเซอร์และโฟตอนนิกส์   ลายมือชื่อนักศึกษา _____
ปีการศึกษา 2548   ลายมือชื่ออาจารย์ที่ปรึกษา _____
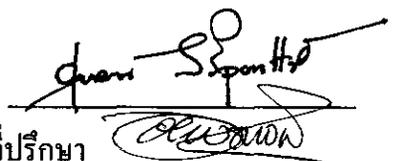
UBON SURIPON : AUTOMATIC TARGET RECOGNITION BY USING JOINT TRANSFORM CORRELATOR WITH COMPRESSED REFERENCE IMAGES. THESIS ADVISOR : ASSOC. PROF. JOEWONO WIDJAJA, Ph.D. 117 PP. ISBN 974-533-490-1

JOINT TRANSFORM CORRELATOR/JPEG-IMAGE COMPRESSION/OPTICAL PATTERN RECOGNITION/RECOGNITION PERFORMANCE

In order to solve storage problems and improve time response of automatic target recognition systems, a real-time joint transform correlator by using JPEG-compressed reference images is proposed. This dissertation discusses the effects of the compression of reference images on recognition performance of the joint transform correlator. Detections of single- and multiple-target detections are performed by computer simulations and experiments, where two types of images with different spatial-frequency contents are used as the test scenes in the presence of noise in the input plane and the contrast difference. The simulation results show that in comparison with the use of the compressed reference with high-spatial-frequency contents, the target detection by using the joint transform correlator with the compressed reference with low-spatial-frequency contents offers better detection performance in that it is robust to noise and contrast difference for a wide range of compression levels. However, due to the limitation of the EASLM on displaying low-contrast image and low dynamic range of the CCD sensor, not all experimental verifications are completely performed.

School of Laser Technology and Photonics   Student's Signature

Academic Year 2005   Advisor's Signature

ACKNOWLEDGEMENTS

CONTENTS

CONTENTS (Continued)

CONTENTS (Continued)

CONTENTS (Continued)

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

CHAPTER I

INTRODUCTION

1.1　Automatic Target Recognition

Automatic target recognition is a process of detection and classification used in various automatic functions (Bhanu and Jones, 1993; Ratches, Walters, Buser, and Guenther, 1997). There are various applications of the automatic target recognition to real world problems, such as personal identification, military surveillance, passport authentication, robotic navigation and manufacturing. In personal identification, human faces, fingerprints, and iris are of targets to be identified, while military surveillance deals widely with the recognition of military vehicles. In general, the classification is done by comparing the target captured by an image sensor against known objects stored in a reference template library until the target is identified. In this sense, the automatic target recognition can be performed by using a correlation method or template matching which is a mathematical method for comparing a similarity between two functions and defined as (Feitelson, 1988)

$$C(x, y) = \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} t(x', y') r^*(x'-x, y'-y) dx' dy'. \tag{1-1}$$

Here $r(x, y)$ and $t(x, y)$ are the reference and the target images, respectively. In the correlation method, the height and sharpness of the correlation function $C(x, y)$ measures the similarity between the target and the reference images. If they match, the integral of Eq. (1-1) has large values over a sharp, narrow correlation peak. In the

unmatched case, the correlation peak is broad. Thus its value is small.

One of the useful methods for calculating image correlation is the Fourier transform method. This method is based on a convolution theorem which states that the multiplication of two functions in the frequency domain is equivalent to their convolution in the space domain (Goodman, 1996), so that

$$C(x,y) = \mathcal{F}^{-1}\left\{\mathcal{F}\{t(x,y)\}\mathcal{F}^{*}\{r(x,y)\}\right\}. \qquad (1\text{-}2)$$

In Eq. (1-2), $\mathcal{F}\{\cdot\}$ and $\mathcal{F}^{-1}\{\cdot\}$ represent the forward and the inverse Fourier transform operators, respectively, and * denotes complex conjugation. According to Eq. (1-2), the image correlation can be computed by taking the Fourier transform of the target and the reference images. After the spectra of the target and the reference are multiplied, the resultant product is inversely Fourier transformed in order to obtain the correlation output.

The computation of the image correlation via the Fourier transform method can be implemented by using either an electronic or optical approach. In order to perform the electronic correlation, the Fourier transformation is computed by using a two-dimensional (2-D) discrete Fourier transform. The discrete Fourier transform of the image $r(m,n)$ with $M \times N$ pixels is defined as (Heckbert, www, 1998)

$$R(k,l) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} r(m,n) e^{-j2\pi(\frac{km}{M}+\frac{ln}{N})}. \qquad (1\text{-}3)$$

This equation shows that the 2-D discrete Fourier transform is computed in serial by first performing the 1-D discrete Fourier transform of M rows, and then is followed by another 1-D discrete Fourier transform of N columns. Because the image correlation requires two forward discrete Fourier transforms and one inverse discrete Fourier transform, the electronic approach must perform very intensive computations

and it is slow due to the serial process. This indicates that the correlation-based system requires an extremely fast parallel-computing machine in order to perform the real-time automatic target recognition.

On the other hand, an optical system provides an effective way to perform the 2-D Fourier transform of images at the speed of light. The optical Fourier transform exploits an inherent parallelism of optics and the Fourier transforming ability of a lens to process input images. In order to perform optically the Fourier transform, a spatial pattern or a transparency is placed at the front focal plane of a thin lens. By illuminating the pattern with a collimated coherent light, the Fourier spectrum of the pattern can be obtained at the back focal plane of the lens. On the basis of this optical property, the 2-D correlation can be real-time implemented by using joint transform correlator. In the real-time joint transform correlator (Yu and Lu, 1984), the target image is detected by a charge-coupled device (CCD) image sensor, while the reference images are stored in the computer system. The target and the reference images are displayed onto an electrically addressed spatial light modulator (EASLM) in order to produce the joint input image. Next, the generated joint power spectrum is captured by the CCD. This captured joint power spectrum is transferred to the computer which will then either display the joint power spectrum onto the EASLM in order to perform the second optical Fourier transform or directly perform digital Fourier transform via fast Fourier transform algorithm.

1.2   Significance of Study

Although the joint transform correlator architecture is indeed useful for implementing the automatic target recognition system, there are several limitations that characterize the system such as:

1. The process of displaying the target and the reference images introduces a time delay which is dependent upon the image size.

2. In real-world applications of the automatic target recognition system to personal identification and military surveillance, the system may deal with a large (and still increasing) number of target images to be recognized. As a consequence, the joint transform correlator must have a huge number of reference images which cover all possible variations of the target such as rotation, orientation, and scale changes. Thus, the joint transform correlator requires considerable storage capability which may cause a storage problem.

3. The time delay and the storage problems become even more severe when a CCD sensor having a mega-pixels resolution is employed to detect the target image, because the size of the detected target may exceed several Mbytes.

    One practical approach to solve these problems is to compress the reference images. By applying a digital image compression, such as the one developed by the Joint Photographic Experts Group (JPEG) (Pennebaker and Mitchell, 1993), there are two advantages that can be obtained. First, the time delay occurred during transferring the image from the computer to the EASLM can be reduced. Second, less capacity will be occupied by storing the compressed references. Therefore, the implementation of the automatic target recognition by using the joint transform correlator with the compressed reference images has advantages in that it saves storage requirement and speeds up the detection time. In general, to perform compression, the number of bits needed to represent an image is reduced by discarding spatial and spectral redundancies as much as possible.

Furthermore, although there have been many research works devoted to the joint transform correlator, they were mainly done to improve the recognition performance of the joint transform correlator when its input target suffered from a poor illumination (Alam and Karim, 1993; Jutamulia, Storti, Gregory, and Kirsch, 1991; Pati and Singh, 1998; Zhang and Karim, 1999) and additive Gaussian noise caused by the sensor (Li, Zhang, and Hu, 1996; Li, Yin, and Yu, 1998; Tanone, Uang, Yu, Tam, and Gregory, 1992; Wang, Shang, and Chatwin, 1996). A study of the performance of the joint transform correlator by using compressed reference images is yet to be done. For these reasons, it is important to study effects of the reference compression on the performance of joint transform correlator.

1.3   Research Objectives

There are two primary objectives of this work. The first objective is to study the effects of lossy image compression on the recognition performance of the joint transform correlator, while the second one is to obtain a guideline to optimize the performance of the joint transform correlator.

1.4   Scope and Limitations of the Study

This dissertation studies the implementation of the automatic target recognition system by using the joint transform correlator with the compressed reference images. The study is carried out by both a computer simulation and an experimental verification. The scope and limitations of the study were defined as follows.

- The JPEG compression algorithm is used for compressing the reference images, because it is one of the digital image compression standards and its format is widely supported by most CCD sensors.

- The target image of interest used in the study consists of single and multiple targets. Each target is characterized by different contrast and spatial-frequency content. In addition, the presence of an additive noise in the input target is also taken into account.

- The metrics for quantifying the recognition performance of the joint transform correlator with compressed reference images are a ratio of correlation peak intensity to the standard deviation of the correlation intensity or peak-to-correlation deviation (PCD) and a ratio of the autocorrelation peak intensity to the intensity of the secondary peak or peak-to-secondary peak ratio (PSR). The PCD is used for the case of single-target detection, while for multiple-target detection the PSR is used.

In order to study the effects of the reference image compression on the recognition performance of the joint transform correlator, the research procedure shown as a diagram in Fig. 1.1 is used, where the image of interest is first prepared and duplicated into the target and the reference images. Second, the reference image is compressed into the JPEG format by using the ACDsee software with different compression levels, while the noise-free and noisy images are generated from the target. The target and the compressed reference images are then input into the joint transform correlator. The resultant output of the joint transform correlator is finally quantified.

The joint transform correlator by using compressed reference images is first simulated by using MATLAB and is then experimentally verified. The experiment used in this dissertation is constructed by a conventional optical Fourier transformer using a Helium-Neon (He-Ne) laser as a coherent light source. An EASLM placed in

```
                         ┌─────────────────────┐
                         │   Generate images   │
                         └─────────────────────┘
              ┌────────────────┴──────────────────┐
   ┌─────────────────────┐            ┌─────────────────────┐
   │      Original       │            │      Original        │
   │   target images     │            │   reference images   │
   └─────────────────────┘            └─────────────────────┘
              │    Noise free target                │
   ┌─────────────────────┐       ┌──────────────────────────────┐
   │     Add noise:      │       │ Compression of reference       │
   │    Noisy target     │       │ images into JPEG format        │
   └─────────────────────┘       │ by using ACDSee software with  │
                                 │ different quality factor (QF)  │
                                 └──────────────────────────────┘
              └──────────┐   ┌───────────┐   ┌───────┘
                  ┌──────────────────────────────────┐
                  │ Perform joint transform correlator:│
                  │ •Computer simulation              │
                  │ •Experimental verification        │
                  └──────────────────────────────────┘
                                 │
                  ┌──────────────────────────────────┐
                  │ Quantify the correlation performance│
                  └──────────────────────────────────┘
```
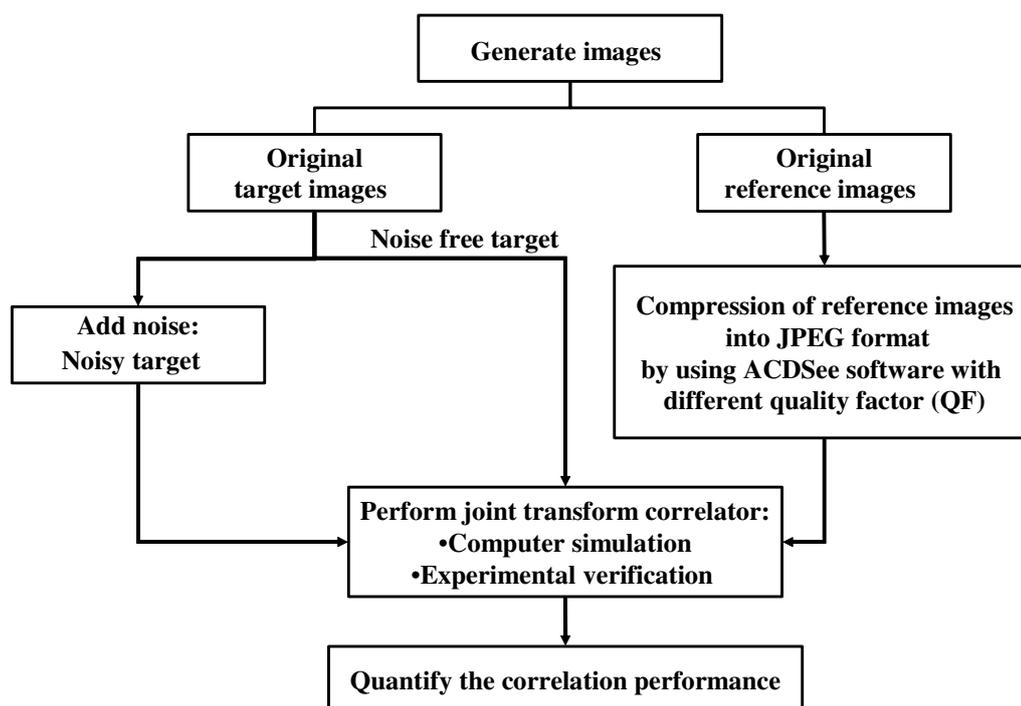
Figure 1.1   Diagram of research procedure.

the front focal-plane of the Fourier transform lens is used to display a joint reference and target images. In order to detect the joint power spectrum and the correlation output intensity, a CCD sensor is placed in the back focal-plane of the lens.

1.5   Organization

This dissertation is organized into six chapters of which this is the first. Chapter II reviews the principles of optical pattern recognition by using the joint transform correlator. The review is started by discussing the basic optical Fourier transform processor and is followed by the mathematical description of the joint transform correlator and its optical implementation. The algorithm of the JPEG image compression and the compression of the reference images of the joint transform correlator are presented in Chapter III. The simulation and the experimental verification of single-target detection by using the joint transform correlator with

compressed reference images are presented in Chapter IV. Chapter V presents the simulation and the experimental results of the multiple-target detection by using joint transform correlator with compressed reference images. Finally, the conclusions of the dissertation are provided in Chapter VI.

CHAPTER II

OPTICAL PATTERN RECOGNITION USING

JOINT TRANSFORM CORRELATOR

In this chapter, the concept of optical pattern recognition by using joint transform correlator is presented. The discussion is started with reviewing principles of optical Fourier transformation by using a thin lens which forms the basis for implementing the correlation-based pattern recognition. This is followed by discussion of the theory of a joint transform correlator and its real-time implementation.

2.1   Introduction

Optical pattern recognition by correlation is one of the most successful applications of coherent optical processing systems and currently remains an active area of research. The reason of this is that the correlation of two functions, reference and target images, can be practically computed by taking an inverse Fourier transform of the product of the spectra of the two functions. Since the coherent optical system can perform Fourier transformations with the speed of light, optical correlators are very useful for real-time image classifications.

In the field of optical pattern recognition, the VanderLugt (Lugt, 1964) and the joint transform correlators are two optical architectures widely used for implementing correlation operations. The VanderLugt correlator is the first optical architecture which uses a complex-valued transfer function of the reference image to detect the

input target image. The transfer function is interferometrically synthesized as a spatial filter. However, due to requirement for prior filter synthesis, the VanderLugt architecture is not suitable for real-time systems. This drawback is not found in the joint transform correlator (Goodman, 1996), in particular when the joint transform corrleator is implemented in conjunction with an electrically addressed spatial light modulator.

2.2　The Optical Fourier Transform

The optical Fourier transform is based on the laws of diffraction and propagation of light. In the optical Fourier transform, a far-field diffraction (i.e., Fraunhofer diffraction) pattern of coherently illuminated spatial mask is mapped into the near-field diffraction by using a positive lens. This can be regarded as taking the 2-D Fourier transform of the spatial mask by using the lens. Figure 2.1 shows a schematic diagram for implementing the optical Fourier transform of the spatial pattern with amplitude transmittance $u(x_1, y_1)$. By illuminating perpendicularly the
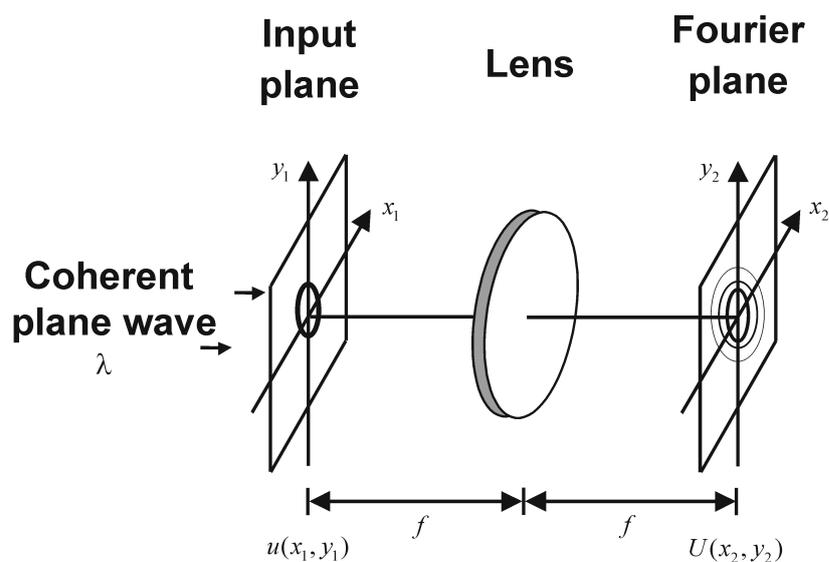
Figure 2.1　Optical Fourier transform by a positive lens.

pattern placed at one focal length f in front of the Fourier transform lens with a coherent plane wave of unity amplitude, the amplitude distribution appearing at one focal length behind the lens is found to be (Goodman, 1996)

$$U(x_2, y_2) = \frac{1}{j\lambda f} \int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} u(x_1, y_1) \exp\left\{-j\frac{2\pi}{\lambda f}(x_1 x_2 + y_1 y_2)\right\} dx_1 dy_1,$$ (2-1)

where λ is the wavelength of the coherent light, and (x_2, y_2) are the actual coordinates in the horizontal and the vertical directions at the back focal plane which is also called as the Fourier plane.

2.3 The Joint Transform Correlator

The basic operation of the joint transform correlator consists of two of step process: Firstly, the generation of the joint power spectrum of the target and the reference images by using the optical Fourier transform shown in Fig. 2.2(a). In order to correlate the target $t(x_1, y_1)$ and the reference $r(x_1, y_1)$, they are placed side-by-side on the input plane. This joint input image can be mathematically expressed as

$$f_{JTC}(x_1, y_1) = r(x_1 - x_0, y_1) + t(x_1 + x_0, y_1),$$ (2-2)

where x_0 and $-x_0$ are the position of the reference and the target images at the input plane (x_1, y_1), respectively. By illuminating perpendicularly this input image with a coherent plane wave, the Fourier transform of the joint input image generated at the Fourier plane becomes

$$F_{JTC}(x_2, y_2) = \frac{1}{j\lambda f} R\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) \exp\left(-j2\pi x_0 \frac{x_2}{\lambda f}\right)$$
$$+ \frac{1}{j\lambda f} T\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) \exp\left(j2\pi x_0 \frac{x_2}{\lambda f}\right),$$ (2-3)

where $R\left(\dfrac{x_2}{\lambda f}, \dfrac{y_2}{\lambda f}\right)$ and $T\left(\dfrac{x_2}{\lambda f}, \dfrac{y_2}{\lambda f}\right)$ correspond to the Fourier spectrum of the

reference and the target images, respectively. The exponential terms in Eq. (2-3) are

(a)

(b)
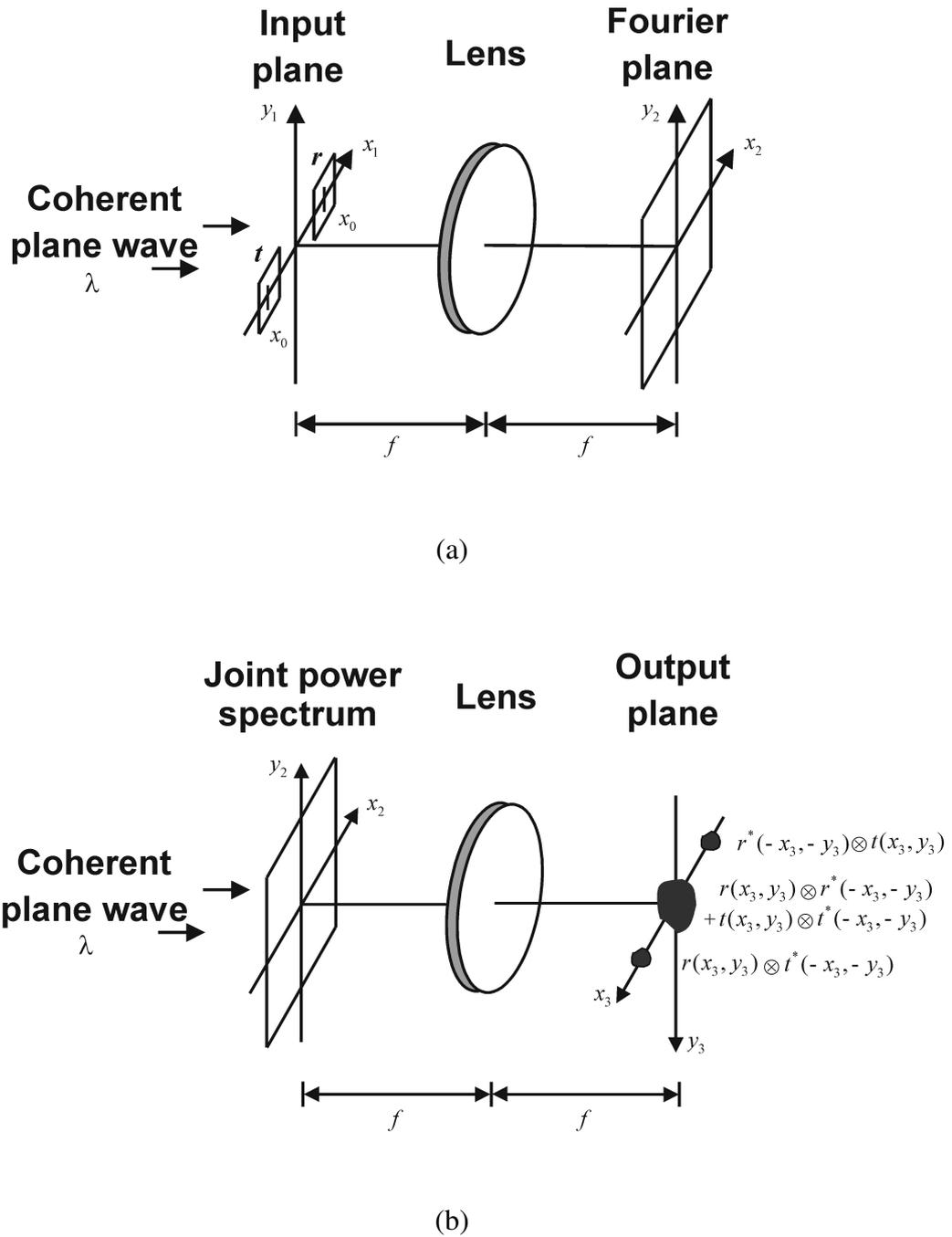
Figure 2.2   Schematic diagram of joint transform correlator (a) recording of the

joint power spectrum, (b) obtaining the correlation output.

caused by the translation of the images in the input plane. By recording the intensity of this complex field distribution on a photographic film or other light-sensitive recording media, its resultant joint power spectrum can be mathematically described by

$$
\begin{aligned}
I(x_2, y_2) &= \left| F_{JTC}(x_2, y_2) \right|^2 \\
&= \frac{1}{\lambda^2 f^2} \left[\left| R\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f} \right) \right|^2 + \left| T\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f} \right) \right|^2 \right. \\
&\quad + R\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f} \right) T^*\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f} \right) \exp\left(-j4\pi x_0 \frac{x_2}{\lambda f} \right) \\
&\quad \left. + R^*\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f} \right) T\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f} \right) \exp\left(+j4\pi x_0 \frac{x_2}{\lambda f} \right) \right].
\end{aligned}
\tag{2-4}
$$

The last two terms in Eq. (2-4) correspond to the multiplications of the Fourier spectra of the target and the reference images that are $R\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f} \right) T^*\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f} \right)$ and $R^*\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f} \right) T\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f} \right)$. These terms are of particular interest, because they provide the desired correlation between the target and the reference images.

The second process is the generation of the cross-correlation between the target and the reference images by Fourier transforming optically the joint power spectrum. Once the film has been developed, the transparency with an amplitude transmittance that is proportional to the recorded joint power spectrum is inserted in the input plane of the second optical Fourier transform shown in Fig. 2.2 (b). As a result, the field distribution generated at the output plane is proportional to

$$
\begin{aligned}
U_3(x_3, y_3) &= \frac{1}{\lambda f} \left[r(x_3, y_3) \otimes r^*(-x_3, -y_3) + t(x_3, y_3) \otimes t^*(-x_3, -y_3) \right. \\
&\quad + r(x_3, y_3) \otimes t^*(-x_3, -y_3) \otimes \delta(x_3 - 2x_0, y_3)
\end{aligned}
$$

$$+ r^*(-x_3, -y_3) \otimes t(x_3, y_3) \otimes \delta(x_3 + 2x_0, y_3) \Big], \qquad (2\text{-}5)$$

where $\otimes$ denotes convolution. The first two terms of Eq. (2-5) correspond to the autocorrelations of the reference and the target images, respectively. They appear on axis of the correlation plane (x_3, y_3). The third and the fourth terms produce the desired cross-correlation signals of the reference $r(x_1, y_1)$ and the target $t(x_1, y_1)$ centered at $x_3 = 2x_0$ and $x_3 = -2x_0$. By measuring the cross-correlation height and its sharpness, the degree of similarity between the target and the reference is finally determined.

2.4   The Real-Time Joint Transform Correlator

With rapid development of detection technology and electro-optics devices, a real-time implementation of the joint transform correlator for adaptive pattern recognition by using a magneto-optic device was firstly proposed by Yu and his coworkers (Yu and Lu, 1984). Since then, several proposals of the real-time joint transform correlator by using liquid crystal television and EASLM have also been reported (Yu, Jutamulia, Lin, and Gregory, 1987). Figure 2.3 shows a schematic diagram of optical setup for implementing the real-time joint transform correlator which is used in this thesis. In this setup, the target image is captured by a CCD image sensor, while the reference images are stored in a computer system. In order to perform correlation, the two images are displayed onto an EASLM placed in a front focal plane of a Fourier transforming lens. By illuminating perpendicularly the EASLM with collimated laser light, the generated joint power spectrum at the back focal is captured by the CCD sensor and is then transferred to the computer. By redisplaying the recorded joint power spectrum onto the EASLM and after subsequent

Fourier transformation, the correlation output can be obtained at the back focal plane of the lens L_1.
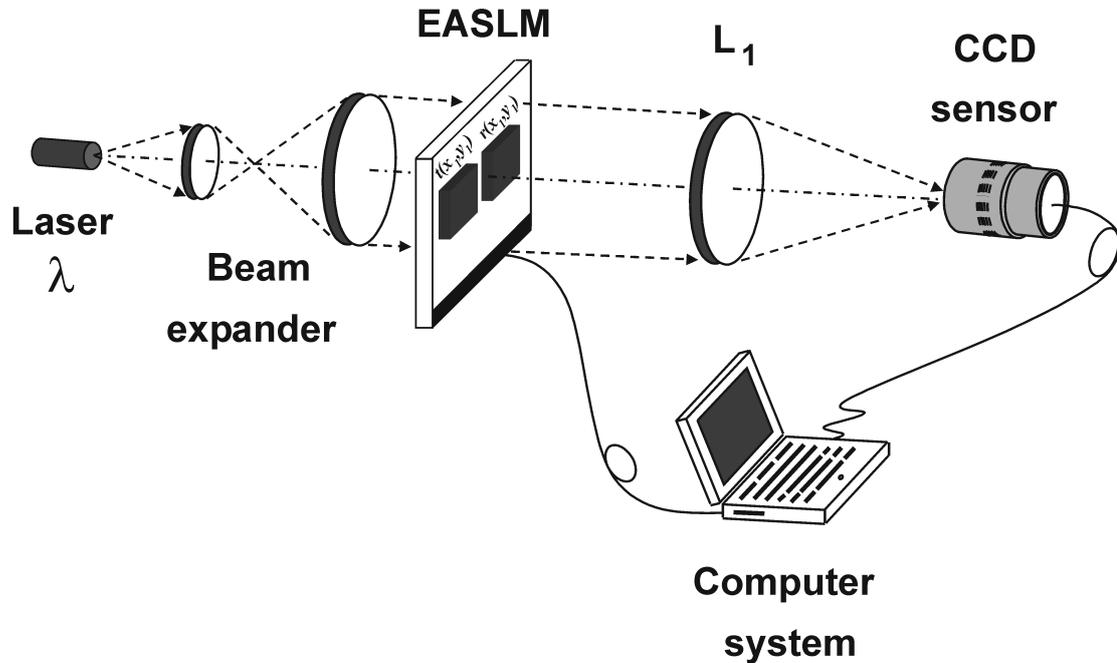
Figure 2.3   A schematic diagram of optical setup for implementing real-time joint transform correlator.

2.4.1   The Optical Fourier Transform of Pixelated EASLM

As discussed in the preceding Section, the real-time joint transform correlator employs the EASLM as a programmable display of the reference and the target images. However, most EASLMs comprise of a matrix of light-modulating elements with rectangular shape as shown in Fig. 2.4. Therefore, unlike photographic films, EASLMs are discrete pixelated devices. They display sampled images rather than continuous images. As for the EASLM with resolution of $2M_{ex} \times 2M_{ey}$ pixels and pixel size of $L_{ex} \times L_{ey}$, its amplitude transmittance can be mathematically written as
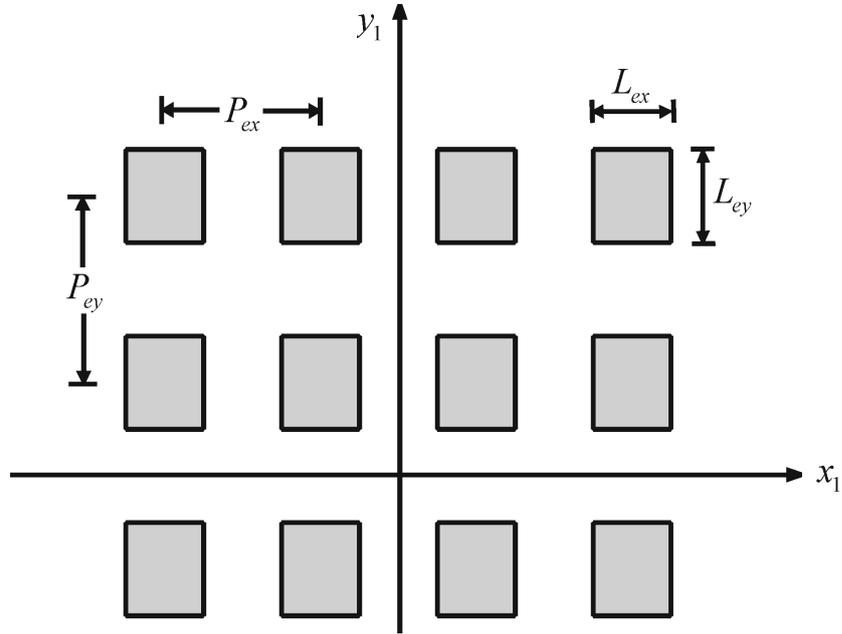
Figure 2.4 Geometry of the EASLM with rectangular pixels.

$$f_{EASLM}(x_1, y_1) = \sum_{n=-M_{ey}}^{M_{ey}} \sum_{m=-M_{ex}}^{M_{ex}} \delta\left(x_1 - mP_{ex}, y_1 - nP_{ey}\right) \otimes \text{rect}\left(\frac{x_1}{L_{ex}}\right) \text{rect}\left(\frac{y_1}{L_{ey}}\right), \quad (2\text{-}6)$$

where P_{ex} and P_{ey} are the pixel pitch in the x_1 and the y_1 directions, respectively. Since the EASLM is used to display the input image of the optical Fourier transformer, it is important to understand the sampling effect of the EASLM on the generated Fourier spectrum of the image. Let us consider rectangular aperture with size of $N_x P_{ex} \times N_y P_{ey}$ as

$$f(x_1, y_1) = \text{rect}\left(\frac{x_1}{N_x P_{ex}}\right) \text{rect}\left(\frac{y_1}{N_y P_{ey}}\right) \quad (2\text{-}7)$$

is displayed on the center of the EASLM. The sampled rectangular aperture can be mathematically expressed as

$$f_s(x_1, y_1) = f_{EASLM}(x_1, y_1) f(x_1, y_1)$$

$$= \left[\sum_{n=-M_{ey}}^{M_{ey}} \sum_{m=-M_{ex}}^{M_{ex}} \delta(x_1 - mP_{ex}, y_1 - nP_{ey}) \otimes \text{rect}\left(\frac{x_1}{L_{ex}}\right) \text{rect}\left(\frac{y_1}{L_{ey}}\right) \right]$$

$$\times \text{rect}\left(\frac{x_1}{N_x P_{ex}}\right) \text{rect}\left(\frac{y_1}{N_y P_{ey}}\right). \tag{2-8}$$

By Fourier transforming optically the sampled aperture, its Fourier spectrum is found to be

$$F_s(x_2, y_2) = \frac{1}{j\lambda f} \left[F(x_2, y_2) \otimes F_{EASLM}(x_2, y_2) \right], \tag{2-9}$$

where $F(x_2, y_2)$ is the Fourier spectrum of the aperture $f(x_1, y_1)$ given by

$$F(x_2, y_2) = N_x N_y P_{ex} P_{ey} \, \text{sinc}\left(\frac{N_x P_{ex}}{\lambda f} x_2\right) \text{sinc}\left(\frac{N_y P_{ey}}{\lambda f} y_2\right), \tag{2-10}$$

and

$$F_{EASLM}(x_2, y_2) = \frac{L_{ex} L_{ey}}{P_{ex} P_{ey}} \text{sinc}\left(\frac{L_{ex}}{\lambda f} x_2\right) \text{sinc}\left(\frac{L_{ey}}{\lambda f} y_2\right)$$

$$\times \sum_{n=-M_{ey}}^{M_{ey}} \sum_{m=-M_{ex}}^{M_{ex}} \delta\left(\frac{x_2}{\lambda f} - \frac{m}{P_{ex}}, \frac{y_2}{\lambda f} - \frac{n}{P_{ey}}\right). \tag{2-11}$$

The field intensity at the back focal plane of the Fourier transform lens can be written as

$$\left| F_s(x_2, y_2) \right|^2 = \left| \frac{1}{j\lambda f} N_x N_y L_{ex} L_{ey} \text{sinc}\left(\frac{L_{ex}}{\lambda f} x_2\right) \text{sinc}\left(\frac{L_{ey}}{\lambda f} y_2\right) \right.$$

$$\times \sum_{m=-M_{ey}}^{M_{ey}} \sum_{n=-M_{ex}}^{M_{ex}} \delta\left(\frac{x_2}{\lambda f} - \frac{m}{P_{ex}}, \frac{y_2}{\lambda f} - \frac{n}{P_{ey}}\right)$$

$$\left. \otimes \text{sinc}\left(\frac{N_x P_{ex}}{\lambda f} x_2\right) \text{sinc}\left(\frac{N_y P_{ey}}{\lambda f} y_2\right) \right|^2. \tag{2-12}$$

Figure 2.5 shows the intensity pattern along the x_2 axis for $N_x = N_y = 20$ pixels. The first two sinc functions correspond to the Fourier transform of the rectangular pixel.

Since the pixel size is the smallest structure in the input plane, this 2-D sinc function will be the broadest signal in the Fourier plane. It can be regarded as an envelope function with a main lobe of $2\dfrac{\lambda f}{L_{ex}}$ wide. This envelope function modulates a convolution of an array of delta functions, spaced at interval of $\dfrac{\lambda f}{P_{ex}}$ and $\dfrac{\lambda f}{P_{ey}}$ in the x_2 and the y_2 directions, with the Fourier spectrum of the rectangular aperture. Therefore, although the Fourier spectrum of the sampled rectangular aperture appears as the replication of the original spectrum of the aperture at $\left(\dfrac{m\lambda f}{P_{ex}},\dfrac{n\lambda f}{P_{ey}}\right)$ in the Fourier plane, the number of replications is limited by the envelope function.
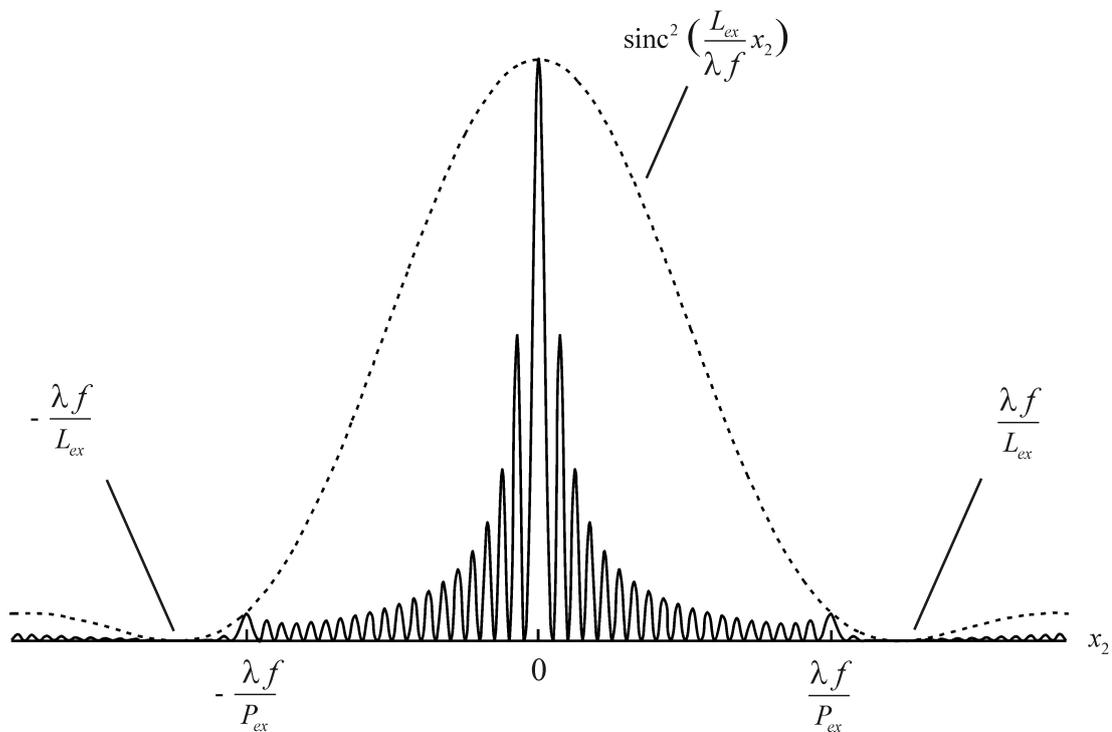
Figure 2.5   1-D cross-sectional scan of the power spectrum of the rectangular aperture displayed on the EASLM.

2.4.2   The Joint Transform Correlator Using a Pixelated EASLM

Consider the joint input image of the joint transform correlator to be displayed on the pixelated EASLM as shown in Fig. 2.6. Let the reference $r(x_1, y_1)$ and the target $t(x_1, y_1)$ images having $N_x \times N_y$ pixels are separated with M pixels from each other. The amplitude transmittance of the joint input image to be displayed on the EASLM is mathematically described as

$$f_{JTC}(x_1, y_1) = t\left(x_1 + \frac{MP_{ex}}{2}, y_1\right) \text{rect}\left(\frac{x_1 + MP_{ex}/2}{N_x P_{ex}}\right) \text{rect}\left(\frac{y_1}{N_y P_{ey}}\right)$$

$$+ r\left(x_1 - \frac{MP_{ex}}{2}, y_1\right) \text{rect}\left(\frac{x_1 - MP_{ex}/2}{N_x P_{ex}}\right) \text{rect}\left(\frac{y_1}{N_y P_{ey}}\right). \qquad (2\text{-}13)$$

By displaying the joint input image onto the EASLM, the sampled image becomes
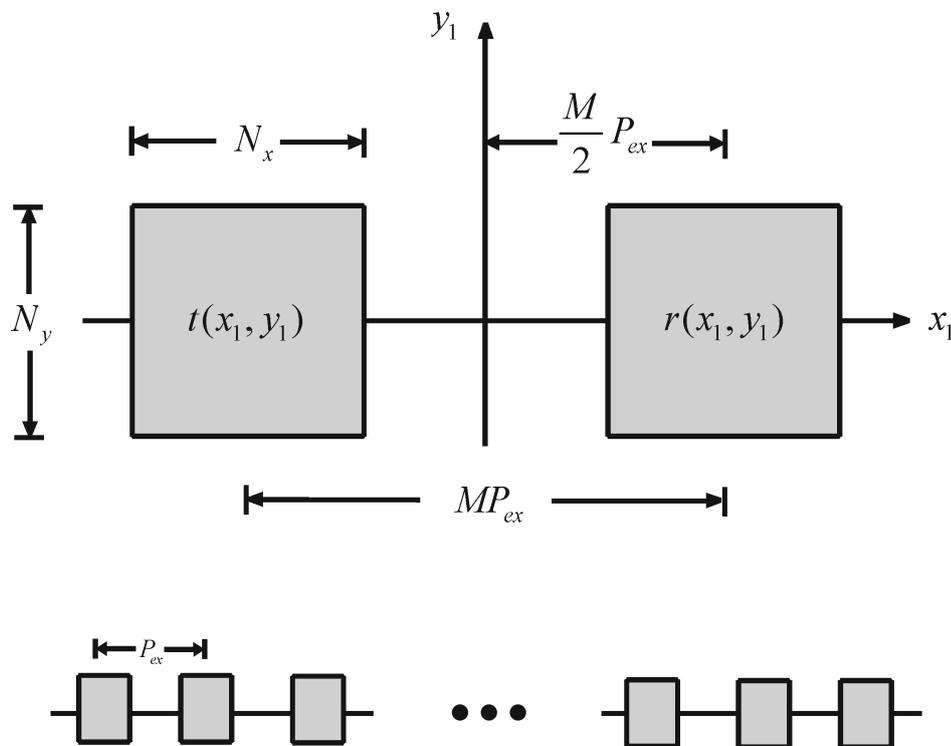
Figure 2.6   Joint input image.

$$f_s(x_1, y_1) = f_{EASLM}(x_1, y_1) f_{JTC}(x_1, y_1)$$

$$= \sum_{m=-M_{ey}}^{M_{ey}} \sum_{n=-M_{ex}}^{M_{ex}} \delta(x_1 - mP_{ex}, y_1 - nP_{ey}) \otimes \text{rect}\left(\frac{x_1}{L_{ex}}\right) \text{rect}\left(\frac{y_1}{L_{ey}}\right)$$

$$\times \left[t\left(x_1 + \frac{MP_{ex}}{2}, y_1\right) \text{rect}\left(\frac{x_1 + MP_{ex}/2}{N_x P_{ex}}\right) \text{rect}\left(\frac{y_1}{N_y P_{ey}}\right) \right.$$

$$\left. + r\left(x_1 - \frac{MP_{ex}}{2}, y_1\right) \text{rect}\left(\frac{x_1 - MP_{ex}/2}{N_x P_{ex}}\right) \text{rect}\left(\frac{y_1}{N_y P_{ey}}\right) \right]. \qquad (2\text{-}14)$$

Perpendicular illumination of the EASLM by a coherent plane wave produces the Fourier spectrum of the sampled joint input image at the Fourier plane as

$$F_s(x_2, y_2) = \frac{1}{j\lambda f} \frac{L_{ex} L_{ey}}{P_{ex} P_{ey}} \left[\text{sinc}\left(\frac{L_{ex}}{\lambda f} x_2\right) \text{sinc}\left(\frac{L_{ex}}{\lambda f} y_2\right) \right.$$

$$\left. \times \sum_{n=-M_{ey}}^{M_{ey}} \sum_{m=-M_{ex}}^{M_{ex}} \delta\left(\frac{x_2}{\lambda f} - \frac{m}{P_{ex}}, \frac{y_2}{\lambda f} - \frac{n}{P_{ey}}\right) \otimes F_{JTC}\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) \right] \qquad (2\text{-}15)$$

where the Fourier spectrum of the

$$F_{JTC}\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) = T'\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) \exp\left(j2\pi \frac{MP_{ex}}{2\lambda f} x_2\right)$$

$$+ R'\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) \exp\left(-j2\pi \frac{MP_{ex}}{2\lambda f} x_2\right), \qquad (2\text{-}16)$$

with

$$T'\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) = T\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) \otimes N_x N_y P_{ex} P_{ey} \text{sinc}\left(\frac{N_x P_{ex}}{\lambda f} x_2\right) \text{sinc}\left(\frac{N_y P_{ey}}{\lambda f} y_2\right) \quad (2\text{-}17a)$$

and

$$R'\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) = R\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) \otimes N_x N_y P_{ex} P_{ey} \text{sinc}\left(\frac{N_x P_{ex}}{\lambda f} x_2\right) \text{sinc}\left(\frac{N_y P_{ey}}{\lambda f} y_2\right). (2\text{-}17b)$$

By a straight forward calculation, the joint power spectrum of the joint transform correlator system is given by

$$|F_s(x_2, y_2)|^2 = \frac{1}{\lambda^2 f^2}\left(\frac{L_{ex}L_{ey}}{P_{ex}P_{ey}}\right)^2 \sum_{m=-M_{ey}}^{M_{ey}} \sum_{n=-M_{ex}}^{M_{ex}} \text{sinc}^2\left(\frac{mL_{ex}}{P_{ex}}\right) \text{sinc}^2\left(\frac{nL_{ey}}{P_{ey}}\right)$$

$$\times \left[\left|R'\left(\frac{x_2}{\lambda f}-\frac{m}{P_{ex}}, \frac{y_2}{\lambda f}-\frac{n}{P_{ey}}\right)\right|^2 + \left|T'\left(\frac{x_2}{\lambda f}-\frac{m}{P_{ex}}, \frac{y_2}{\lambda f}-\frac{n}{P_{ey}}\right)\right|^2\right.$$

$$+ R'\left(\frac{x_2}{\lambda f}-\frac{m}{P_{ex}}, \frac{y_2}{\lambda f}-\frac{n}{P_{ey}}\right)T'^*\left(\frac{x_2}{\lambda f}-\frac{m}{P_{ex}}, \frac{y_2}{\lambda f}-\frac{n}{P_{ey}}\right)e^{-j4\pi\frac{MP_{ex}}{2}\left(\frac{x_2}{\lambda f}-\frac{m}{P_{ex}}\right)}$$

$$+ R'^*\left(\frac{x_2}{\lambda f}-\frac{m}{P_{ex}}, \frac{y_2}{\lambda f}-\frac{n}{P_{ey}}\right)T'\left(\frac{x_2}{\lambda f}-\frac{m}{P_{ex}}, \frac{y_2}{\lambda f}-\frac{n}{P_{ey}}\right)$$

$$\left.\times e^{+j4\pi\frac{MP_{ex}}{2}\left(\frac{x_2}{\lambda f}-\frac{m}{P_{ex}}\right)}\right]. \tag{2-18}$$

Equation (2-18) shows that the pixel structure of the EASLM produces multiple power spectra of the joint input image $F_{JTC}(x_2, y_2)$ in the Fourier plane. However because of the modulation by the envelope function, only the zero order spectrum $(m=0, n=0)$ of the joint power spectrum $|F(x_2, y_2)|^2$ has the highest intensity, while the higher orders are attenuated. Since its intensity is the highest, the zero order spectrum is recorded by the CCD sensor for the next computation purpose.

The zero order spectrum can be mathematically written as

$$F_{JTC_0}(x_2, y_2) = \left|R'\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right)\right|^2 + \left|T'\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right)\right|^2$$

$$+ R'\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right)T'^*\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right)\exp\left(-j4\pi\frac{MP_{ex}}{2}\frac{x_2}{\lambda f}\right)$$

$$+ R'^*\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right)T'\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right)\exp\left(+j4\pi\frac{MP_{ex}}{2}\frac{x_2}{\lambda f}\right). \tag{2-19}$$

By expressing the complex field distribution $R'\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right)$ and $T'\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right)$ into

their amplitude and phase distributions

$$R'\left(\frac{x_2}{\lambda f},\frac{y_2}{\lambda f}\right)=\left|R'\left(\frac{x_2}{\lambda f},\frac{y_2}{\lambda f}\right)\right|\exp\left\{j\phi_R\left(\frac{x_2}{\lambda f},\frac{y_2}{\lambda f}\right)\right\}$$

and

$$T'\left(\frac{x_2}{\lambda f},\frac{y_2}{\lambda f}\right)=\left|T'\left(\frac{x_2}{\lambda f},\frac{y_2}{\lambda f}\right)\right|\exp\left\{j\phi_T\left(\frac{x_2}{\lambda f},\frac{y_2}{\lambda f}\right)\right\},$$

the zero order of the joint power spectrum can be rewritten as

$$F_{JTC_0}\left(\frac{x_2}{\lambda f},\frac{y_2}{\lambda f}\right)=\left|R'\left(\frac{x_2}{\lambda f},\frac{y_2}{\lambda f}\right)\right|^2+\left|T'\left(\frac{x_2}{\lambda f},\frac{y_2}{\lambda f}\right)\right|^2+2\left|R'\left(\frac{x_2}{\lambda f},\frac{y_2}{\lambda f}\right)\right|\left|T'\left(\frac{x_2}{\lambda f},\frac{y_2}{\lambda f}\right)\right|$$

$$\times\cos\left\{2\pi MP_{ex}\frac{x_2}{\lambda f}+\phi_R\left(\frac{x_2}{\lambda f},\frac{y_2}{\lambda f}\right)-\phi_T\left(\frac{x_2}{\lambda f},\frac{y_2}{\lambda f}\right)\right\}. \qquad (2\text{-}20)$$

Equation (2-20) shows that the third term contains the desired product of the Fourier spectra of the target and the reference images which is sampled by the cosine carrier fringes. Since the product of the spectra determines the correlation output, the cosine fringes must be faithfully recorded.

By recording the continuous joint power spectrum using the CCD sensor with the spatial resolution of $2N_{CCD_x}\times2N_{CCD_y}$ and pitch of P_{CCD_x} and P_{CCD_y}, the sampled zero order of the joint power spectrum can be mathematically described as

$$F'_{JTC_0}(x_2,y_2)=F_{JTC_0}(x_2,y_2)f_{CCD}(x_2,y_2)$$

$$=\left\{\left|R'\left(\frac{x_2}{\lambda f},\frac{y_2}{\lambda f}\right)\right|^2+\left|T'\left(\frac{x_2}{\lambda f},\frac{y_2}{\lambda f}\right)\right|^2+2\left|R'\left(\frac{x_2}{\lambda f},\frac{y_2}{\lambda f}\right)\right|\left|T'\left(\frac{x_2}{\lambda f},\frac{y_2}{\lambda f}\right)\right|\right.$$

$$\left.\times\cos\left\{2\pi MP_{ex}\frac{x_2}{\lambda f}+\phi_R\left(\frac{x_2}{\lambda f},\frac{y_2}{\lambda f}\right)-\phi_T\left(\frac{x_2}{\lambda f},\frac{y_2}{\lambda f}\right)\right\}\right\}$$

$$\times\sum_{n=-N_{CCD_y}}^{N_{CCD_y}}\sum_{m=-N_{CCD_x}}^{N_{CCD_x}}\delta\left(x_2-mP_{CCDx},y_2-mP_{CCDy}\right)$$

$$\otimes\mathrm{rect}\left(\frac{x_2}{L_{cx}}\right)\mathrm{rect}\left(\frac{y_2}{L_{cy}}\right), \qquad (2\text{-}21)$$

where $L_{cx}\times L_{cy}$ represents the size of the light-detecting rectangular element of the

CCD. Since the CCD sensor used for recording the joint power spectrum is generally characterized by both finite resolution and finite active area, in order to record the zero order spectra the CCD sensor must fulfill two requirements. First, its resolution must satisfy the sampling theorem (Oppenheim and Schafer, 1989) in order to sample the detail of the spectra that is the cosine function of Eq. (2-20). Second, its active area must be wider than the size of the zero order spectra. The first requirement determines the relationship between the sampling frequency of the CCD with the frequency of the cosine function such that

$$f_{CCD} \geq 2 f_{\text{cosine}} \tag{2-22a}$$

or

$$\frac{\lambda f}{M P_{ex}} \geq 2 P_{CCD_x} . \tag{2-22b}$$

Equation (2-22) indicates that the spatial separation of the target and the reference images in the input plane is determined by not only the focal length of the Fourier transform lens and the operating wavelength but also the spatial resolution P_{CCD_x} of the CCD sensor. The second requirement can be mathematically expressed as

$$2 N_{CCDx} P_{CCDx} \geq \frac{\lambda f}{P_{ex}} \tag{2-23a}$$

$$2 N_{CCDy} P_{CCDy} \geq \frac{\lambda f}{P_{ey}} . \tag{2-23b}$$

Next, the digitized joint power spectrum is redisplayed on the EASLM in order to produce correlation output. However, since the pixel pitch and size of the CCD sensor are different than that of the EASLM, the spatial extension of the redisplayed joint power spectrum are scaled by the ratio of the pixel pitch given by

$$R_{px} = \frac{P_{CCDx}}{P_{ex}} \tag{2-24a}$$

$$R_{py} = \frac{P_{CCDy}}{P_{ey}} \tag{2-24b}$$

and the ratio of the pixel size given by

$$R_{Lx} = \frac{L_{cx}}{L_{ex}} \tag{2-25a}$$

$$R_{Ly} = \frac{L_{cy}}{L_{ey}}, \tag{2-25b}$$

respectively. By taking the scaling factors into account, the amplitude transmittance of the redisplayed joint power spectrum becomes

$$
\begin{aligned}
F'_{JTC_0}(x_2, y_2) = & \left\{ \left| R'\left(R_{px}\frac{x_2}{\lambda f}, R_{py}\frac{y_2}{\lambda f} \right) \right|^2 + \left| T'\left(R_{px}\frac{x_2}{\lambda f}, R_{py}\frac{y_2}{\lambda f} \right) \right|^2 \right. \\
& + R'\left(R_{px}\frac{x_2}{\lambda f}, R_{py}\frac{y_2}{\lambda f} \right) T'^*\left(R_{px}\frac{x_2}{\lambda f}, R_{py}\frac{y_2}{\lambda f} \right) \\
& \times \exp\left(-j4\pi R_{px}\frac{MP_{ex}}{2}\frac{x_2}{\lambda f} \right) \\
& + R'^*\left(R_{px}\frac{x_2}{\lambda f}, R_{py}\frac{y_2}{\lambda f} \right) T'\left(R_{px}\frac{x_2}{\lambda f}, R_{py}\frac{y_2}{\lambda f} \right) \\
& \left. \times \exp\left(+j4\pi R_{px}\frac{MP_{ex}}{2}\frac{x_2}{\lambda f} \right) \right\} \\
& \times \sum_{n=-M_{ey}}^{M_{ey}} \sum_{m=-M_{ex}}^{M_{ex}} \delta\left(R_{px}x_2 - mP_{CCDx}, R_{py}y_2 - mP_{CCDy} \right) \\
& \otimes \mathrm{rect}\left(R_{Lx}\frac{x_2}{L_{cx}} \right) \mathrm{rect}\left(R_{Ly}\frac{y_2}{L_{cy}} \right).
\end{aligned}
\tag{2-26}
$$

By Fourier transforming the redisplayed joint power spectrum using the lens L$_1$, the correlation output intensity is mathematically found to be

$$U(x_3, y_3) = \frac{1}{\lambda^2 f^2} \left| c(x_3, y_3) \otimes \frac{L_{ex} L_{ey}}{P_{CCDx} P_{CCDy}} \operatorname{sinc}\left(\frac{L_{ex}}{\lambda f} x_3\right) \operatorname{sinc}\left(\frac{L_{ey}}{\lambda f} y_3\right) \right.$$

$$\left. \times \sum_{m=-M_{ey}}^{M_{ey}} \sum_{n=-M_{ex}}^{M_{ex}} \delta\left(\frac{x_3}{\lambda f} - \frac{m}{P_{ex}}, \frac{y_3}{\lambda f} - \frac{n}{P_{ey}}\right) \right|^2, \tag{2-27}$$

where $c(x_3, y_3)$ is the correlation signal defined as

$$c(x_3, y_3) = \frac{P_{ex}^2 P_{ey}^2}{P_{CCDx}^2 P_{CCDy}^2} \left\{ \left[r\left(\frac{P_{ex}}{P_{CCDx}} x_3, \frac{P_{ey}}{P_{CCDy}} y_3\right) \operatorname{rect}\left(\frac{x_3}{N_x P_{CCDx}}\right) \operatorname{rect}\left(\frac{y_3}{N_y P_{CCDy}}\right) \right.\right.$$

$$\otimes r^*\left(\frac{-P_{ex}}{P_{CCDx}} x_3, \frac{-P_{ey}}{P_{CCDy}} y_3\right) \operatorname{rect}\left(\frac{x_3}{N_x P_{CCDx}}\right) \operatorname{rect}\left(\frac{y_3}{N_y P_{CCDy}}\right) \bigg]$$

$$+ \left[t\left(\frac{P_{ex}}{P_{CCDx}} x_3, \frac{P_{ey}}{P_{CCDy}} y_3\right) \operatorname{rect}\left(\frac{x_3}{N_x P_{CCDx}}\right) \operatorname{rect}\left(\frac{y_3}{N_y P_{CCDy}}\right) \right.$$

$$\otimes t^*\left(\frac{-P_{ex}}{P_{CCDx}} x_3, \frac{-P_{ey}}{P_{CCDy}} y_3\right) \operatorname{rect}\left(\frac{x_3}{N_x P_{CCDx}}\right) \operatorname{rect}\left(\frac{y_3}{N_y P_{CCDy}}\right) \bigg]$$

$$+ \left[r\left(\frac{P_{ex}}{P_{CCDx}} x_3, \frac{P_{ey}}{P_{CCDy}} y_3\right) \operatorname{rect}\left(\frac{x_3}{N_x P_{CCDx}}\right) \operatorname{rect}\left(\frac{y_3}{N_y P_{CCDy}}\right) \right.$$

$$\otimes t^*\left(\frac{-P_{ex}}{P_{CCDx}} x_3, \frac{-P_{ey}}{P_{CCDy}} y_3\right) \operatorname{rect}\left(\frac{x_3}{N_x P_{CCDx}}\right) \operatorname{rect}\left(\frac{y_3}{N_y P_{CCDy}}\right) \bigg]$$

$$\otimes \delta\left(x_3 - MP_{CCDx}, y_3\right)$$

$$+ \left[r^*\left(\frac{-P_{ex}}{P_{CCDx}} x_3, \frac{-P_{ey}}{P_{CCDy_2}} y_3\right) \operatorname{rect}\left(\frac{x_3}{N_x P_{CCDx}}\right) \operatorname{rect}\left(\frac{y_3}{N_y P_{CCDy}}\right) \right.$$

$$\otimes t\left(\frac{P_{ex}}{P_{CCDx}} x_3, \frac{P_{ey}}{P_{CCDy}} y_3\right) \operatorname{rect}\left(\frac{x_3}{N_x P_{CCDx}}\right) \operatorname{rect}\left(\frac{y_3}{N_y P_{CCDy}}\right) \bigg]$$

$$\otimes \delta\left(x_3 + MP_{CCDx}, y_3\right) \Bigg\}. \tag{2-28}$$

The first two terms of Eq. (2-28) correspond to the autocorrelations of the reference

and the target images appeared on the origin of the correlation output plane. The last

two terms are the desired cross-correlation signals centered at coordinates $(MP_{CCD_x}, 0)$

and $(-MP_{CCD_x}, 0)$, respectively. The maximum width of each correlation signal is confined in the area of $2N_x P_{CCDx} \times 2N_y P_{CCDy}$. Thus, the cross-correlation and the autocorrelation signal can be completely separated provided $M \geq 2N_x$. However due to the pixel structure of the EASLM, the correlation output is periodically replicated. According to Eq. (2-27), the separation of the replicated correlation output in the x and the y directions are $\dfrac{\lambda f}{p_{ex}}$ and $\dfrac{\lambda f}{P_{ey}}$, respectively. Therefore, these separation must be chosen sufficiently wide in order to prevent overlapping of the adjacent replicas. Figure 2.7 shows the cross-sectional scan of the replicated correlation output along the $y_3 = 0$. It is clear that the separation of each correlation signal depends on the characteristic of the EASLM and the CCD sensor, the focal length of the Fourier transform lens and the wavelength of the laser light. It can be seen from Fig. 2.7 that in order to avoid the overlapping of the adjacent replicas, their separation must follow
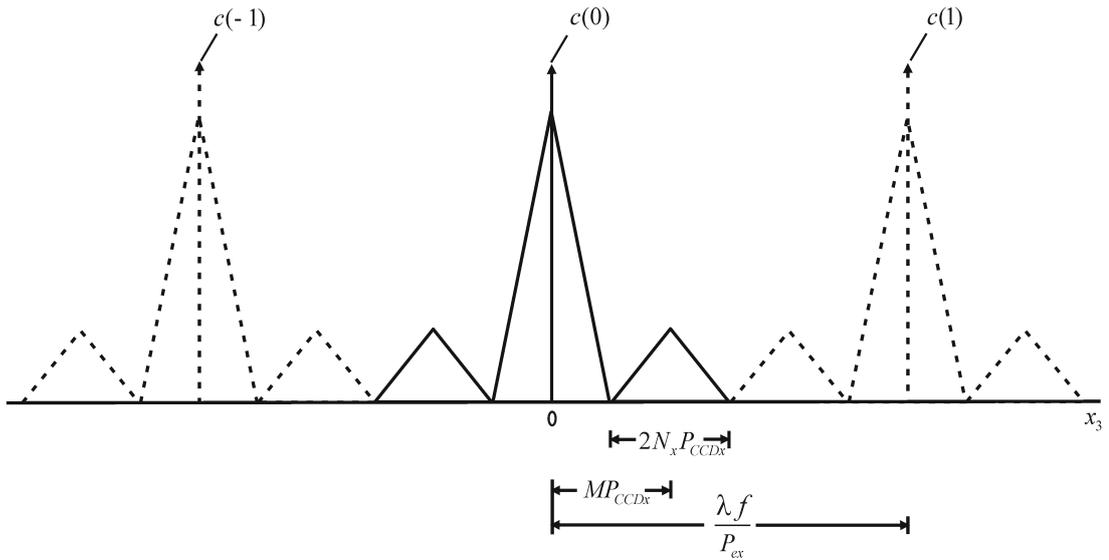
Figure 2.7 Location and width of correlation output of the joint transform correlator.

$$\frac{\lambda f}{P_{ex}} \geq 2MP_{CCDx} + 2N_x P_{CCDx}. \tag{2-30}$$

When $M = 2N_x$, the focal length of the lens L$_1$ must satisfy

$$f \geq \frac{6N_x P_{CCDx} P_{ex}}{\lambda}. \tag{2-31}$$

CHAPTER III

JPEG IMAGE COMPRESSION

In this chapter we review the algorithm of JPEG image compression and apply the algorithm to compress images with different spatial frequency contents and contrasts. Quantitative measurements of the effects of compression on image quality will be presented and used as the basis for analyzing the performance of the joint transform correlator with compressed reference images.

3.1   Introduction

Digital data is now prevalent in many multimedia applications such as internet, photography, mobile communication, medical imagery, digital libraries, and displays. However the storage of uncompressed digital data requires considerable space, while its transmission can be very time consuming such that it prohibits real-time systems. In order to solve storage and access time problems, the data could be compressed. In the case of digital image compression, the image will be represented with less data, while maintaining good image quality. Image compression operations reduce the data contents of the digital images and represent the image in a more compact form. There are two types of image compression, namely lossless and lossy compressions. A lossless compression retains the exact data of the original image bit for bit, while lossy compression returns the decompressed image being similar but not exactly same as the original image. As a result, lossy compression provides a higher compression level than that of the lossless compression. Among the various lossy

compression schemes, the JPEG is one of the most popular and well-established image compression standards.

This thesis employs the JPEG compression scheme to compress the reference images of the joint transform correlator. Four different sets of test images were first compressed into a JPEG file with various compression levels and were used as the reference images of the joint transform correlator. The compression performance and the image quality of the compressed images were then assessed. The recognition performance of the joint transform correlator by using the compressed reference images was finally investigated and the results will be presented in the next chapter.

3.2   The JPEG Compression

The JPEG compression algorithm is one of the digital image compression standards designed for compressing either full-color or gray scale images. The JPEG algorithm exploits the limitation of the human eye which is more sensitive to intensity than to color. Thus, the JPEG intends to discard information that is not easily perceived by the human eye such as high spatial-frequency components and small variation in color of images. The JPEG performs well on the continuous-tone images, while an image with many sudden jumps in intensity or color will not be compressed well. In the JPEG algorithm, compression is achieved in a series of steps (Pennebaker and Mitchell, 1993) as shown in Fig. 3.1:

- The pixels of the image are grouped into blocks of 8×8 pixels.

- A discrete cosine transform is applied to each block. The transform generates 1 DC and 63 AC components of spatial frequency.

- Each of the 64 discrete cosine transform components is quantized in conjunction with a quantization table. This is done in such a way that high

spatial-frequency components are quantized with higher quantization coefficient than the lower one. After quantization, the results are rounded to an integer. As a consequence, these AC components almost become zero. This cause irretrievably lost of information.

- The 64 quantized frequency components are encoded by using a combination of the run length encoding (RLE) and the Huffman coding.
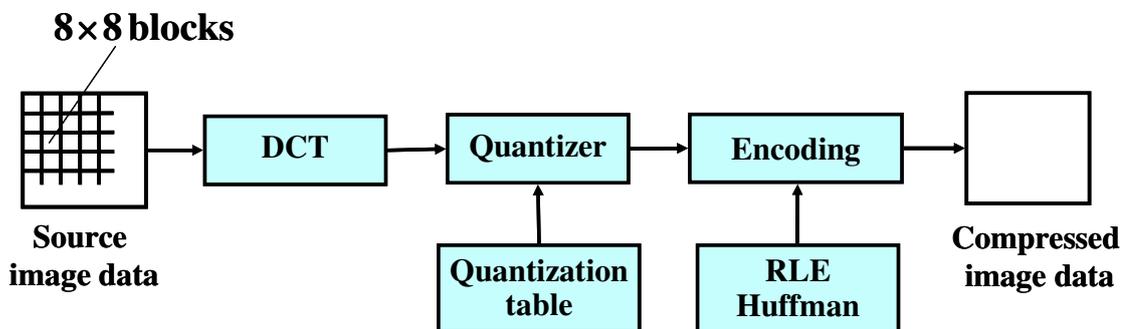
Figure 3.1   A block diagram of the JPEG encoder (Salomon, 1998).

3.3   JPEG-Compressed Reference Images

In order to study the effect of compression of the reference image on the recognition performance of the joint transform correlator, two input images with different spatial-frequency content and contrast were prepared as the reference images. All of images consisted of 124×186 pixels with 8-bit gray scale levels and their size were 23 Kbytes. Figures 3.2(a) and (b) show the fingerprint and the human face images employed as the high-contrast scenes with high and low spatial-frequency contents, respectively, while its low contrast versions are shown in Figs. 3.2(c) and (d). In comparison to Figs. 3.2(a) and (b), the low contrast images have smaller intensity variation of the luminance from the uniform background. Figures 3.3(a), (b),

(a)   (b)

(c)   (d)

Figure 3.2   Original images as test scenes: (a) high-contrast fingerprint, (b) high-

contrast human face, (c) low-contrast fingerprint, and (d) low-contrast

human face.

(c), and (d) correspond to the logarithmic value of the power spectra of the high-

contrast fingerprint, the high-contrast human face, the low-contrast fingerprint, and

the low-contrast human face, respectively. Since the fingerprint and the human face

are complex images, the contrast function defined as (Hess, Bradley, and Piotrowski,

1983)

$$C(u,v) = \frac{2A(u,v)}{DC} \qquad (3\text{-}1)$$

(a)   (b)

(c)   (d)
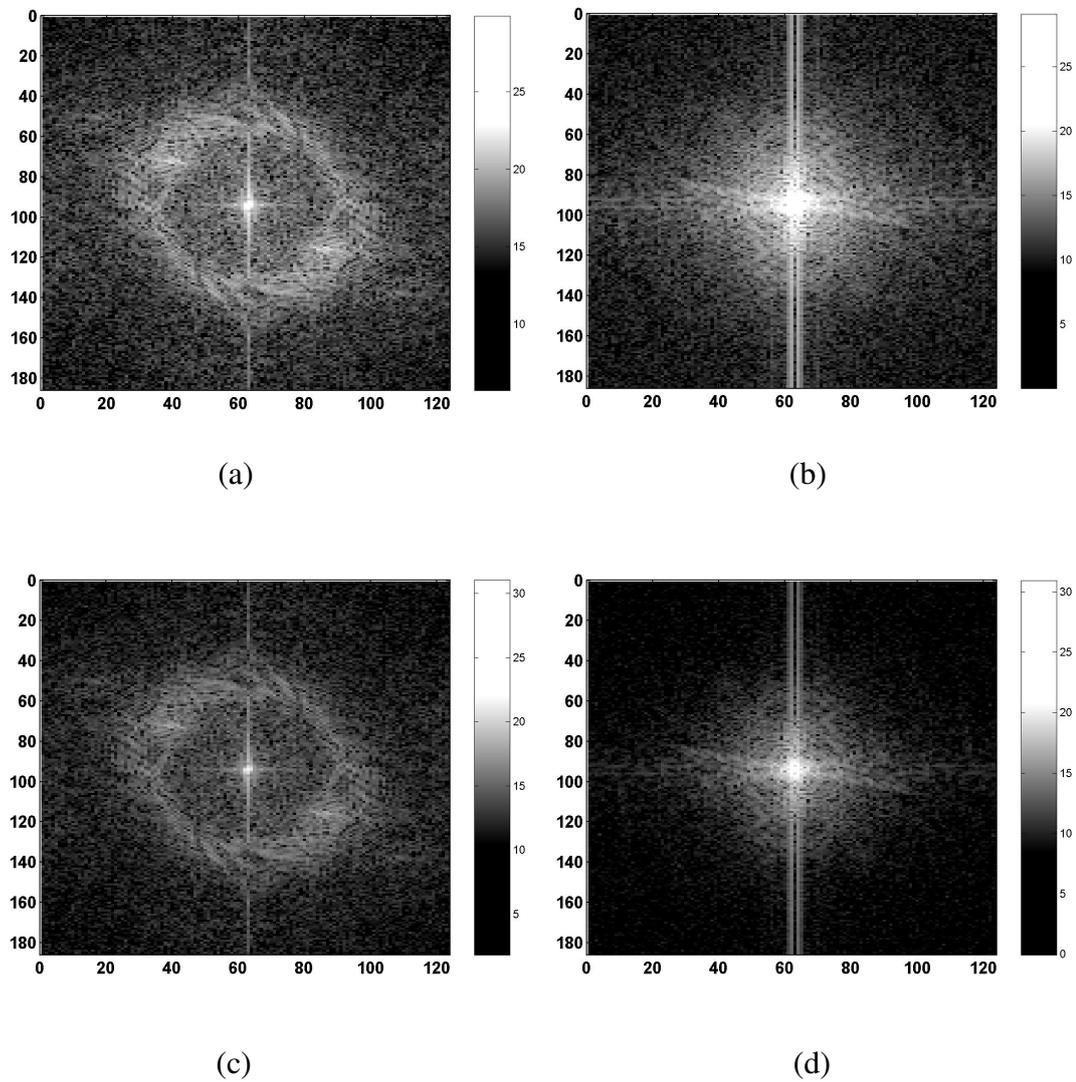
Figure 3.3   The power spectra of (a) high-contrast fingerprint, (b) high-contrast human face, (c) low-contrast fingerprint, and (d) low-contrast human face.

is used to confirm the difference between their contrasts, where $A(u,v)$ is the amplitude of the Fourier spectrum of the image. Figures 3.4(a) and (b) show the 1-dimensional (1-D) plots of the contrast function of the images with high and low spatial-frequency contents, respectively. It is clear from the figure that in the Fourier
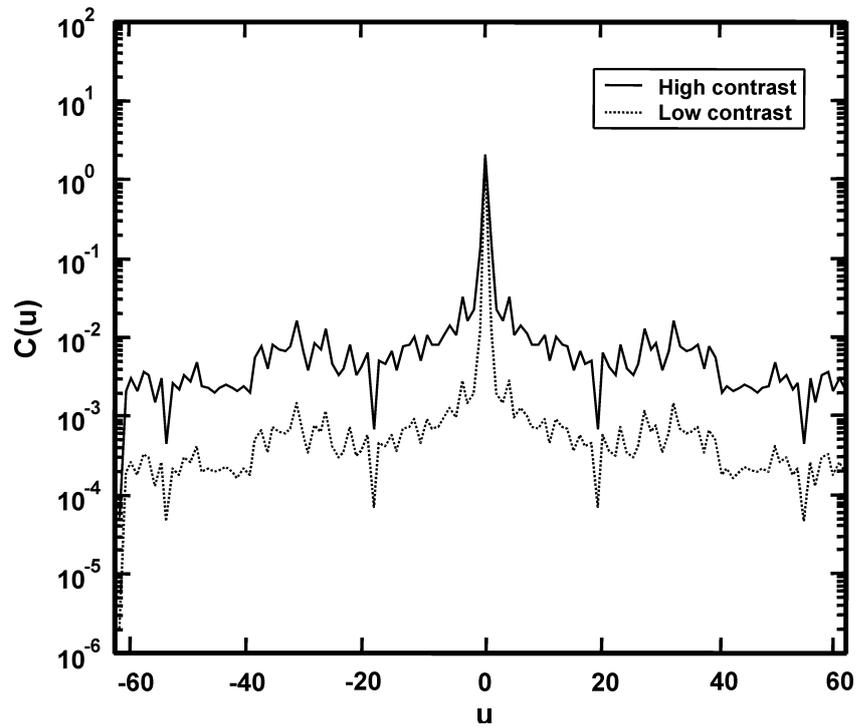
(a)
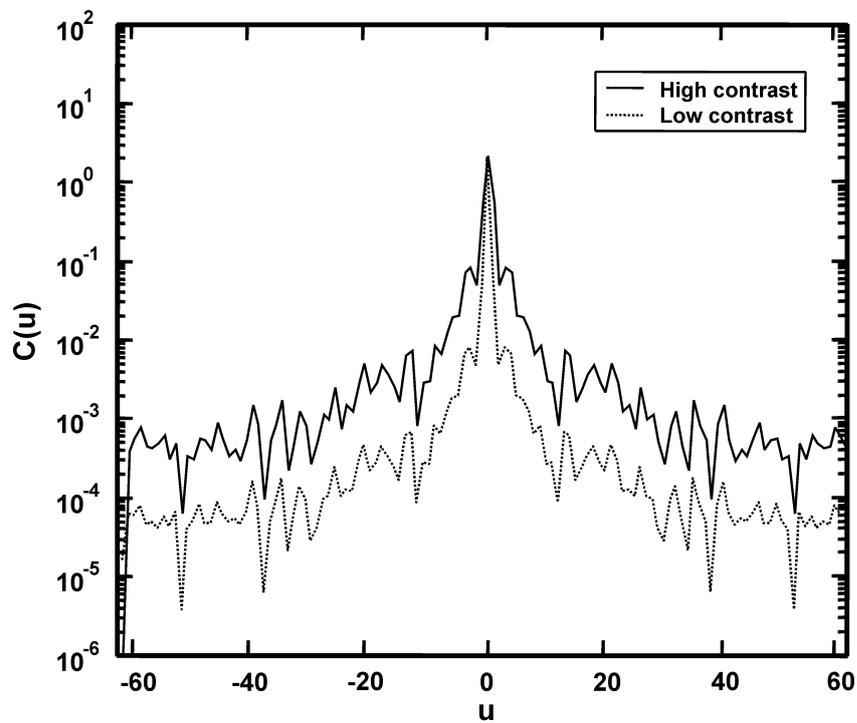
(b)

Figure 3.4 1-D plot of the contrast function of (a) fingerprints and (b) human faces.

domain, the contrast difference determines the amplitude of the spatial-frequency content, such that the high-contrast image has higher amplitude of the spatial frequency than that of the low contrast. This implies that the high-contrast images contain more high spatial- frequency components. The figures confirm also that in comparison with the fingerprint, the human face has less high spatial-frequency components, because their amplitudes are much lower.

The test images of fingerprints and human faces were compressed into the JPEG format by using the ACDsee software (The 2000 ACD systems, Ltd.) with different compression quality. In this software, the compression quality is determined by a parameter called the quality factor (QF) whose value can be varied from 100 to 0. High value of the QF discards less information than that of the 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 will be. Figure 3.5 shows compressed images with QF equals to 0, 10, 50, and 100. It is clear that at low QF, the quantization done independently on 8×8 blocks of pixels generates visible gray scale discontinuities along the block boundaries. These are called the blocking artifacts (Pennebaker and Mitchell, 1993). The subjective evaluation of the compressed test images is shown in Table I. Regardless the spatial frequency contents of the image, the results show that the degradation of the compressed low-contrast images is more severe than that of the high contrast images.

3.3.1   Compression Performance

The compression produced by the ACDsee software was assessed by using the compression ratio (CR) defined as the ratio of the uncompressed to the compressed

QF = 0   QF = 10   QF = 50   QF = 100

(a)

(b)

(c)

(d)
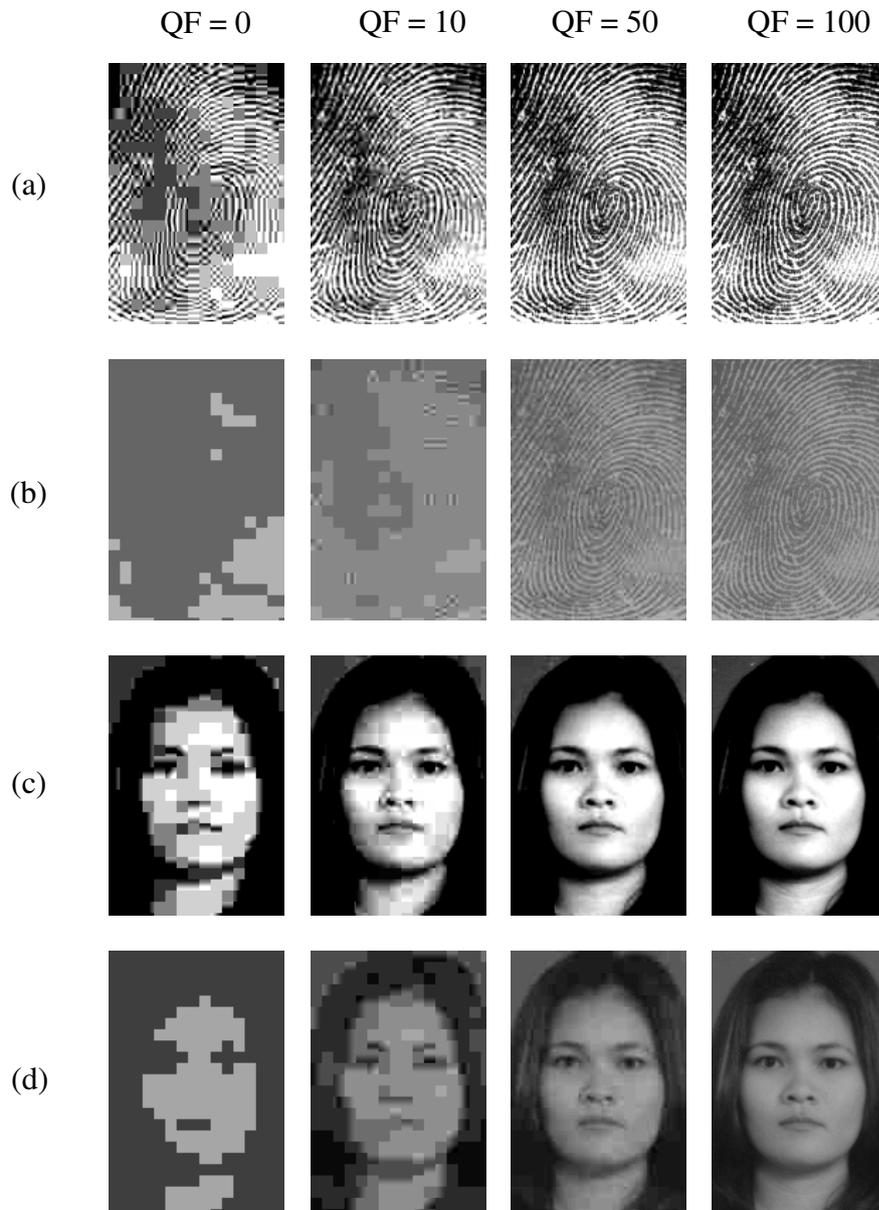
Figure 3.5   JPEG compressions of (a) high-contrast fingerprint, (b) low-contrast

fingerprint, (c) high-contrast human face, and (d) low-contrast human

face.

file sizes (Pennebaker and Mitchell, 1993). High CR corresponds to the small size of

the compressed image. Figure 3.6 shows the CR of the reference images as a function

Table 3.1 The subjective evaluation of the test images.

Test images	Subjective Ranking			
	QF=0	QF=10	QF=50	QF=100
High-contrast fingerprint	Poor	Fair	Good	Excellent
Low-contrast fingerprint	Worst	Very poor	Fair	Excellent
High-contrast human face	Poor	Fair	Good	Excellent
Low-contrast human face	Very poor	Poor	Fair	Excellent
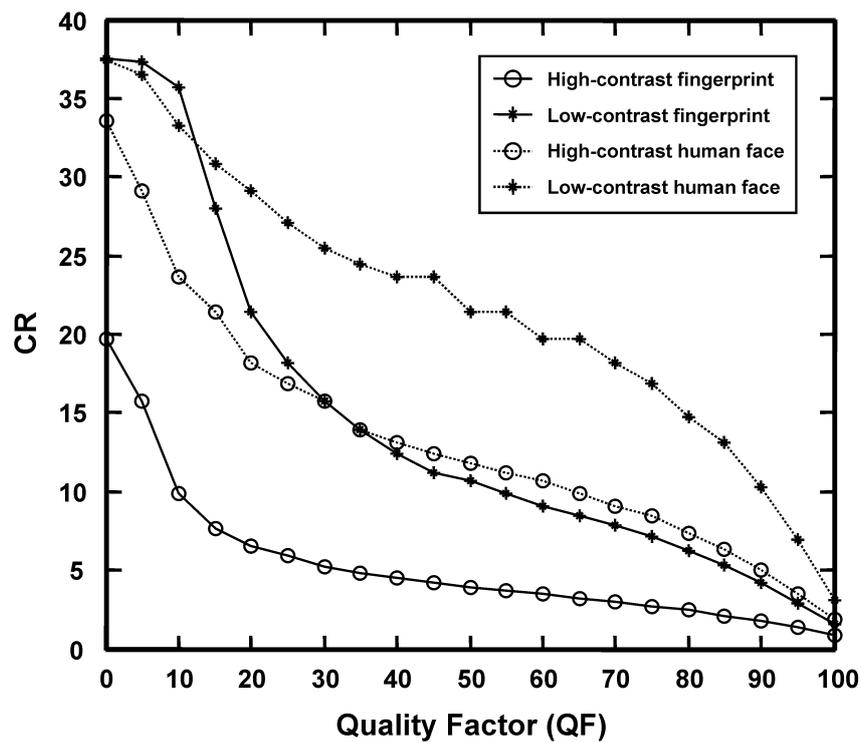
Figure 3.6   The CR as function of the QF.

of the QF. It is clear from Fig. 3.6 that regardless of the spatial-frequency content, the

CR of the low-contrast image is higher than that of the high-contrast image. This is

due to the fact that as shown in Fig. 3.4 the low-contrast image contains less high

spatial-frequency components than the high-contrast image does. The quantization process done on the low contrast image causes more AC components of the spatial frequency to become zero. As a consequence, the RLE and the Huffman coding can encode efficiently the redundant zeroes. For the same reason, the CR of the high-contrast human face is higher than that of the fingerprint.

3.3.2   Image Quality Measurements

The effect of image compression on the quality of the image was objectively evaluated by using the peak-signal-to-noise ratio (PSNR) defined as (Yang, Zhang, and Mitra, 1999)

$$PSNR = 10\log_{10} \frac{255^2}{\frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [f(i,j) - \hat{f}(i,j)]^2} \tag{3-2}$$

Here $f(i,j)$ and $\hat{f}(i,j)$ are the original and the compressed images with $M \times N$ pixels, respectively. $f(i,j)$ takes integer values $0 \leq f(i,j) \leq 255$ for 8-bit gray scale images. The PSNR measures the similarity of the images in decibels (dB). The large PSNR means that the degree of similarity between the original and it compressed versions is high. Based on Eq. (3-2), the PSNR for each test image was calculated and the results are shown in Fig. 3.6. It can be seen from the figure that the PSNRs increase as the QF increases. In addition regardless the spatial frequency contents of the images, the PSNRs of the low contrast image is higher than that of the high contrast image. This is because the high-contrast image contains more high spatial-frequency components than the low-contrast image. Since the high spatial-frequency determines fine details of the image, discarding high spatial-frequency components of the high-contrast image degrades the fine details more significantly than that of the
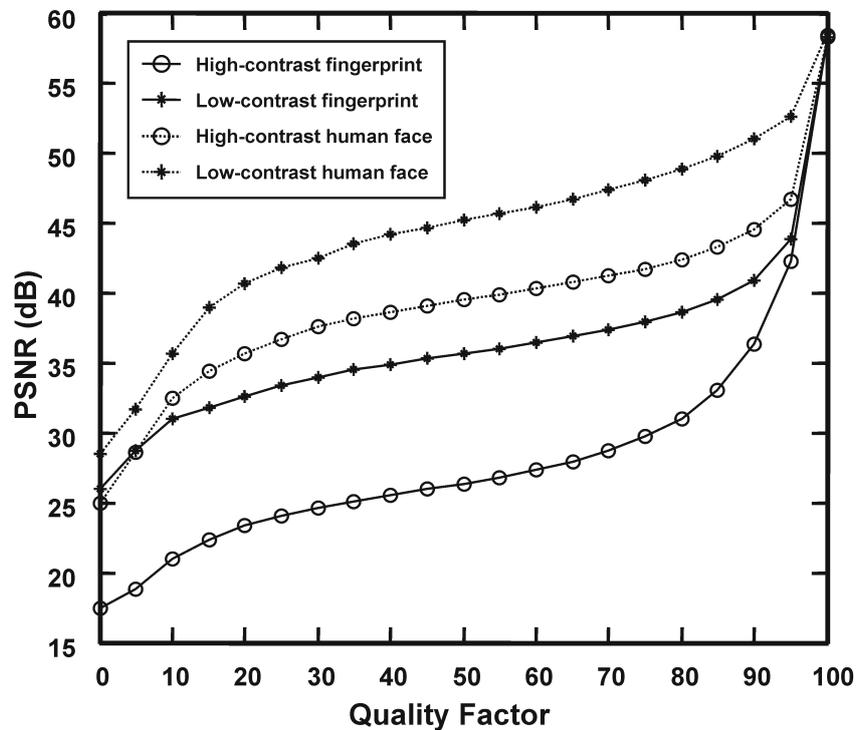
Figure 3.7 The PSNRs as function of the QF.

low-contrast image. Therefore, the PSNR of the compressed high-contrast fingerprint is the lowest, because the degradation is the most severe. However, based on the subjective evaluation provided in Sect. 3.3, the image quality of the compressed low-contrast image is poorer than that of the compressed high-contrast images especially at small QF. This difference may be caused by the limitation of the human eyes in detecting degradation of fine details.

CHAPTER IV

SINGLE-TARGET DETECTION

In this chapter, the proposed real-time implementation of joint transform correlator with JPEG-compressed reference images for single-target detection is studied by using computer simulations and experiments. The simulation results show that in comparison with the compressed high spatial-frequency image, the joint transform correlator by using the compressed low spatial-frequency reference image offers a better recognition performance in that it is robust to noise and contrast difference for a wide range of compression levels. However, due to the limitation of the EASLM on displaying low contrast image and low dynamic range of the CCD sensor, not all experimental results for different situations can be successfully obtained.

4.1   The Joint Transform Correlator with Compressed Reference Images

This study is based on the optical setup for implementing the real-time joint transform correlator with compressed reference image shown in Fig. 2.3. If the compressed reference is represented by $r_c(x_1, y_1)$, the joint input image of Eq. (2-2) can be mathematically rewritten as

$$f_{JTC}(x_1, y_1) = r_c(x_1 - x_0, y_1) + t(x_1 + x_0, y_1) .\tag{4-1}$$

Under a presence of additive white Gaussian noise $n(x_1, y_1)$ at the input target and a

contrast difference between the target and the compressed reference images, the joint input image can be rewritten as

$$f_{JTC}(x_1, y_1) = r_c(x_1 - x_0, y_1) + c_T t(x_1 + x_0, y_1) + n(x_1 + x_0, y_1),\qquad(4\text{-}2)$$

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

$$
\begin{aligned}
U(x_2, y_2) &= \left| R_c\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) \right|^2 + c_T^2 \left| T\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) \right|^2 + \left| N\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) \right|^2 \\
&+ c_T \left[T^*\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) N\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) + T\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) N^*\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) \right] \\
&+ c_T \left[R_c\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) T^*\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) \exp\left(-j2\pi x_0 \frac{x_2}{\lambda f}\right) \right. \\
&\left. + T\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) R_c^*\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) \exp\left(j2\pi x_0 \frac{x_2}{\lambda f}\right) \right] \\
&+ R_c\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) N^*\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) \exp\left(-j2\pi x_0 \frac{x_2}{\lambda f}\right) \\
&+ N\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) R_c^*\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) \exp\left(j2\pi x_0 \frac{x_2}{\lambda f}\right),\qquad(4\text{-}3)
\end{aligned}
$$

where (x_2, y_2) are the coordinates at the Fourier plane. $R_c\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right)$ and $N\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right)$ are the Fourier transforms of the compressed reference, and the noise, respectively. By displaying the captured joint power spectrum 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 sixth, seventh, eighth, and

ninth terms of Eq. (4-3) can be expressed as

$$I(x_3, y_3) = c_T[r_c(x_3, y_3) * t(x_3, y_3) \otimes \delta(x_3 \pm 2x_0)]$$
$$+ r_c(x_3, y_3) * n(x_3, y_3) \otimes \delta(x_3 \pm 2x_0), \qquad (4\text{-}4)$$

where $*$ denote correlation. The first term of Eq. (4-4) corresponds to the desired correlation of the input target with the compressed reference which is scaled by the contrast difference, while the second one is the unwanted correlation of the compressed reference with the noise. Since both terms appear at the same position $\pm 2x_0$, Eq. (4-4) indicates that besides the image quality of the compressed reference $r_c(x_1, y_1)$, the correlation output depends on both the contrast and the noise. Therefore, it is important to measure the effects of image compression on the correlation performance of the joint transform correlator.

4.2   Simulation Results

In the computer simulation, the target and the compressed reference images were combined to form the joint input image of 832×624 pixels with a separation of $2x_0 = 248$ pixels. Each set of compressed references consisted of 21 images with QF varying from 0 to 100 with 5 incremental steps. Each target consisted of images with no noise, noise with variance $\sigma^2 = 0.01$ and variance $\sigma^2 = 1$, respectively. All noises were generated by using the IMNOISE command of MATLAB software. The algorithm for performing computer simulation of the joint transform correlator with compressed reference images is shown in Fig. 4.1. The target and the compressed reference with QF = 0 were first combined to form the joint input image. The joint transform correlator of this joint image was computed and its resultant output was quantified by measuring a ratio of the correlation peak intensity to the standard
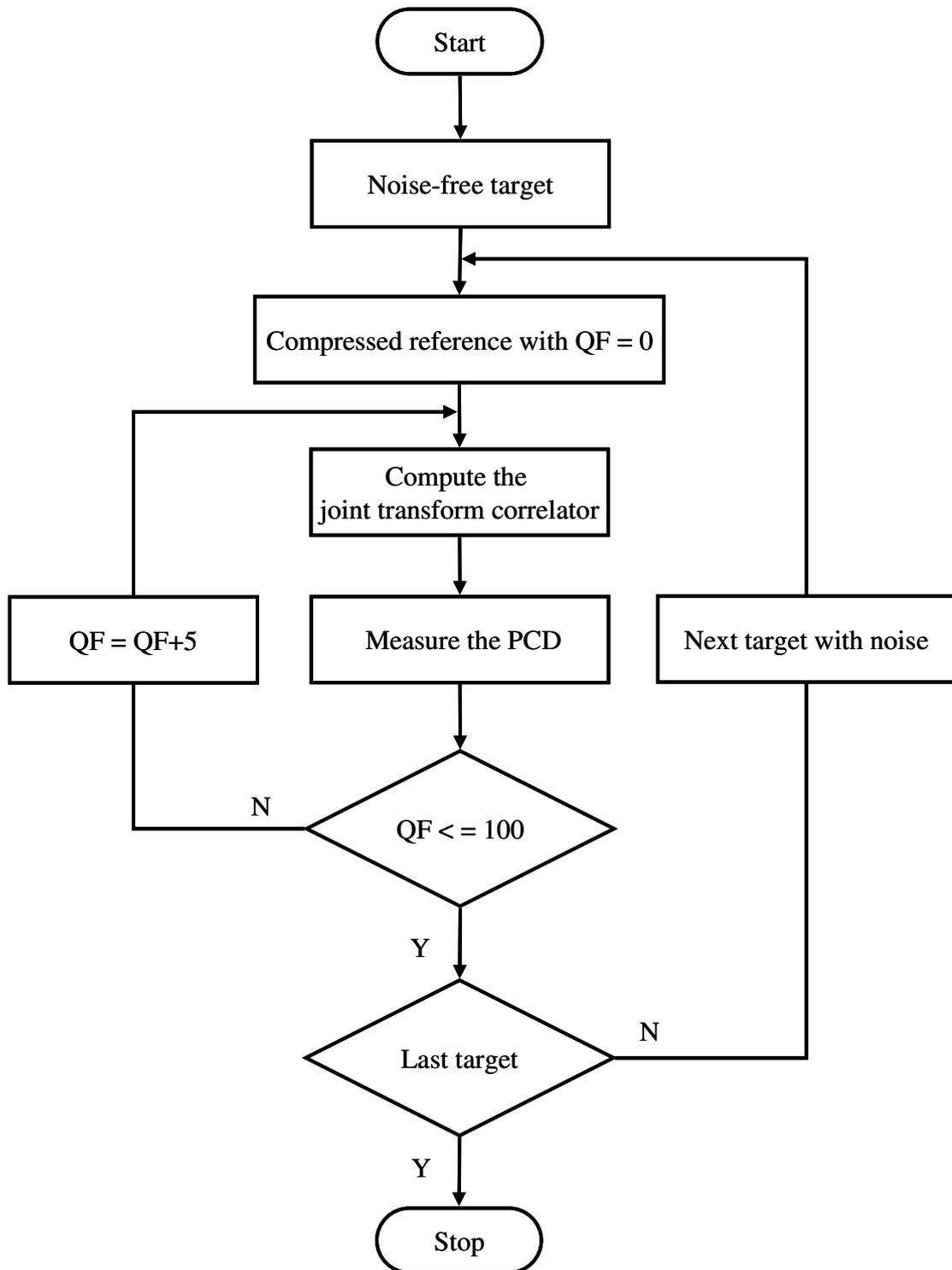
Figure 4.1   Flowchart for conducting computer simulation of the joint transform

correlator with compressed reference images.

deviation of the correlation intensity or peak-to-correlation deviation (PCD) given by
(Roberge and Sheng, 1994)

$$PCD = \frac{I(i,j)_{max}}{\left\{ \frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \left[I(i,j) - E\{I(i,j)\} \right]^2 \right\}^{1/2}} . \qquad (4\text{-}5)$$

Here, $I(i,j)_{max}$ is the maximum intensity of the correlation output, while $E\{I(i,j)\}$ is
the mean of the correlation intensity. In the case that the target matches the
compressed reference, the correlation function has a sharp peak and its standard
deviation is small. Thus, the PCD of Eq. (4-5) is large. In the unmatched case, the
correlation output is broad and its peak is low. Since the standard deviation is large,
the PCD is small. In order to compare the recognition performance at different
compression levels, each PCD is normalized by the value of the autocorrelation peak.
The computation of the joint transform correlator is repeated for different compressed
reference images. After QF = 100, the computation of the joint transform correlator is
restarted for the next target image with the reference having QF = 0.

Figure 4.2 illustrates a flowchart for computing the joint transform correlator.
The Fourier transform of the joint input image is calculated by using the 2-D FFT
command of the MATLAB 6.0 run on a Windows-based personal computer. The joint
power spectrum is generated by taking the modulus square of the resultant Fourier
spectrum. By calculating the inverse Fourier transform of the joint power spectrum
and followed by taking the modulus square of the result, the correlation output
intensity is produced.
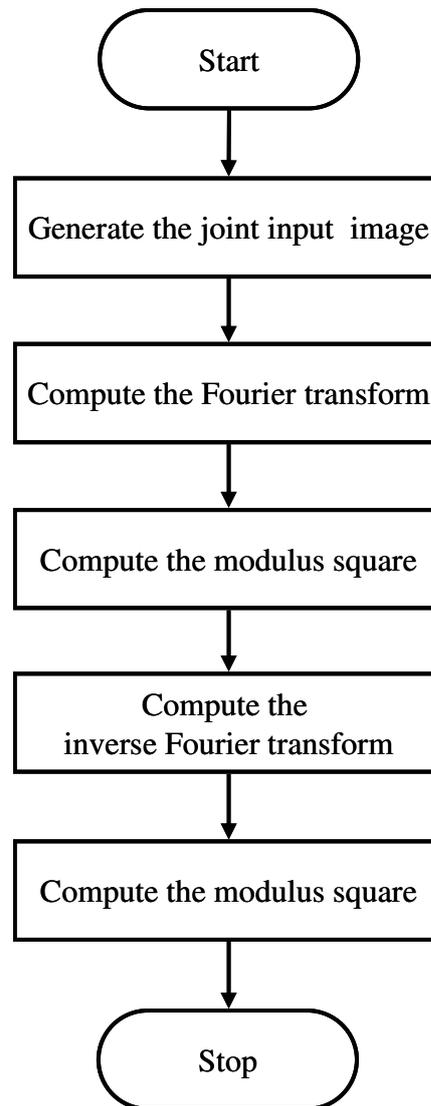
Figure 4.2   Flowchart for computing the joint transform correlator.

4.2.1   Compressed High-Contrast Fingerprint as the Reference Images

The 3-D plot of the autocorrelation of the uncompressed high-contrast fingerprint is shown in Fig. 4.3(a), while the correlation outputs of the joint transform correlator by using the compressed high-contrast fingerprint with the QF = 10 as the reference are illustrated in Figs. 4.3(b), (c), and (d). In Fig. 4.3(b), the target is the noise-free high-contrast fingerprint image. It is obvious that besides the correlation
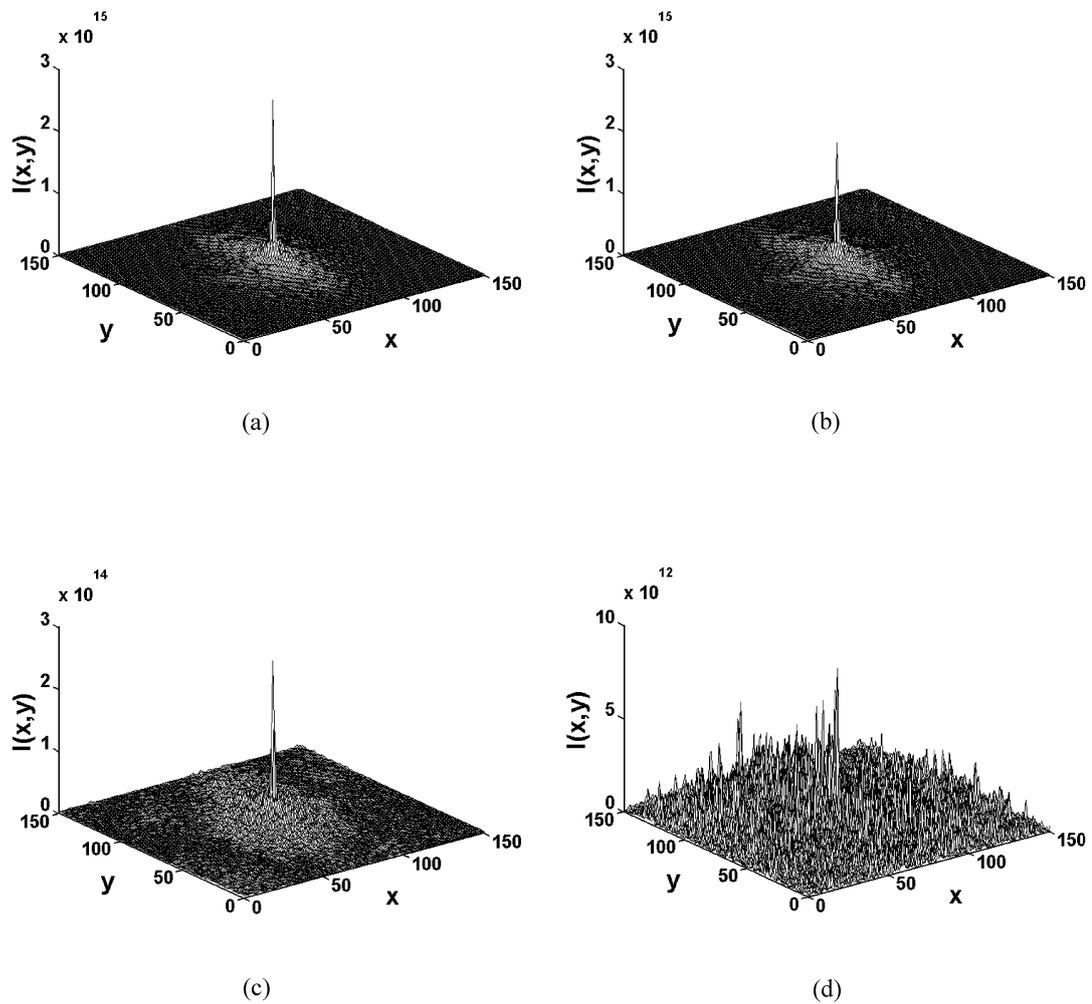
Figure 4.3   Simulation results of the joint transform correlator. (a) autocorrelation of the uncompressed high-contrast fingerprint; and cross-correlation outputs by using the compressed high-contrast fingerprint reference (QF = 10) under a situation that the target is: (b) noise-free high-contrast fingerprint, (c) noisy high-contrast fingerprint ($\sigma^2 = 1$), and (d) noisy low-contrast fingerprint ($\sigma^2 = 1$).

signal is almost as sharp as the autocorrelation output shown in Fig. 4.3(a), its peak decrease slightly. This is the effect of the compression on the reference image (Widjaja, 2003). However, since the CR of the high-contrast fingerprint is small (see Chapter II), the decrease of the degree of similarity between the two images is not significant. The correlation outputs caused by the noise-corrupted high- and low-contrast fingerprint targets with the variance $\sigma^2 = 1$ are shown in Figs. 4.3(c) and (d), respectively. When the target is the noisy high-contrast fingerprint, the peak intensity falls about one order of magnitude compared with the noise-free target, and its correlation plane is slightly noisy. However, in the case of the noisy low-contrast fingerprint target, besides the peak intensity falling by about two orders of magnitude, the correlation plane appears to be noisier. These occur because first, since the target has a lower contrast than the reference, a contribution of the correlation term $r_c(x, y) * t(x, y)$ of Eq. (4-4) is reduced by the factor c_T which is smaller than 1. Second, the small luminance value of the low contrast image is sensitive to noise. In particular when the noise variance σ^2 is 1, the noise level is stronger than the image luminance. Therefore, under the presence of the same noise level, the correlation plane of Fig. 4.3(d) appears to be more noisy.

Figure 4.4 shows the variation of the normalized PCDs as a function of the QF of the compressed high-contrast fingerprint reference image for different target scenes. The normalized PCD is calculated from the ratio of the PCD of the cross correlation between the target and the compressed reference images to the PCD of the autocorrelation of the uncompressed test images. In general regardless of the contrast, the PCDs decrease gradually as the QF becomes smaller, because when more

information is discarded from the compressed reference the degree of similarity between the reference and the target images decreases. However, in a situation where the variance of the noise is $\sigma^2 = 1$, the PCD of the low-contrast fingerprint target reduces drastically to less than 0.3. As shown in Fig. 4.3(d), this is caused by the decrease of the correlation term $r_c(x, y) * t(x, y)$ and the increase of the noise in the correlation plane. This PCD does not increase as the QF becomes higher. Therefore, the recognition of the noisy low-contrast fingerprint target is dependent upon the noise rather than the compression level of the reference image, while the recognition of the high-contrast fingerprint depends only on the compression level.
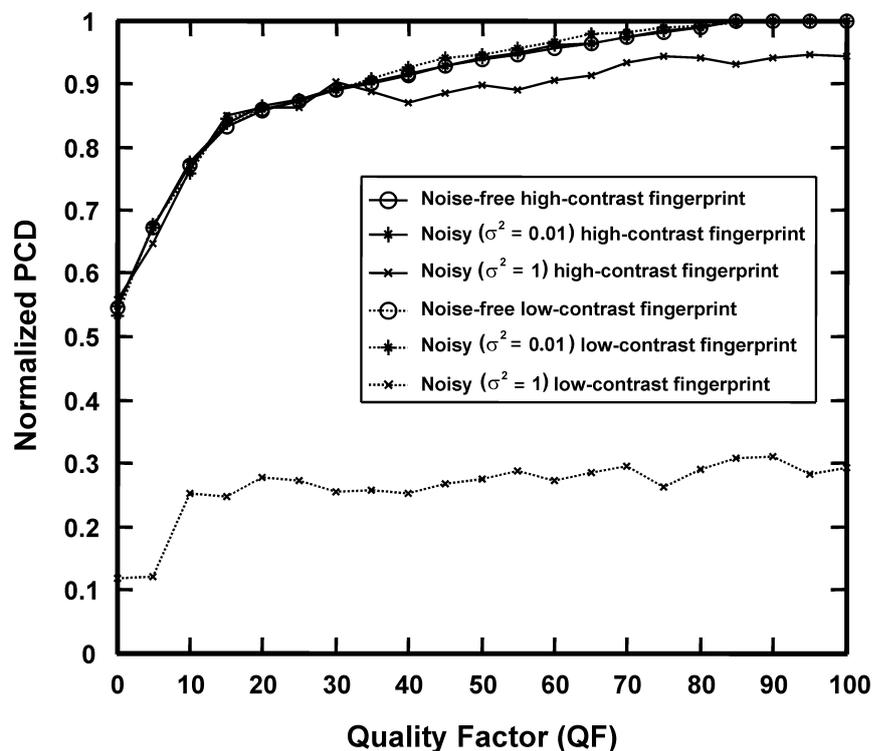
Figure 4.4　The PCD-based recognition performance of the joint transform correlator as a function of the QF of the compressed high-contrast fingerprint reference.

4.2.2   Compressed Low-Contrast Fingerprint as the Reference Images

Figures 4.5(a), (b), (c), and (d) show the 3-D correlation outputs of the joint transform correlator by using the low-contrast fingerprint as the reference image. Since the image contrast is low, it is expected that the autocorrelation peak of the uncompressed low-contrast fingerprint is lower than the autocorrelation of the high-contrast image. This can be clearly observed in Fig. 4.5(a) which shows that although the output is sharp, its peak reduces by three orders of magnitude compared to the autocorrelation of the high-contrast image shown in Fig. 4.3(a). When the reference is the compressed low-contrast fingerprint with the QF = 10 and the target is the noise-free low-contrast fingerprint, the correlation output shown in Fig. 4.5(b) becomes broad and its peak decreases. This is mainly caused by the effect of compression of the low-contrast fingerprint image which suffers from more loss of the high spatial-frequency components than that of the high-contrast fingerprint. Since the reference image is the impulse response of the joint transform correlator, this loss yields a broad impulse response. In the other cases of using the same compressed reference with the QF = 10, the joint transform correlator of the noisy low-contrast fingerprint target with the variance $\sigma^2 = 1$ shown in Fig. 4.5(c) gives the correlation output which is noisier than that of the noisy high-contrast fingerprint target of Fig. 4.5(d). This occurs because the noise is stronger than the target luminance. Thus although the contrast ratio c_T is unity, the contribution of the correlation term $r_c(x, y) * n(x, y)$ is greater than the term $r_c(x, y) * t(x, y)$. As a consequence; the desired correlation peak is indistinguishable as it is buried in a strong noise. Figure 4.5(d) shows the correlation output produced by the noisy high-contrast fingerprint target with the
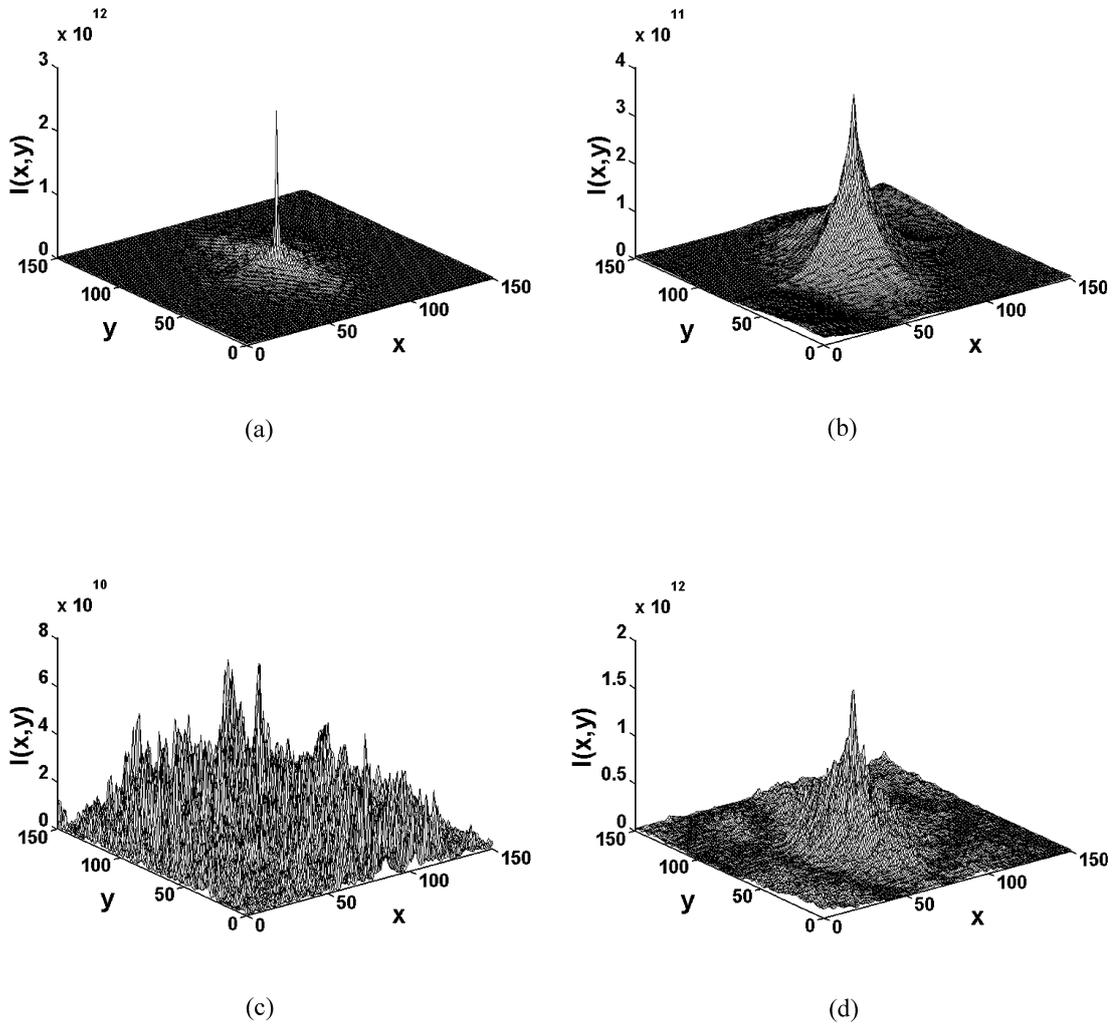
Figure 4.5   Simulation results of the joint transform correlator. (a) autocorrelation of the uncompressed low-contrast fingerprint; and cross-correlation outputs by using the compressed low-contrast fingerprint reference (QF = 10) under a situation that the target is: (b) noise-free low-contrast fingerprint, (c) noisy low-contrast fingerprint ($\sigma^2 = 1$), and (d) noisy high-contrast fingerprint ($\sigma^2 = 1$).

variance $\sigma^2 = 1$. In this case, the luminance value of the high-contrast fingerprint is stronger than the noise, and the contrast ratio is greater than unity. As a consequence, the correlation term $c_T[r_c(x, y) * t(x, y)]$ is stronger than the term $r_c(x, y) * n(x, y)$. Therefore, although the correlation plane appears to be noisy, the correlation peak can still be detected.

Figure 4.6 shows the normalized PCDs as a function of the QF of the compressed reference for different target scenes. It is clear that as the QF becomes smaller, the PCDs decrease rapidly regardless of the image property of the targets. Thus, in comparison with the compressed high-contrast reference, the compression of
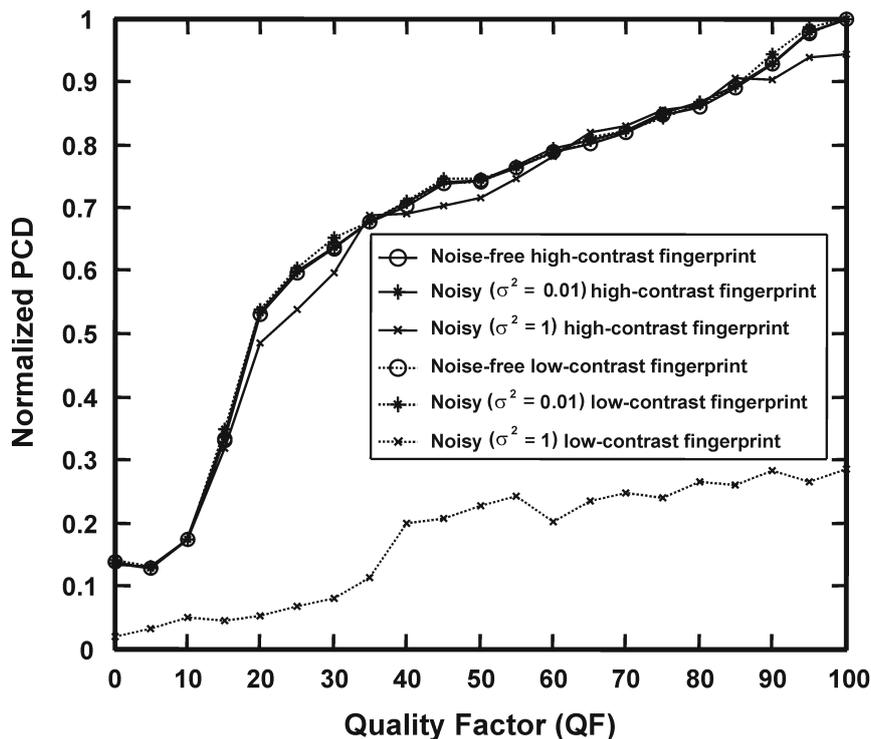
Figure 4.6   The PCD-based recognition performance of the joint transform correlator as a function of the QF of the compressed low-contrast fingerprint reference.

the low-contrast fingerprint reference degrades significantly the correlation performance of the joint transform correlator. This is in agreement with the discussion in the previous paragraph that the compressed low-contrast fingerprint contains less high spatial-frequency information than that of the compressed high-contrast fingerprint. As a result, the degree of similarity between the compressed reference and the target decreases rapidly. Furthermore, as shown in Fig. 4.5(c), when the low-contrast fingerprint target is corrupted by the noise with the variance $\sigma^2 = 1$, the noise level is stronger than the target. Therefore, the PCD falls below 0.3. In comparison with Fig. 4.4, the degradation of the performance of the joint transform correlator caused by compressing the low-contrast fingerprint reference is more severe.

4.2.3 Compressed High-Contrast Human Face as the Reference Images

Figures 4.7(a), (b), (c), and (d) illustrate the 3-D correlation outputs of the joint transform correlator by using the compressed high-contrast human face as the reference image. The autocorrelation output of the uncompressed high-contrast human face shown in Fig. 4.7(a) is broader than that of the high-contrast fingerprint shown in Fig. 4.3(a). This is caused by the characteristic of the human face image which contains less high spatial-frequency information. As a consequence, the impulse response of the joint transform correlator by using the human face image as the reference becomes broader than that of the high-contrast fingerprint reference. When the reference is the compressed high-contrast human face with the QF = 10 the following results are obtained: As shown in Fig. 4.7(b), the joint transform correlator of the noise-free high-contrast human face target gives slightly broader correlation output compared to the autocorrelation output, because of the compression. The
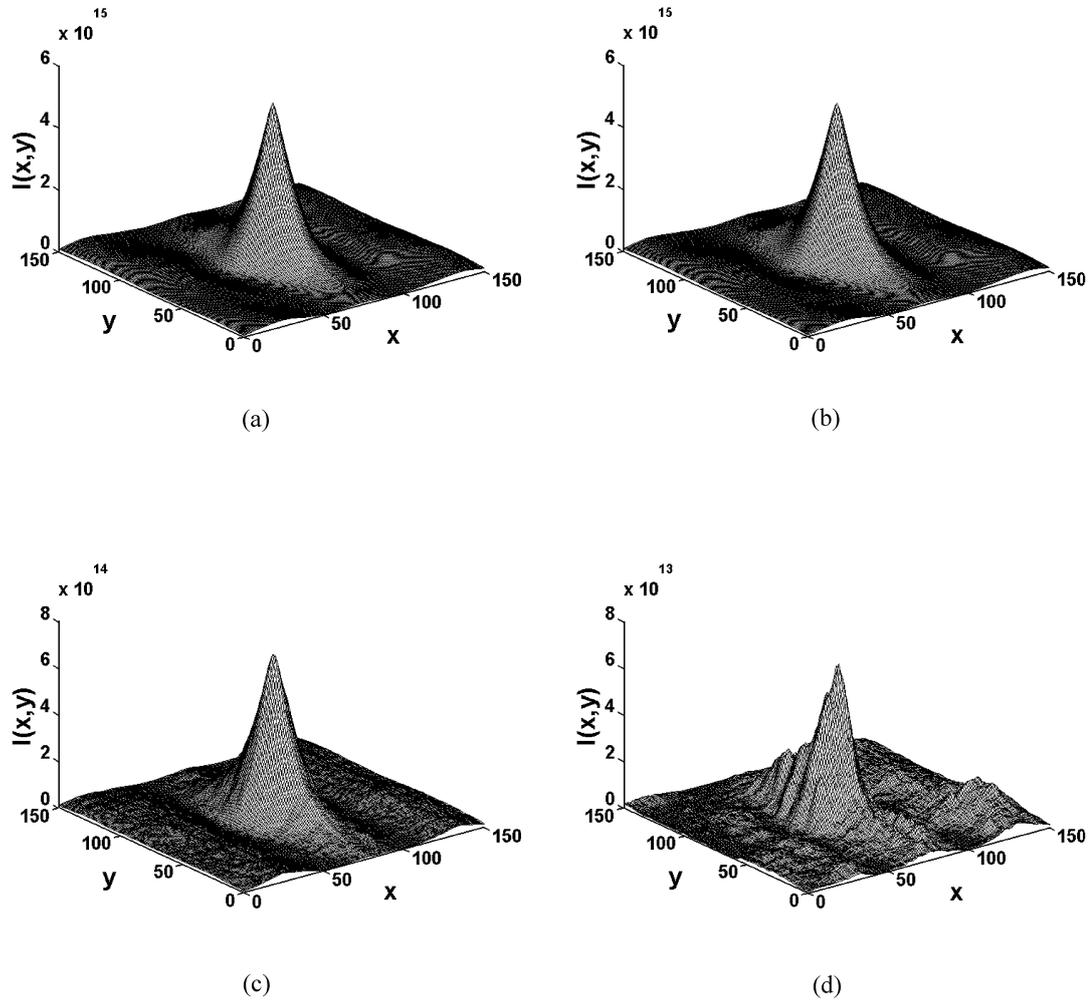
Figure 4.7   Simulation results of the joint transform correlator. (a) autocorrelation of the uncompressed high-contrast human face; and cross-correlation outputs by using the compressed high-contrast human face reference (QF = 10) under a situation that the target is: (b) noise-free high-contrast human face, (c) noisy high-contrast human face ($\sigma^2 = 1$), and (d) noisy low-contrast human face ($\sigma^2 = 1$).

correlation outputs of the noise-corrupted high- and low-contrast human face targets with the variance $\sigma^2 = 1$ are depicted in Figs. 4.7(c) and (d), respectively. It is obvious that due to the lower contrast, the correlation peak of Fig. 4.7(d) is lower than that of Fig. 4.7(c). In comparison to Figs. 4.3(c) and (d), these correlation planes are less noisy. This is mainly caused by the broad impulse response of the joint transform correlator which confines the output of the correlation term $n(x, y) * r_c(x, y)$ within its impulse area. Since the correlation term $t(x, y) * r_c(x, y)$ is broad and its peak is high, the effect of the noise is not significant.

Figure 4.8 shows the normalized PCDs as a function of the QF of the
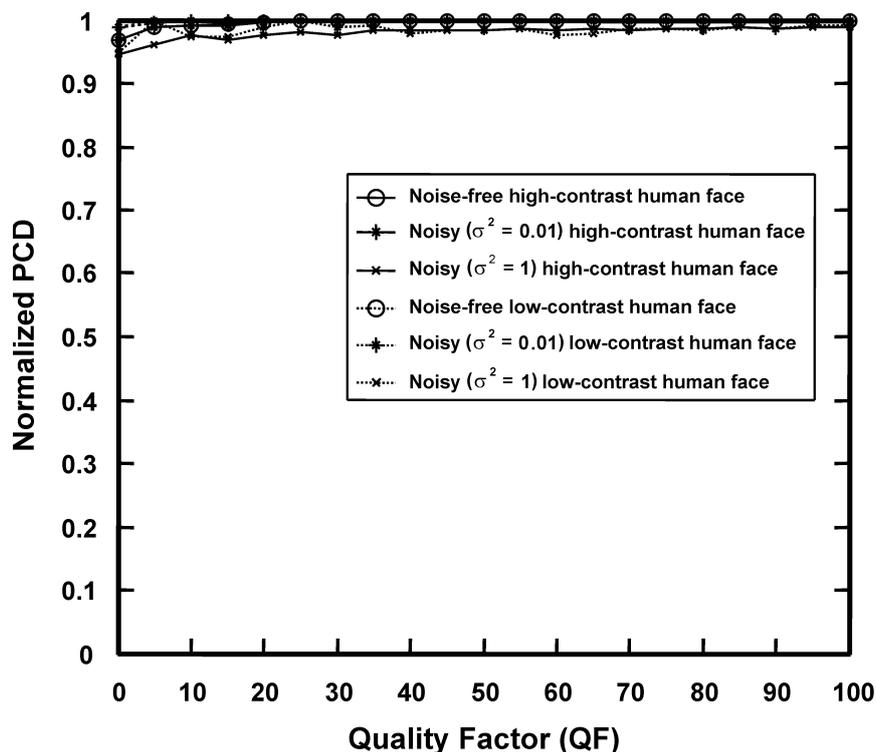
Figure 4.8   The PCD-based recognition performance of the joint transform correlator as a function of the QF of the compressed high-contrast human face reference.

compressed high-contrast human face for different target scenes. It is clear that for all given types of targets, the normalized PCDs almost do not vary with respect to compression level. This indicates that the presence of the noise and the contrast difference between the target and the reference do not affect the correlation performance of the joint transform correlator by using the compressed high-contrast human face reference. As shown in Figs. 4.7(b), (c), and (d) since the term $t(x, y) * r_c(x, y)$ gives broad correlation output, the degradation caused by the noise and the contrast difference do not change significantly the standard deviation of the correlation output. As a result, the normalized PCDs are always maximum for nearly all target scenes.

4.2.4   Compressed Low-Contrast Human Face as the Reference Images

Figure 4.9(a) shows that the autocorrelation of the uncompressed low-contrast human face is as broad as Fig. 4.7(a). However, its peak reduces by about two orders of magnitude, because the image contrast is low. Note that although the autocorrelation of the low-contrast human face is broader than that of the low-contrast fingerprint shown in Fig. 4.5(a), its peak is higher. This is caused by the luminance of the low-contrast human face which is higher than the fingerprint. Figures 4.9(b), (c), and (d) show the correlation peaks of the joint transform correlator by using the compressed low-contrast human face with QF = 10 as the reference. In Fig. 4.9(b), the target is the noise-free low-contrast human face image. The correlation output further decreases and broadens. This is the effect of the compression of the reference image. Figures 4.9(c) and (d) show the correlation output of the noisy low- and high-contrast human face targets with the noise variance $\sigma^2 = 1$, respectively. When the input
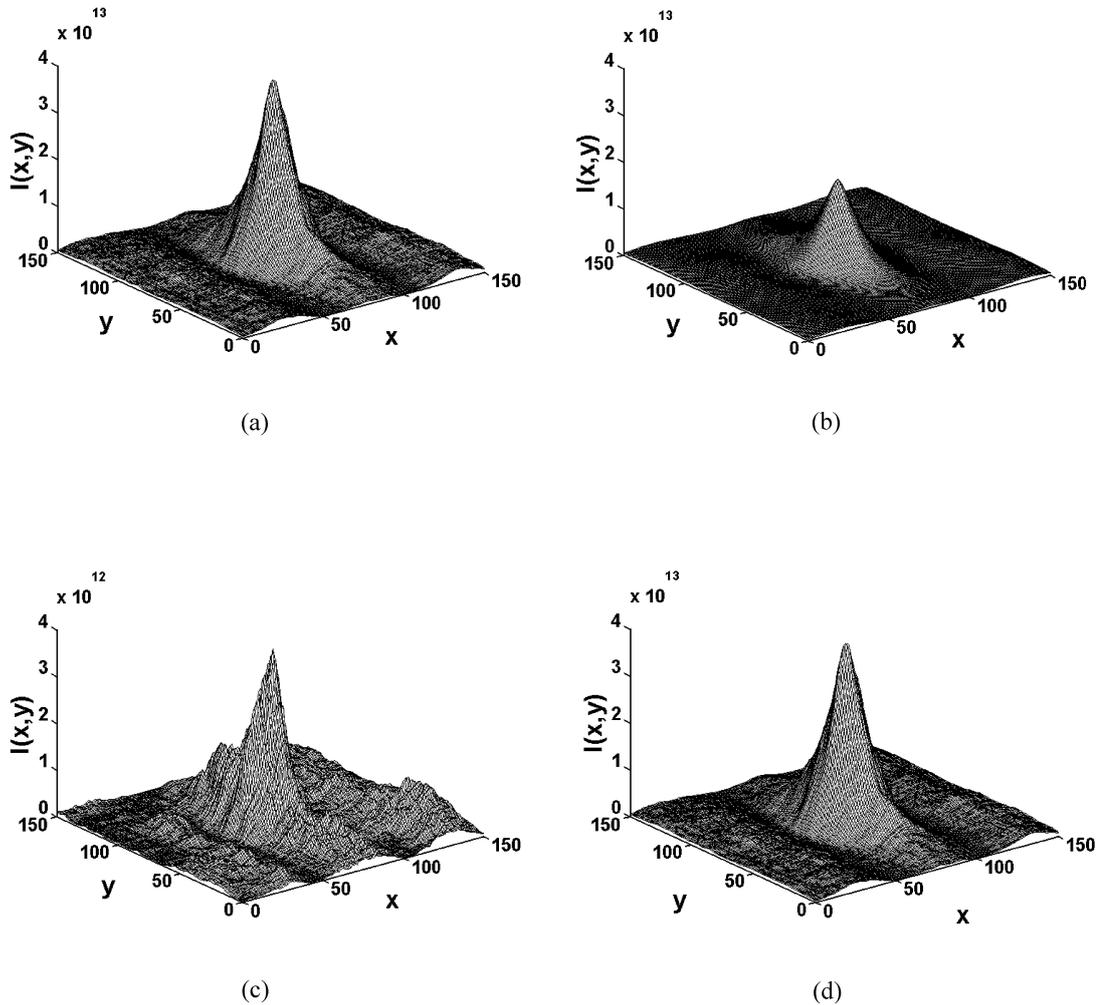
Figure 4.9 Simulation results of the joint transform correlator. (a) autocorrelation of the uncompressed low-contrast human face; and cross-correlation outputs by using the compressed low-contrast human face reference (QF = 10) under a situation that the target is: (b) noise-free low-contrast human face, (c) noisy low-contrast human face ($\sigma^2 = 1$), and (d) noisy high-contrast human face ($\sigma^2 = 1$).

target is the low-contrast human face, the degradation of the correlation output caused by the noise is more severe than that of the high-contrast target shown in Fig. 4.9(d). This is in agreement with the resultant correlations discussed in Sect. 4.2.2. Since the luminance of the low-contrast target is smaller than the noise, the presence of the correlation term $n(x, y) * r_c(x, y)$ is obvious in the correlation plane. Furthermore, in comparison with Figs. 4.5(c) and (d), the joint transform correlator by using the compressed low-contrast human face has broader impulse response. As a consequence, the correlation output of the low-contrast human face is less affected by noise than that of the low-contrast fingerprint.

Figure 4.10 shows the variation of the normalized PCDs as a function of the
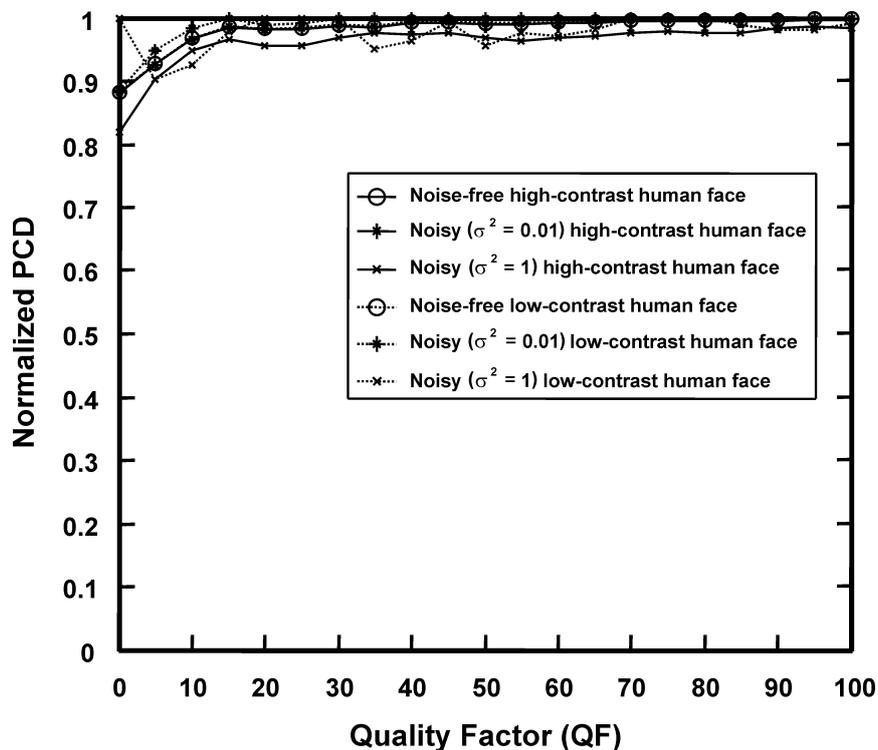
Figure 4.10 The PCD-based recognition performance of the joint transform correlator as a function of the QF of the compressed low-contrast human face reference.

QF of the compressed low-contrast human face for different target scenes. The normalized PCDs are greater than 0.9 and almost constant for all compression levels, except when the QF is less than 20. The results also show that the recognition performance of the joint transform correlator by using the compressed low-contrast human face as the reference is affected by the noise. Since the CR of the low-contrast human face is the highest compared with the others, therefore, it is worth mentioning that although by using a small file size of the compressed reference, the joint transform correlator gives a good correlation performance.

4.3   Experimental Verifications

Experimental verifications of the joint transform correlator with compressed reference images were performed by using the optical setup shown in Fig 4.11. A He-Ne laser Uniphase: 1507P-0 operating at wavelength of 632.8 nm was used as a coherent light source. A combination of a spatial filter and a beam expander consisting of 20x microscope objective lens ($f = 8.3$ mm), $25\,\mu$m pinhole aperture and a collimating lens ($f = 300$ mm), was used to produce a plane wave with a diameter of about 36 mm. A twisted-nematic liquid crystal display Jenoptik SLM-M/460 having resolution of 832×624 pixels with pixel aperture of $27 \times 23\,\mu$m, pixel pitch of $32 \times 32\,\mu$m, and a contrast ratio of 200 was used as the EASLM. Because its contrast ratio is low, the EASLM cannot display efficiently a small variation of gray level images. As a consequence, the optical Fourier spectrum of the images cannot be faithfully generated. An eight-bit CCD sensor PULNiX TM-2016-8 having pixels resolution of 1920×1080, pixel size of $7.4 \times 7.4\,\mu$m, and pixel pitch of $7.4 \times 7.4\,\mu$m was used to capture the generated joint power spectrum. The focal length of the

Fourier transform lens was chosen by considering non-overlapping condition of the adjacent replicas of the correlation output given by Eq. (2-31). By substituting the corresponding values into Eq. (2-31), the required focal length is found to be $f \geq 278$ mm. The experiment used a lens with focal length $f = 300$ mm.
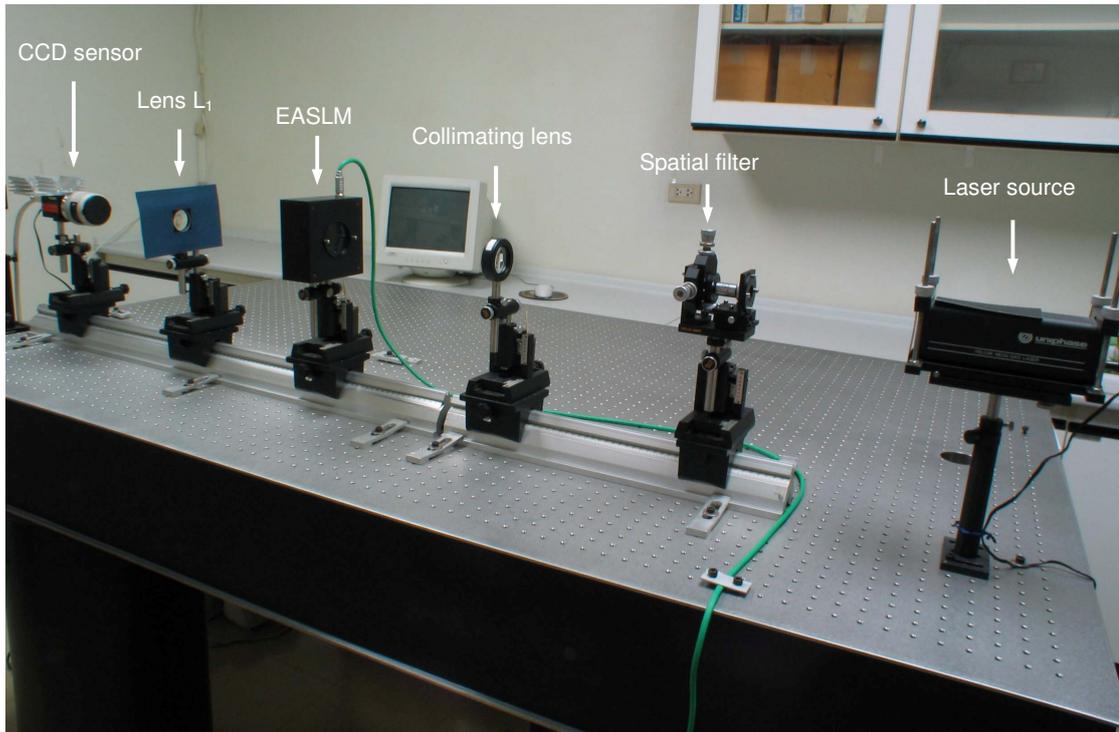
Figure 4.11   Actual optical setup for experimental verifications of the real-time joint transform correlator with compressed reference images.

As a preliminary study, the implementation of the joint transform correlator was done by using two identical uncompressed high-contrast fingerprint images with the size of $124P_{ex} \times 186P_{ey}$ as the reference and the target images. They were separated by the distance $2x_0 = 248P_{ex}$. The generated joint power spectrum was captured by the CCD sensor. In order to record faithfully the joint power spectrum, a neutral

density (ND) filter with an optical density = 2.0 was installed in front of the CCD

sensor, because the dynamic range of the sensor is limited. Saturation of the sensor by

high intensity of light clips the cosine fringes of the joint power spectrum. The

captured joint power spectrum is shown in Fig. 4.12. It is clear that, there are three

bright spots along the horizontal direction. The brightest spot corresponds to the zero

order of the joint power spectrum, while the leftmost and the rightmost spots are the

$(-1,0)^{th}$ and the $(+1,0)^{th}$ orders, respectively. They are separated from the zero order

by the distance $\dfrac{\lambda f}{P_{ex}}$ in horizontal direction. The intensity of the zero order of the joint

power spectrum is brighter than that of the $(-1,0)^{th}$ and the $(+1,0)^{th}$ orders, because of

the intensity attenuation by the envelope sinc function. Furthermore, the generation of

vertically oriented fringes caused by the interference between the target beam and the

reference beam can be clearly observed from Fig. 4.12(b) which shows the enlarged

zero order of the joint power spectrum. The fringes are obvious, because the spatial

resolution of the CCD sensor is finer than the fringe period that is $\dfrac{\lambda f}{MP_{ex}} = 23.92\ \mu\text{m}$.

By redisplaying the zero order of the joint power spectrum on the EASLM and

taking its optical Fourier transform, the correlation output is obtained at the back focal

plane of the lens L_1. Figure 4.13 shows the zero, the $(-1,0)^{th}$ and the $(+1,0)^{th}$ orders

of the correlation signals. They are separated by $\dfrac{\lambda f}{P_{ex}}$ in the x direction. In the

detection of the correlation output by the CCD sensor, the ND filter with an optical

density of 3.0 was used to prevent clipping of the correlation peak intensities. The
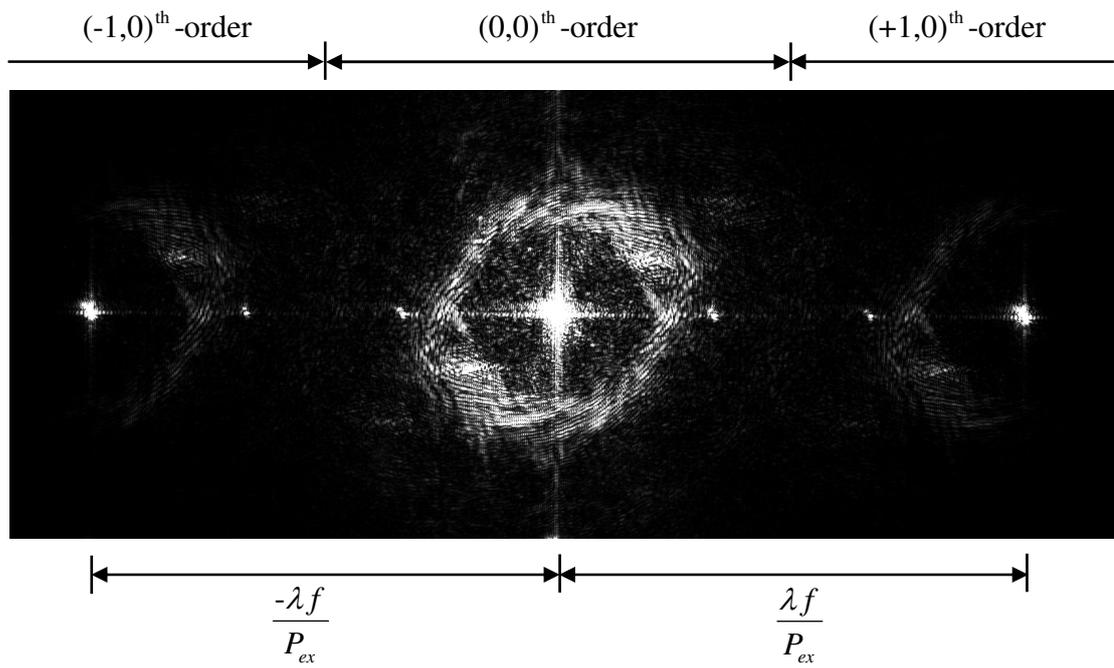
(a)
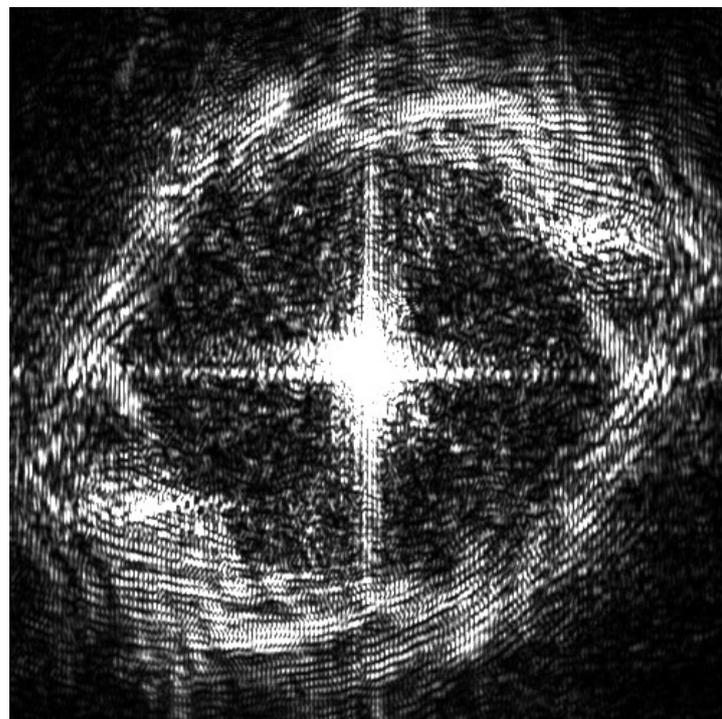
(b)

Figure 4.12 (a) joint power spectrum of two identical uncompressed high-contrast

fingerprints and (b) its enlarged zero order of the joint power spectrum.

zero order contains the DC signal corresponding to the two autocorrelation signals and the two identical cross-correlation signals. The separation of the autocorrelation and the cross-correlation signals is equal to MP_{CCDx}. The quantitative evaluations by using the PCD were then done on one of the cross-correlation output confined in the area of $2N_x P_{CCDx} \times 2N_y P_{CCDy}$. Based on this experimental setup, the fingerprint detections by using the joint transform correlator with compressed reference images were experimentally verified.
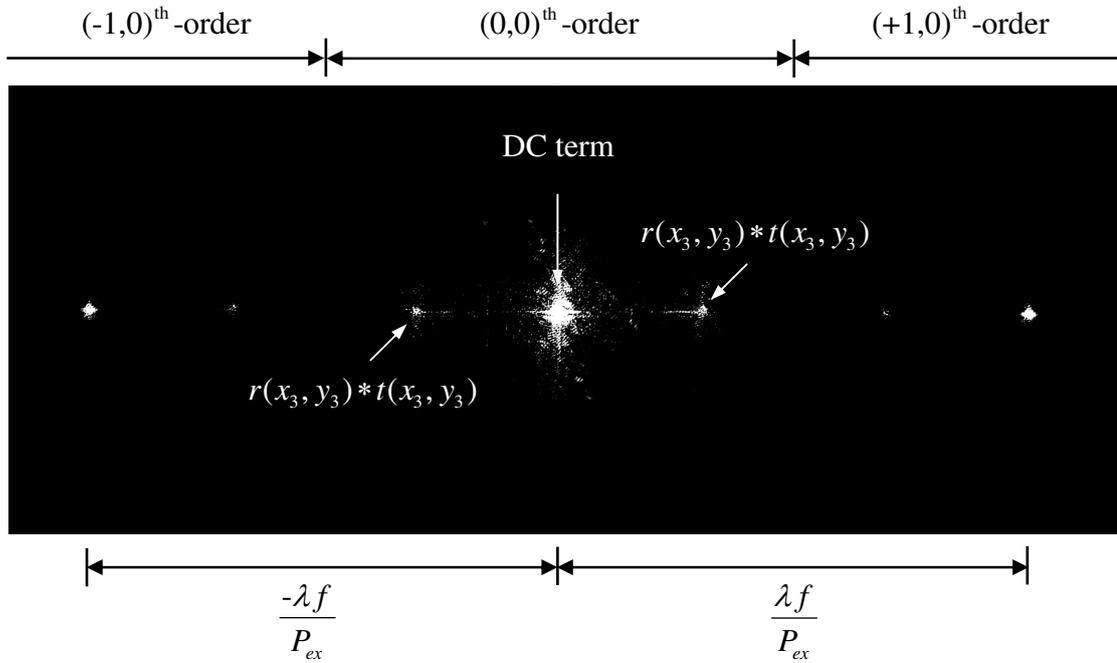
Figure 4.13   Autocorrelation output of uncompressed high-contrast fingerprint.

As for the high-contrast human face, the joint power spectrum of the two identical uncompressed images is shown in Fig. 4.14(a). Following the previous discussion, the zero order of the joint power spectrum is separated from the higher orders by a distance $\dfrac{\lambda f}{P_{ex}}$. This joint power spectrum was directly captured by the CCD sensor without the use of the ND filter, because the use of the filter suppresses

(a)
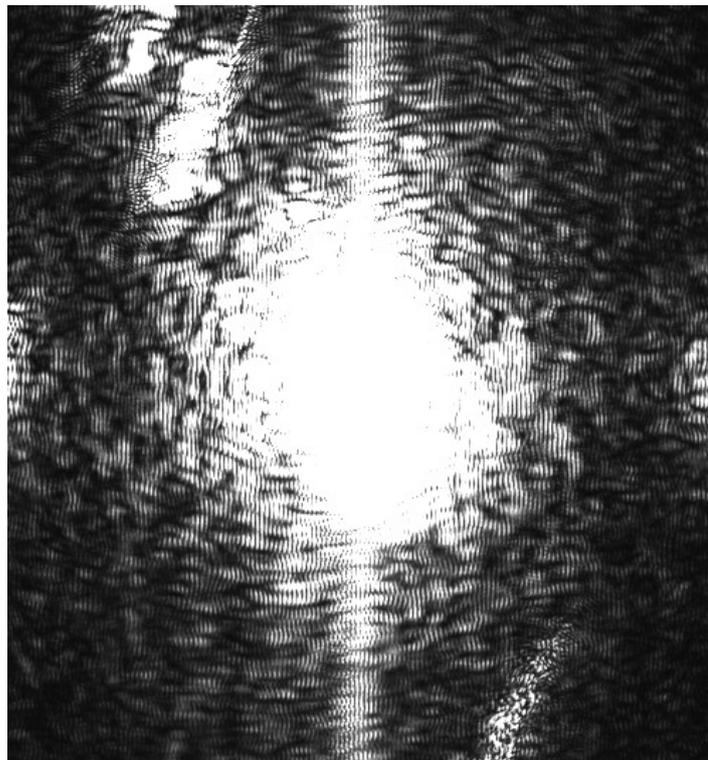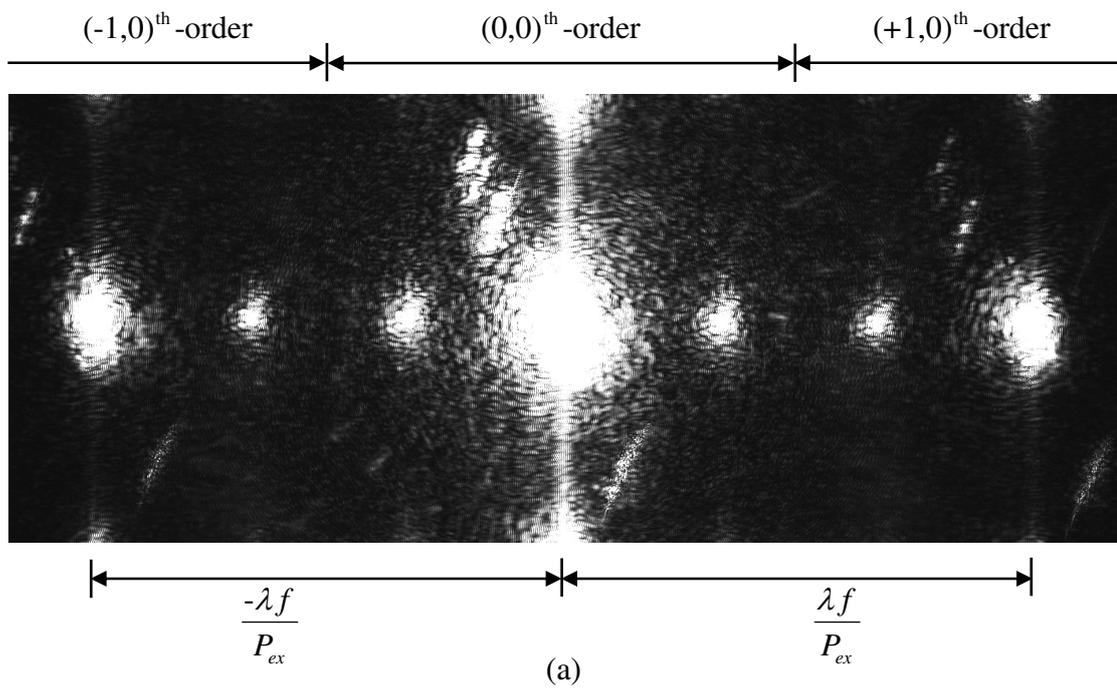
(b)

Figure 4.14 (a) joint power spectrum of two identical uncompressed high-contrast human faces and (b) its enlarged zero order of the joint power spectrum.

the high frequency spectra carried by the cosine fringes. Consequently, the correlation output will not be generated. Note that as discussed in Sect. 3.3, the amplitude of the high frequency components of the human face image is lower than the fingerprint image. Therefore, the amplitude of the fringes modulates by the high frequency spectra is weak. Figure 4.14(b) shows the enlarged zero order of the joint power spectrum. In comparison with Fig. 4.12(b), this joint power spectrum has the same spacing of the fringes, however the spread of the spectra around the origin, which corresponds to the DC signal and the low frequency spectra, is larger. Since their intensities exceed the maximum dynamic range of the CCD sensor, they are clipped.

Figure 4.15 shows the correlation output of this joint power spectrum which was generated by our optical setup shown in Fig. 4.11. This output was captured by using the ND filter with the density = 4.0. In comparison with Fig. 4.13, the

$(-1,0)^{th}$-order $\qquad$ $(0,0)^{th}$-order $\qquad$ $(+1,0)^{th}$-order
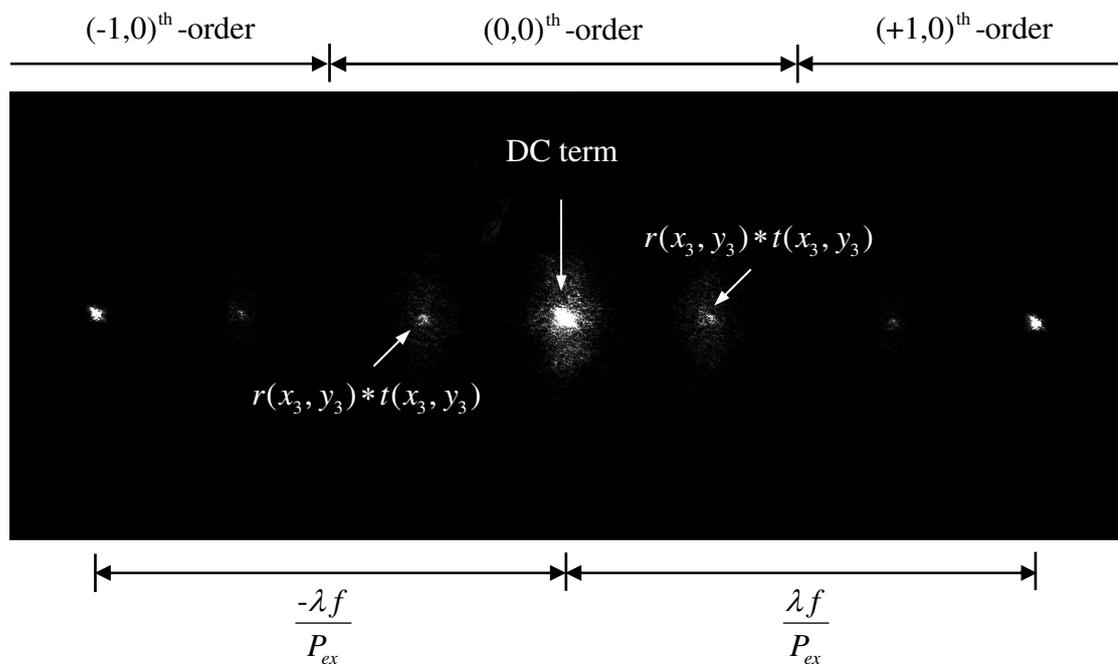
Figure 4.15 Autocorrelation output of uncompressed high-contrast human face.

correlation plane appears to be corrupted by more noise that might be speckle noise. When the EASLM redisplayed the zero order of the joint power spectrum, the light modulating elements displaying clipped intensities of the joint power spectrum had maximum transmittance. Consequently, stronger light totally transmitted by the EASLM generated more speckle noise. Experimental verification of the human face detections were conducted by using this optical setup.

4.3.1 Compressed High-Contrast Fingerprint as the Reference Images

Figures 4.16(a), (b), (c), and (d) show the 3-D output correlations of the high-contrast fingerprint as the reference images obtained from the experiment. The autocorrelation peak of the uncompressed high-contrast fingerprint shown in Fig. 4.16(a) is higher than the cross-correlation by using the compressed reference with QF = 10 shown in Fig. 4.16(b). This is caused by the degradation of the reference image by compression. This result is in agreement with the computer simulation depicted in Fig. 4.3. Figures 4.16(c) shows the correlation output obtained when the reference is the compressed high-contrast fingerprint with QF = 10 and the target is noisy high-contrast fingerprint. It can be seen from the figure that the correlation peak is buried in the noise. This can be explained by the fact that the amplitude of high spatial-frequency components of the noise are higher than that of the target and the reference images. The effect of intensity reduction by the ND filter on the high spatial-frequency components of the target and the reference is more significant than that of the noise. As a consequence, the CCD sensor captured the spatial frequency components of the noise, instead of the interference fringes generated by the spectra of the target and the reference images. As a result, the correlation signal was not
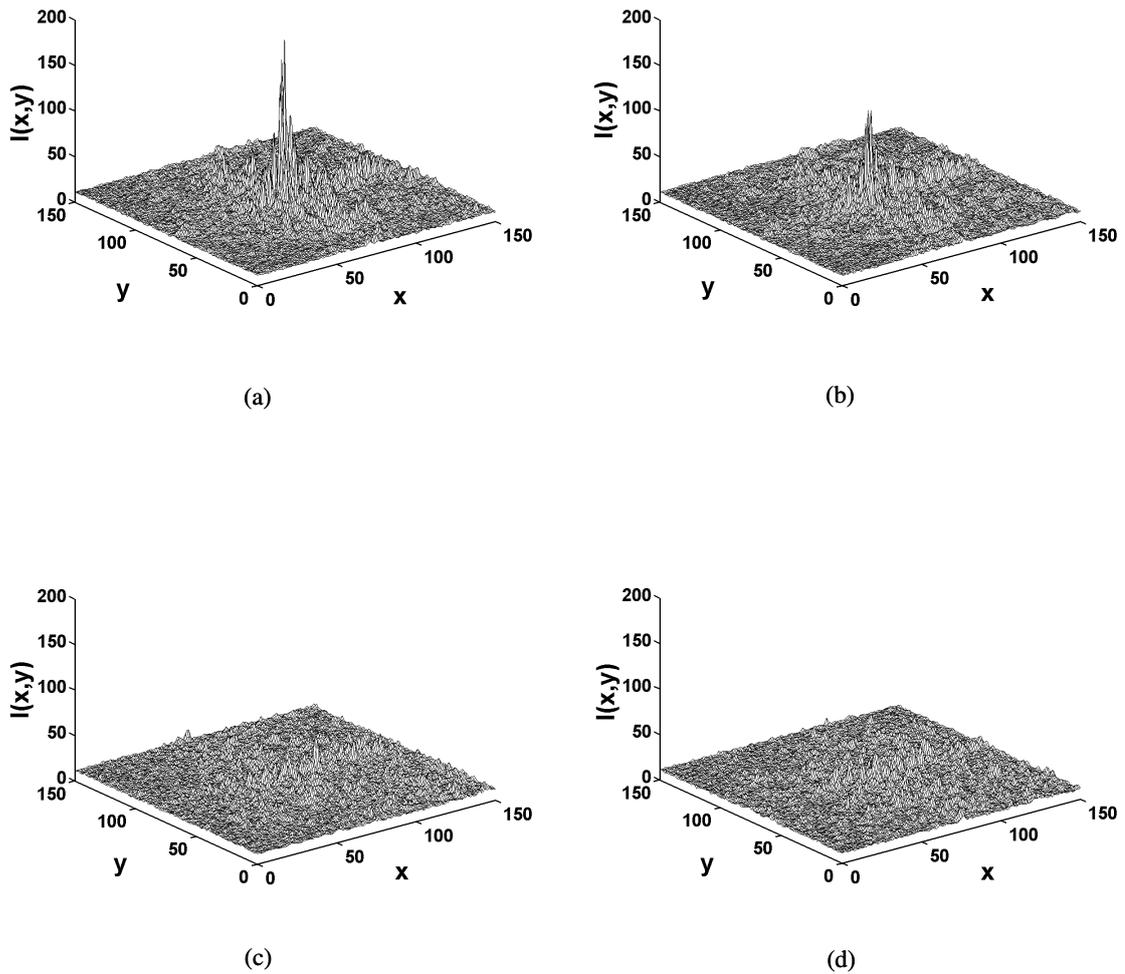
Figure 4.16 Experimental results of the joint transform correlator. (a) autocorrelation of the uncompressed high-contrast fingerprint; and cross-correlation outputs by using the compressed high-contrast fingerprint reference (QF = 10) under a situation that the target is: (b) noise-free high-contrast fingerprint, (c) noisy high-contrast fingerprint ($\sigma^2 = 1$), and (d) noisy low-contrast fingerprint ($\sigma^2 = 1$).

generated. Figure 4.16(d) shows the correlation output when the target is noisy low-contrast fingerprint with variance $\sigma^2 = 1$. There is no observable correlation peak generated in this detection, because of the low contrast ratio of the EASLM.

Figure 4.17 shows the variation of the normalized PCDs as a function of the QF of the compressed high-contrast fingerprint for different target scenes obtained from the experiment. The PCDs were computed by using the intensity value detected at the original position of the correlation peak. When the target is noise-free high-contrast fingerprint, the PCD increases as the QF increases. However when QF becomes larger than 50 the PCD fluctuates. This occurs because the images captured by the CCD sensor were shifted by a few pixels. This image shift may be caused by jitters in a frame-grabber system. In the case of capturing the correlation outputs, the spatial shift of the correlation peak from its original position introduces wrong computation of the PCD. In order to overcome this problem, an average of the correlation peak intensities is computed over 5×5 pixels neighborhood of the original correlation peak. Figure 4.18 shows the normalized PCDs by using the averaged correlation peak intensities. It is clear that the fluctuation of the PCDs becomes smaller than that of Fig. 4.17. The PCDs of the high-contrast fingerprint target without and with noise having variance $\sigma^2 = 0.01$ increase gradually as the QF increase. This is in agreement with the simulation results discussed in Sect. 4.2.1. However when the fingerprint target is corrupted by the strong noise, the PCD reduces drastically regardless of the compression quality. This indicates that the recognition performance of the joint transform correlator depends more on noise level than that on the compression.
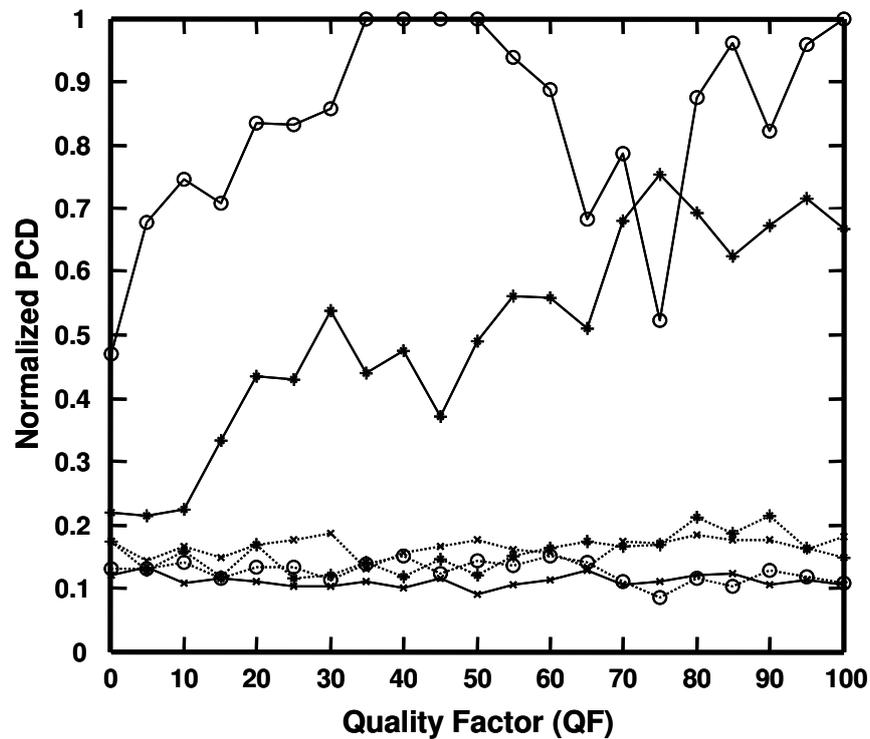
Figure 4.17 The variation of the normalized PCDs as a function of the QF of the compressed high-contrast fingerprint for different target scenes obtained from the experiment (—⊖—: noise-free high-contrast fingerprint target, —✳—: noisy high-contrast fingerprint target with $\sigma^2 = 0.01$, —✗—: noisy high-contrast fingerprint target with $\sigma^2 = 1$, ··⊙·· : noise-free low-contrast fingerprint target, ···✳···: noisy low-contrast fingerprint target with $\sigma^2 = 0.01$, and ···✗···: noisy low-contrast fingerprint target with $\sigma^2 = 1$).

In the case of the low-contrast target detection, all PCDs are low. This is because the limitation of the used EASLM in displaying efficiently low-contrast images. According to the characteristic of the EASLM (see Technical Documentation for Matrix Modulator SLM-M, Jenoptik, 19 September 1999), the transmission is
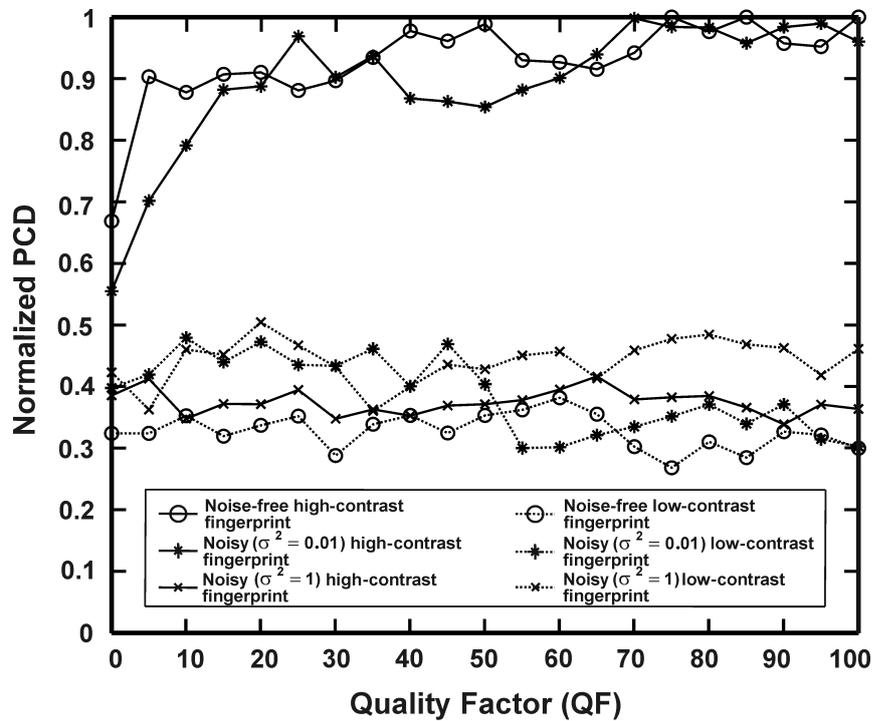
Figure 4.18 The variation of the PCDs as a function of the QF of the compressed high-contrast fingerprint by using the averaged intensities over 5×5 pixels neighborhood of the correlation peak.

maximum when its light-modulating elements are driven by signal with gray level amplitude greater than 210. Since the pixel values of the low-contrast fingerprint image are between 210 to 255, the spatial light modulation by the fingerprint image displayed on the EASLM cannot be done effectively. This can be verified from Fig. 4.19 which shows the zero order of the joint power spectrum of the detection of the low-contrast noise-free fingerprint target by the compressed reference with QF = 100. In comparison with Fig. 4.12(b), it is obvious that the fringes appear only at the cross pattern which corresponds to the power spectra of the backgrounds of the target and the reference images.
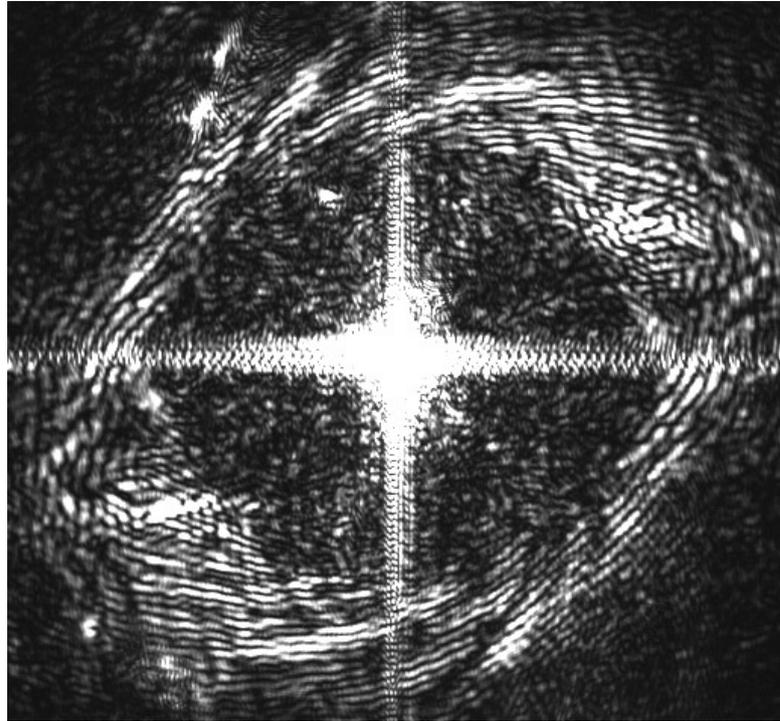
Figure 4.19   Zero order of the joint power spectrum of the low-contrast fingerprint

target.

4.3.2   Compressed Low-Contrast Fingerprint as the Reference Images

The experimental verifications of the fingerprint detection by using the joint

transform correlator with compressed low-contrast reference could not be successfully

accomplished, because the EASLM cannot modulate totally the incident light. Figures

4.20(a) and (b) show the joint power spectrum of the low-contrast fingerprint

generated by the digital and the optical computations, respectively. The digitally

computed joint power spectrum is plotted in logarithmic scale. It is obvious that the

optically generated joint power spectrum does not contain the spectral pattern shown

in Fig. 4.20(a).

<div align="center">(a)   (b)</div>
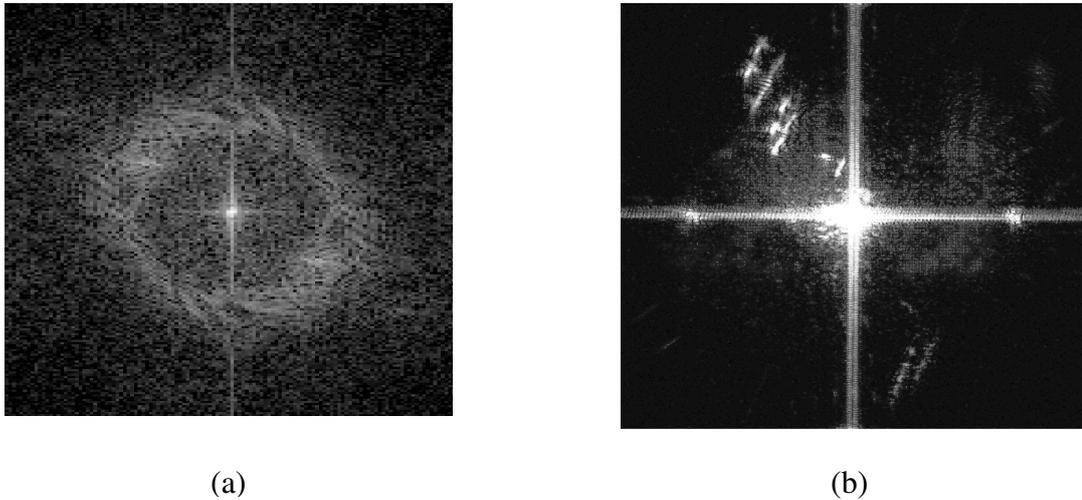
Figure 4.20   The joint power spectrum of low-contrast fingerprint obtained from (a) the simulation and (b) the experiment.

4.3.3   Compressed High-Contrast Human Face as the Reference Images

The 3-D plot of the autocorrelation of the uncompressed high-contrast human face illustrated in Fig. 4.15 is shown in Fig. 4.21(a), while the correlation outputs of the joint transform correlator by using the compressed high-contrast human face with QF = 10 as the reference are illustrated in Figs. 4.21(b), (c), and (d). Note that the experimentally generated correlation signals are not as broad as those obtained by the computer simulation shown in Fig. 4.7. This is because the maximum correlation peaks intensities were reduced by using the ND filter with density = 4.0 installed in the front of the CCD sensor. As a consequence, the peak height and width of the correlation signal were greatly reduced. When the target is noise free high-contrast human face, the correlation output shown in Fig. 4.21(b) is almost as sharp as the autocorrelation output shown in Fig. 4.21(a), however its peak slightly decreased. This is the effect of the compression on the reference image. Furthermore, the correlation planes shown in Figs. 4.21(a) and (b) contain stronger noise than that in
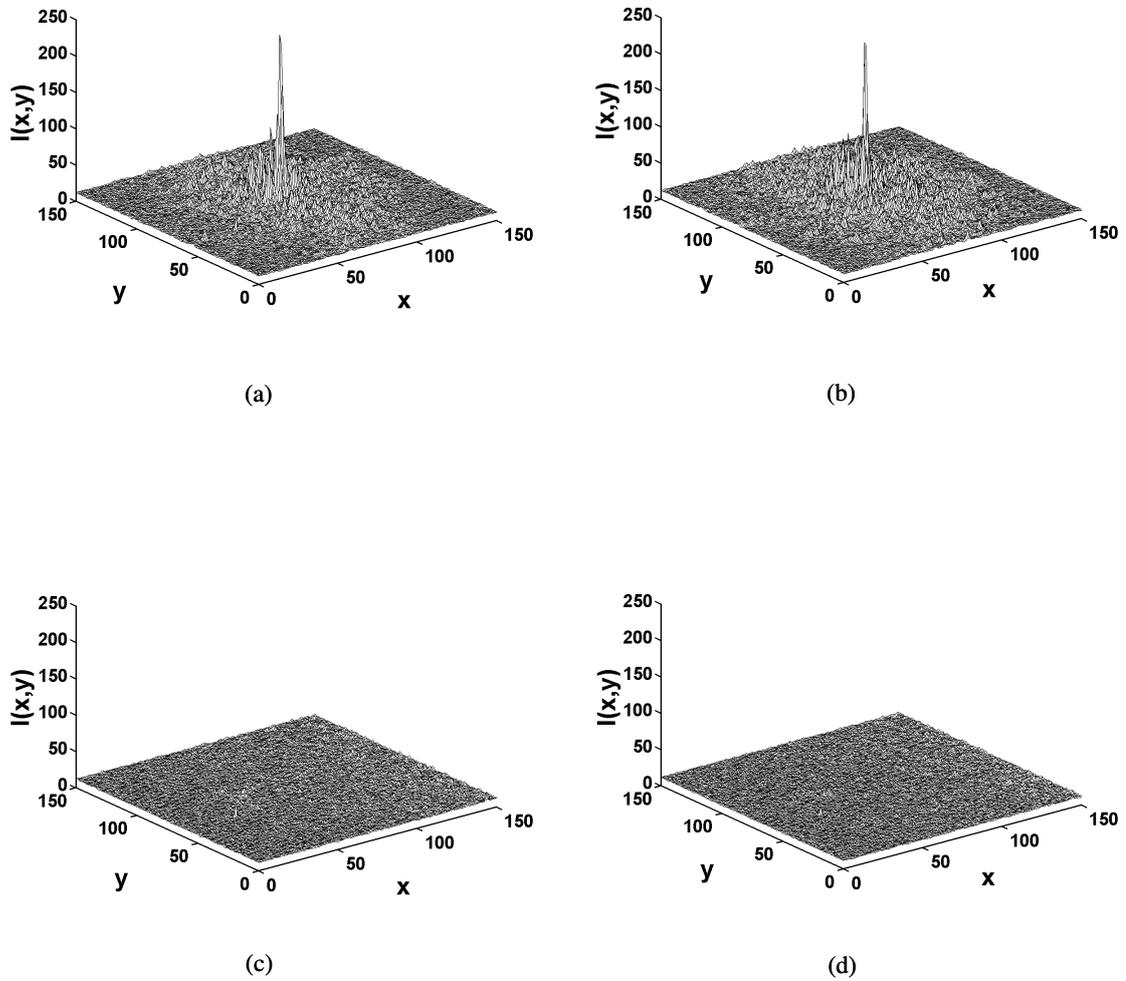
Figure 4.21 Experimental results of the joint transform correlator. (a) autocorrelation of the uncompressed high-contrast human face; and cross-correlation outputs by using the compressed high-contrast human face reference (QF = 10) under a situation that the target is: (b) noise-free high-contrast human face, (c) noisy high-contrast human face ($\sigma^2 = 1$), and (d) noisy low-contrast human face ($\sigma^2 = 1$).

Figs. 4.16(a) and (b). As discussed in Sect. 4.3, this is mainly caused by more speckle noise which is generated by light totally transmitted through the EASLM. Figures 4.21(c) and (d) show the correlation output of the noise-corrupted high- and low-contrast human face targets with the noise variance $\sigma^2 = 1$, respectively. It is obvious that correlation peaks could not be observed. This is caused by the same reasons discussed in Sect. 4.3.1.

Figures 4.22(a) and (b) show the normalized PCDs as a function of the QF of the compressed high-contrast human face for difference target scenes obtained without and with averaging of intensities over 5×5 pixels neighborhood of the original correlation peak. Figure 4.22(b) shows that besides reducing the fluctuation of the PCDs, the averaging of the correlation peaks increases the value of the PCDs, because the averaged correlation peak may become approximately equal to the correct value. When the target is noise-free high-contrast human face, the PCD decreases as the QF becomes smaller. However, when the QF is less than 10, the PCDs decrease rapidly. This is due to the fact that the blocking artifacts carry high-spatial frequency information. Therefore unlike the compressed fingerprint image, the generated blocking artifacts at low QF increase significantly the frequency content of the compressed human face reference. Since there might be frequency content mismatch between the target and the reference, the correlation peaks reduce. As a result, the PCD decreases drastically. In the presence of noise with variance $\sigma^2 = 0.01$ in the input target, the PCDs drop below 0.5. This is because the joint power spectrum recorded without the use of the ND filter was corrupted by additional noise. Consequently, the correlation plane also appears noisy. When the noise presence is
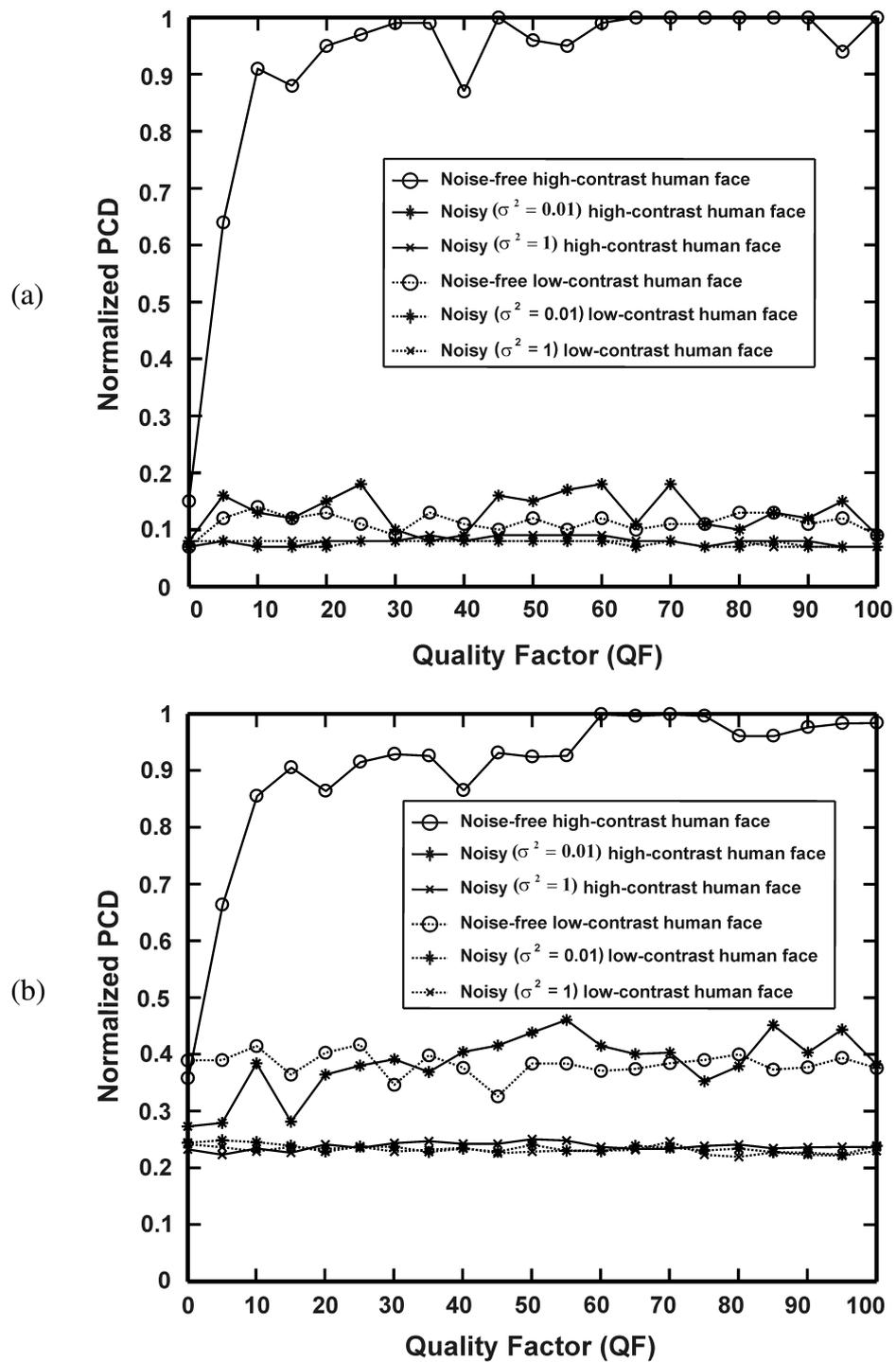
Figure 4.22 The variation of the normalized PCDs as a function of the QF of the compressed high-contrast human face for different target scenes: (a) without and (b) with averaging of intensities over 5×5 pixels neighborhood of the original correlation peak.

stronger, the PCDs decrease further. Because of the limitation of the EASLM, the detections of the low-contrast human face targets give low PCDs.

4.3.4 Compressed Low-Contrast Human Face as the Reference Images

The experimental verifications of the joint transform correlator with the compressed low-contrast human face as the reference could not also be accomplished, because of the limitation of the EASLM. This is confirmed by comparing Fig. 4.23(a) with Fig. 4.23(b) which shows the difference of the joint power spectrum of the low-contrast human face obtained from the computer simulation and the experiment, respectively. It is obvious that the joint power spectrum obtained from the experiment has no spectral information as the digitally generated joint power spectrum.

(a)   (b)
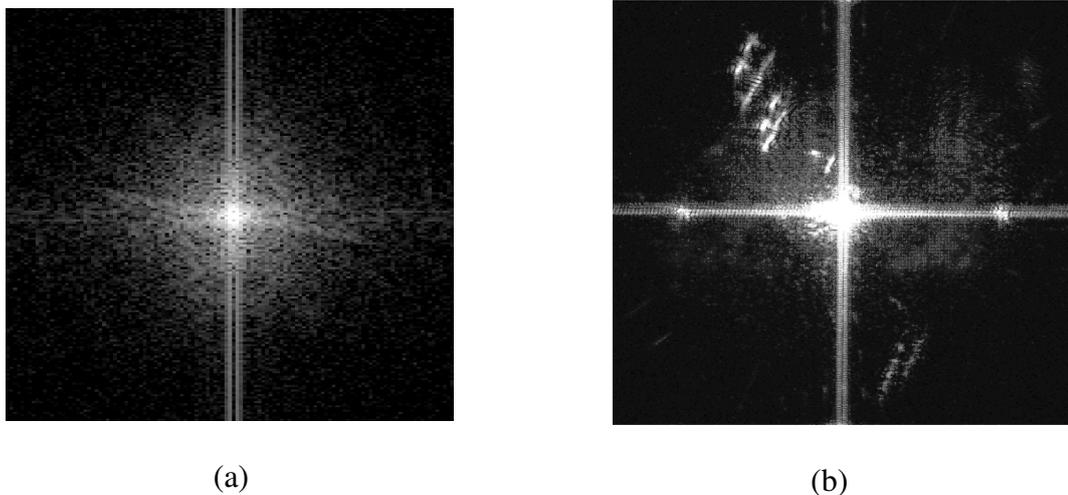
Figure 4.23   The joint power spectrum of low-contrast human face obtained from (a) the simulation and (b) the experiment.

CHAPTER V

MULTIPLE-TARET DETECTION

In Chapter IV, single-target detection by using the joint transform correlator with JPEG-compressed reference images was studied and demonstrated. The system employed two successive optical Fourier transformations to compute the correlation of the target and the compressed reference images. In order to investigate the effects of reference compression on the recognition of the joint transform correlator on using compressed reference images, the correlation output was evaluated by measuring the PCD. To further verify the feasibility of our proposed method, this chapter studies the effects of JPEG-compression of the reference image on the performance of multiple-target detection via both computer simulation and experimental verifications. In this study, the detection performance is measured by using the desired primary correlation peak to the secondary peak intensity ratio (PSR).

5.1 Multiple-Target Detection by Using the Joint Transform Correlator with Compressed Reference Images

The real-time multiple-target detection by using the joint transform correlator with compressed reference images can be implemented by using the optical setup shown in Fig. 2.3. By displaying the compressed reference $r_c(x_1, y_1)$ and the input scene consisting of N target images $t_i(x_1, y_1)$ side-by-side with a separation of $2x_0$ onto the EASLM, the joint input image can be mathematically written as

$$f_{JTC}(x_1, y_1) = r_c(x_1 - x_0, y_1) + \sum_{i=1}^{N} t_i(x_1 + x_i + x_0, y_1 + y_i), \qquad (5\text{-}1)$$

where x_i and y_i correspond to the relative position of the target in the x_1 and the y_1 directions, respectively. When the additive white Gaussian noise $n(x_1, y_1)$ is present at the input scene and there is a contrast difference between the target and the compressed reference images, the joint input image can be rewritten as

$$f_{JTC}(x_1, y_1) = r_c(x_1 - x_0, y) + c_T \sum_{i=1}^{N} t_i(x_1 + x_i + x_0, y_1 + y_i) + n(x_1 + x_0, y_1). \qquad (5\text{-}2)$$

After a Fourier transformation by the lens L_1, the joint power spectrum captured by the CCD sensor is found to be

$$
\begin{aligned}
U(x_2, y_2) = & \left| R_c\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) \right|^2 + c_T^2 \sum_{i=1}^{N} \left| T_i\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) \right|^2 + \left| N\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) \right|^2 \\
& + c_T N\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) \left[\sum_{i=1}^{N} T_i^*\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) e^{-j\frac{2\pi}{\lambda f}(x_i x_2 + y_i y_2)} \right] \\
& + c_T N^*\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) \left[\sum_{i} T_i\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) e^{j\frac{2\pi}{\lambda f}(x_i x_2 + y_i y_2)} \right] \\
& + c_T^2 \left[\sum_{i=1}^{N} \sum_{\substack{k=1 \\ k \neq i}}^{N} T_i^*\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) T_k\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) e^{-j\frac{2\pi}{\lambda f}[(x_i - x_k)x_2 + (y_i - y_k)y_2]} \right] \\
& + c_T^2 \left[\sum_{i=1}^{N} \sum_{\substack{k=1 \\ k \neq i}}^{N} T_i\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) T_k^*\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) e^{j\frac{2\pi}{\lambda f}[(x_i - x_k)x_2 + (y_i - y_k)y_2]} \right] \\
& + c_T R_c\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) \left[\sum_{i=1}^{N} T_i^*\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) e^{-j\frac{2\pi}{\lambda f}[(x_i + 2x_0)x_2 + y_i y_2]} \right] \\
& + c_T R_c^*\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) \left[\sum_{i=1}^{N} T_i\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) e^{j\frac{2\pi}{\lambda f}[(x_i + 2x_0)x_2 + y_i y_2)]} \right] \\
& + R_c\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) N^*\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) e^{-j\frac{4\pi}{\lambda f}x_0 x_2}
\end{aligned}
$$

$$+ N\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) R_c^*\left(\frac{x_2}{\lambda f}, \frac{y_2}{\lambda f}\right) e^{j\frac{4\pi}{\lambda f}x_0 x_2}, \qquad (5\text{-}3)$$

where $T_i\left(\dfrac{x_2}{\lambda f}, \dfrac{y_2}{\lambda f}\right)$ is the Fourier transform of the ith target. The first three terms of

Eq. (5-3) 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 joint power spectrum 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:

$$\begin{aligned}
C(x_3, y_3) = r_c(x_3, y_3) * c_T \sum_{n=1}^{N} t_i(x_3, y_3) \otimes \delta\left[x_3 \pm (x_i + 2x_0), y_3 \pm y_i\right] \\
+ r_c(x_3, y_3) * n(x_3, y_3) \otimes \delta(x_3 \pm 2x_0, y_3).
\end{aligned} \qquad (5\text{-}4)$$

The first term of Eq. (5-4) 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.

5.2   Computer Simulation

The effects of the reference image compression on the multiple-target detection using the joint transform correlator were studied by using the same fingerprint and human face images. The input target scene contained two different images with one of them identical to the original reference image. Figures 5.1(a) and (b) show the multiple target and the compressed reference with QF = 10 for the high-contrast fingerprint and the human face, respectively. Their low-contrast images are shown in Figs. 5.1(c) and (d). The JPEG-compressed reference images were generated by using the same ACDsee software (The 2000 ACD systems, Ltd.).

In the simulation, the multiple-object input scene was first generated and then was combined with compressed reference images 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. In the simulation, the same algorithm for computing the joint transform correlator used in Chapter IV was employed. However, the correlation output of the multiple-target detection was quantified by measuring the PSR in the correlation plane.

5.2.1   Compressed High-Contrast Fingerprint as the Reference Images

The 3-D plot of the correlation outputs of the multiple-fingerprint detection by using the joint transform correlator with the high-contrast fingerprint reference are shown in Fig. 5.2. The correlation output shown in Fig. 5.2(a) consists of two peaks, because there are two targets. The primary peak 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
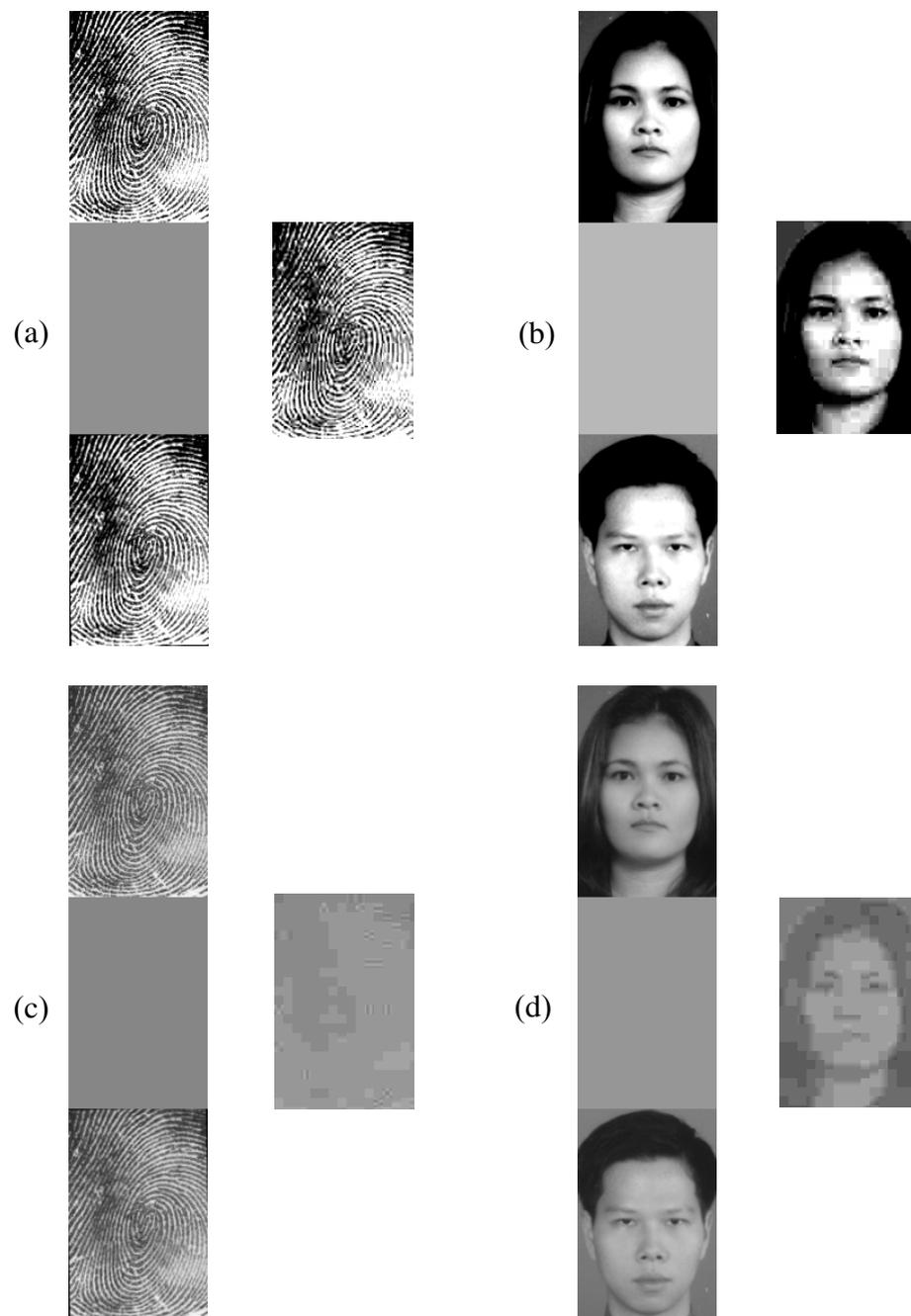
Figure 5.1 Multiple target input scenes and the compressed reference with the QF = 10: (a) high-contrast fingerprint, (b) low-contrast fingerprint, (c) high-contrast human face, (d) low-contrast human face.
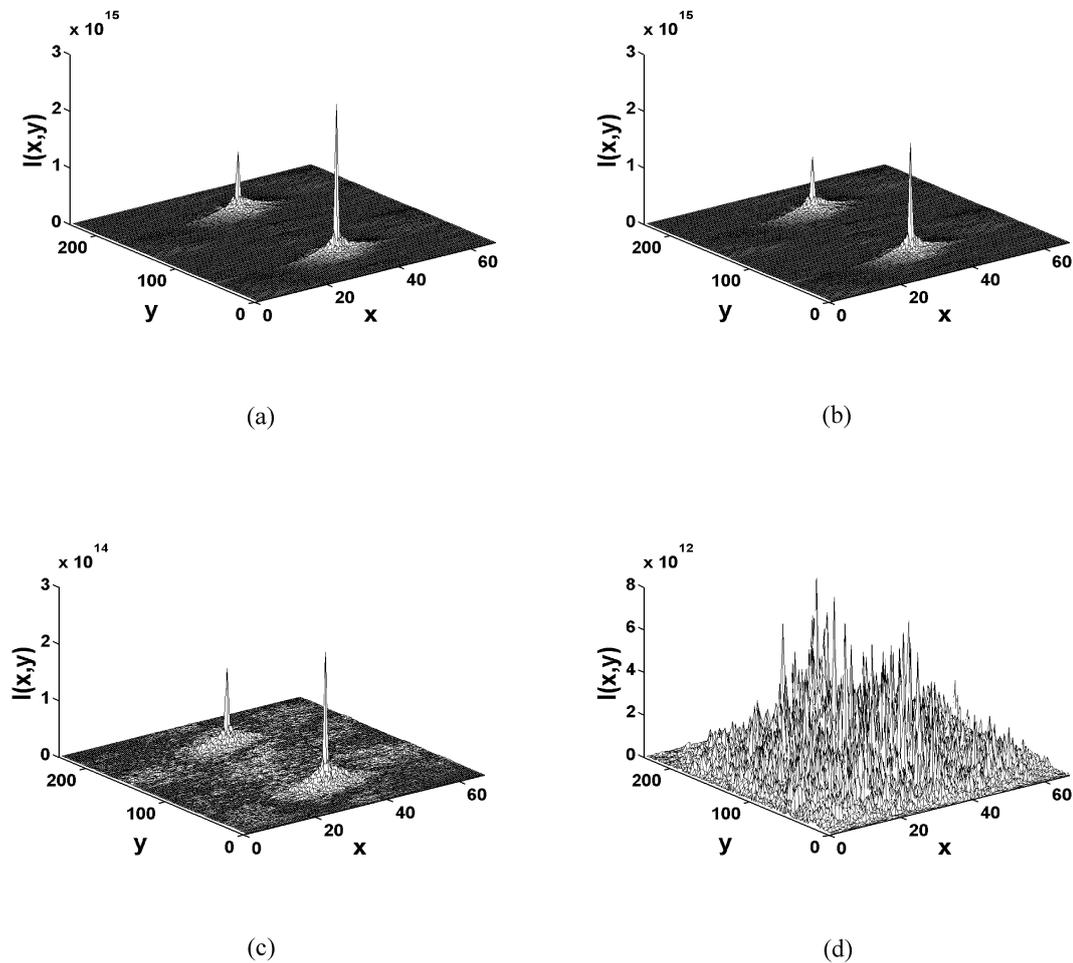
Figure 5.2 Simulation results of the multiple-target joint transform correlator. (a) autocorrelation of the uncompressed high-contrast fingerprint; and cross-correlation outputs by using the compressed high-contrast fingerprint reference (QF = 10) under a situation that the multiple-target scene is: (b) noise-free high-contrast fingerprint image, (c) noisy high-contrast fingerprint image ($\sigma^2 = 1$), and (d) noisy low-contrast fingerprint image ($\sigma^2 = 1$).

primary peak is higher and sharper than the undesired secondary peak. Figures 5.2(b), (c), and (d) show the correlation outputs of the joint transform correlator by using the compressed high-contrast fingerprint at QF = 10 as the reference. In Fig. 5.2(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. 5.2(a). This is the effect of the compression on the reference image (Widjaja, 2003; Widjaja and Suripon, 2004). Figures 5.2(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 (Widjaja and Suripon, 2004). This produces noisy correlation plane. 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.

Figure 5.3 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.
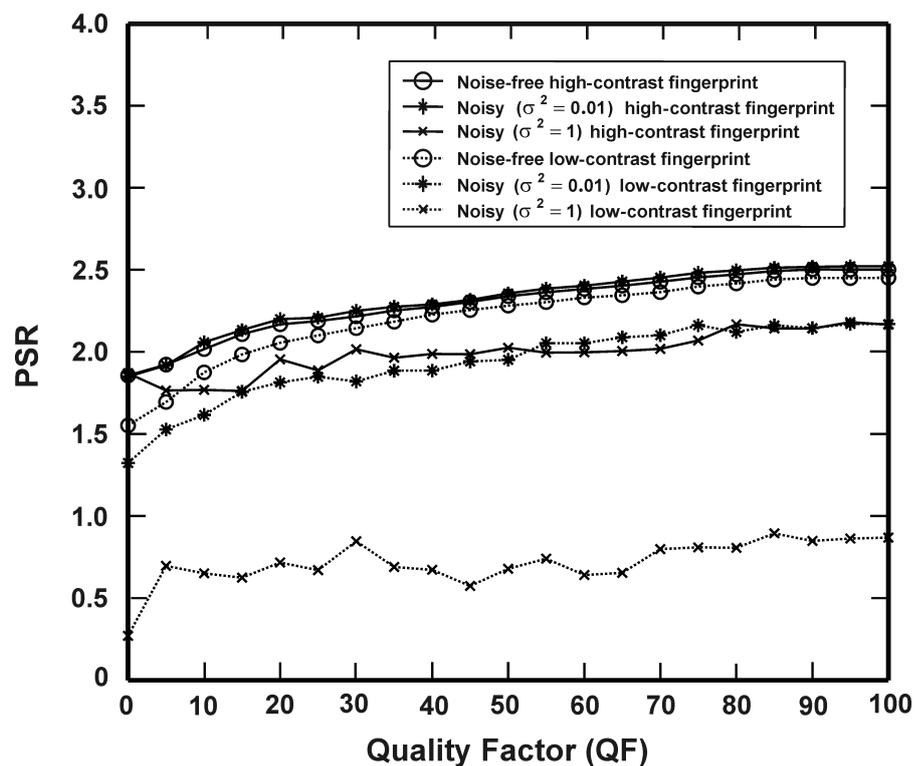
Figure 5.3 The PSR-based measurement of the detection performance of the joint transform correlator as a function of the QF of the compressed high-contrast fingerprint reference.

5.2.2 Compressed Low-Contrast Fingerprint as the Reference Images

Figures 5.4(a), (b), (c), and (d) illustrate the 3-D correlation outputs of the multiple-fingerprint detection by using joint transform correlator with the low-contrast fingerprint as the reference image. The autocorrelation peak of the uncompressed low-contrast fingerprint image shown in Fig. 5.4(a) is almost as sharp as the autocorrelation of the high-contrast target shown in Fig. 5.2(a). However, its peak intensity reduces by three orders of magnitude. The secondary peak is also lower and broader than that of the high-contrast fingerprint. Figure 5.4(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 joint transform correlator with this compressed reference image becomes broad. Figure 5.4(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. Figures 5.4(c) and (d) show 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
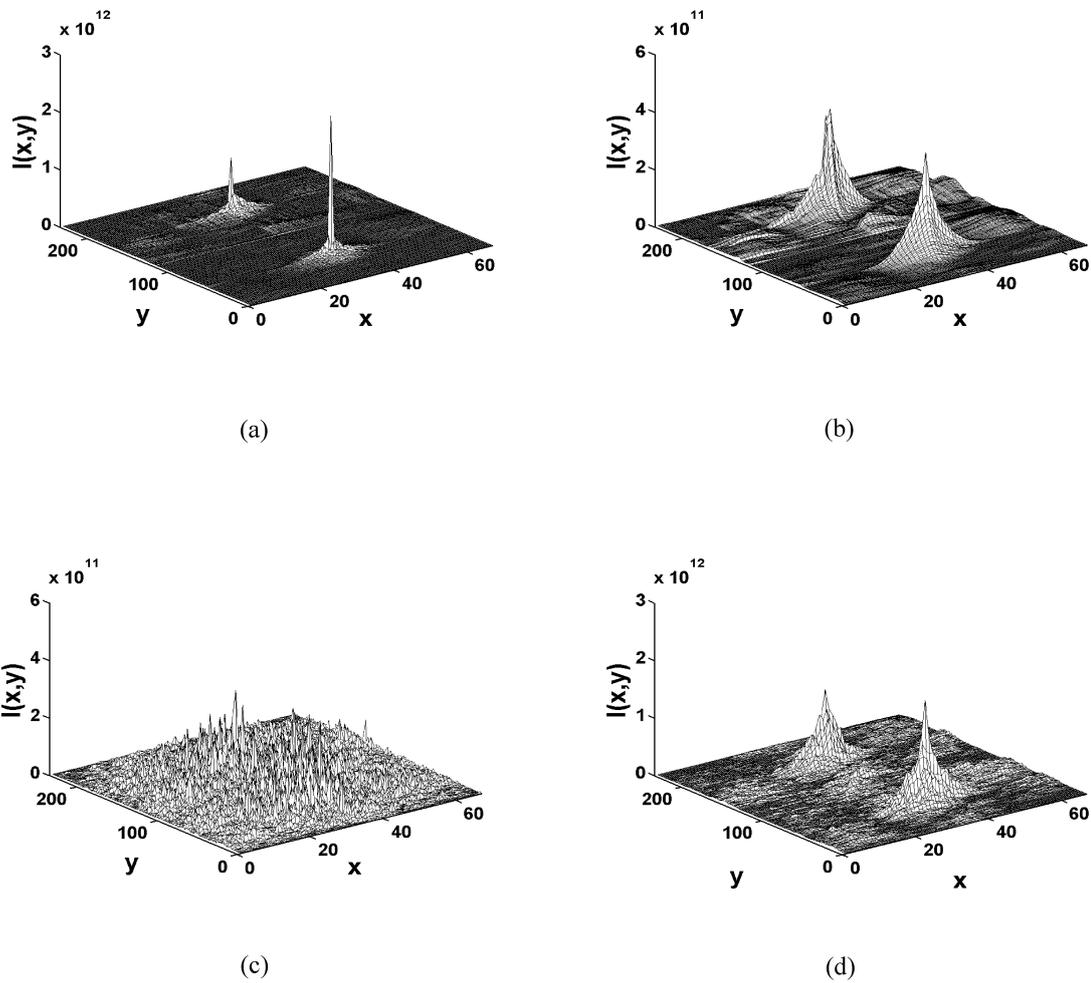
Figure 5.4 Simulation results of the multiple-target joint transform correlator. (a) autocorrelation of the uncompressed low-contrast fingerprint; and cross-correlation outputs by using the compressed low-contrast fingerprint reference (QF = 10) under a situation that the multiple-target scene is: (b) noise-free low-contrast fingerprint, (c) noisy low-contrast fingerprint ($\sigma^2 = 1$), and (d) noisy high-contrast fingerprint ($\sigma^2 = 1$).

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.

Figure 5.5 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
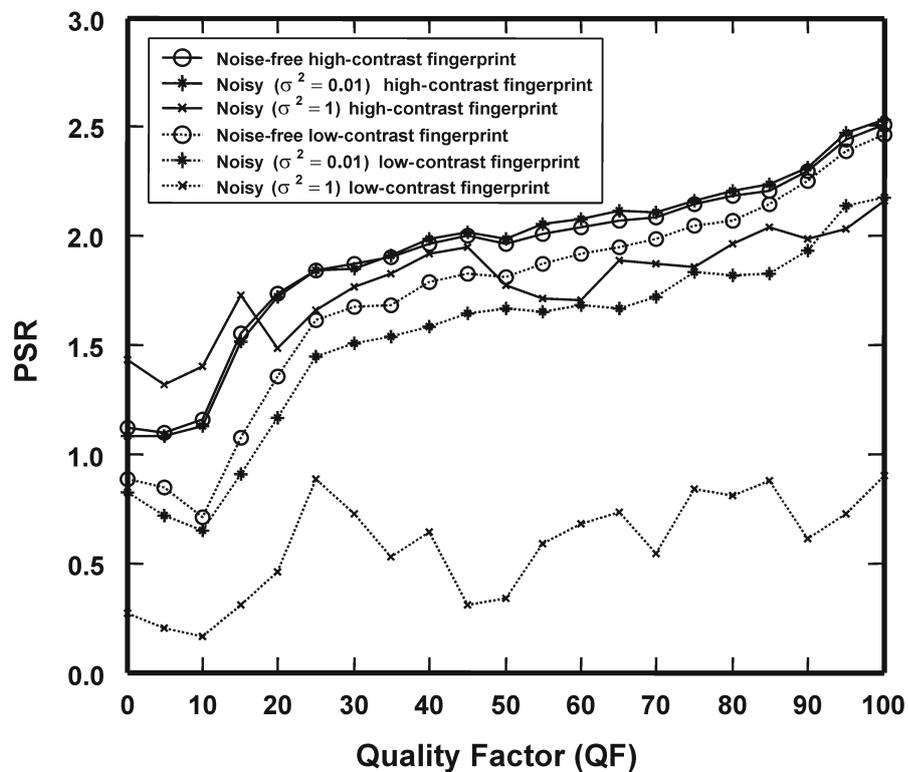
Figure 5.5 The PSR-based measurement of the detection performance of the joint transform correlator as a function of the QF of the compressed low-contrast fingerprint reference.

the QF increases, the PSRs increase more rapidly than that of the compressed high-contrast fingerprint shown in Fig. 5.3. This indicates that the use of the compressed low-contrast fingerprint reference degrades significantly the performance of multiple-target detection by the joint transform correlator. 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.

5.2.3 Compressed High-Contrast Human Face as the Reference Images

Figure 5.6(a) shows the 3-D plot of the output correlation of the high-contrast multiple-human face 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. Figures 5.6(b), (c), and (d) illustrate the correlation outputs of the joint transform correlator by using the compressed high-contrast human face as the reference at QF = 10 for different multiple-target scenes. Figure 5.6(b) illustrates the output detection of the noise-free high-contrast multiple-target which is slightly affected by the compression. Figures 5.6(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
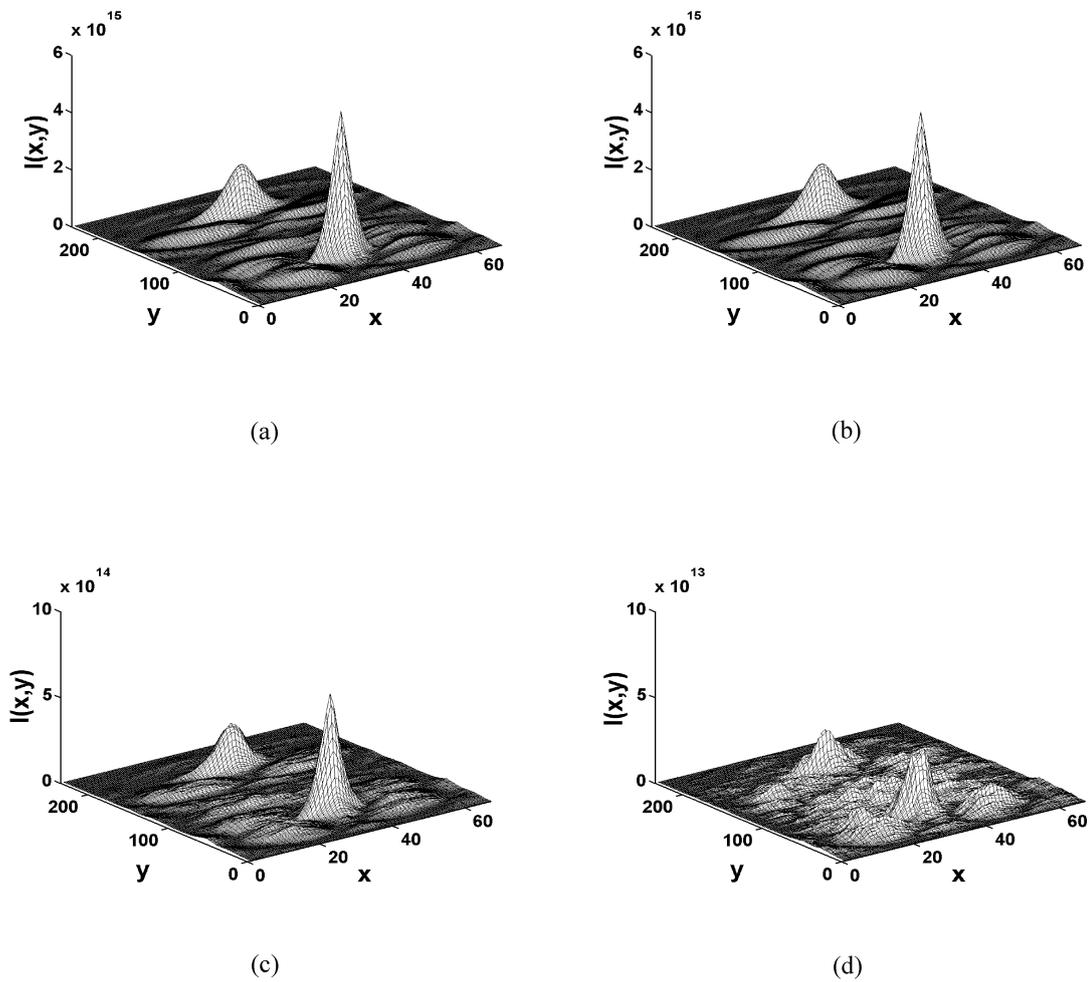
Figure 5.6   Simulation results of the multiple-target joint transform correlator. (a) autocorrelation of the uncompressed high-contrast human face; and cross-correlation outputs by using the compressed high-contrast human face reference (QF = 10) under a situation that the multiple-target scene is: (b) noise-free high-contrast human face, (c) noisy high-contrast human face ($\sigma^2 = 1$), and (d) noisy low-contrast human face ($\sigma^2 = 1$).

contrast, the peak intensities of Fig. 5.6(d) are lower and its correlation plane is more noisy than that of Fig. 5.6(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. 5.7. It is obvious that all PSRs are almost independent of the compression. Their magnitudes which are always greater than 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
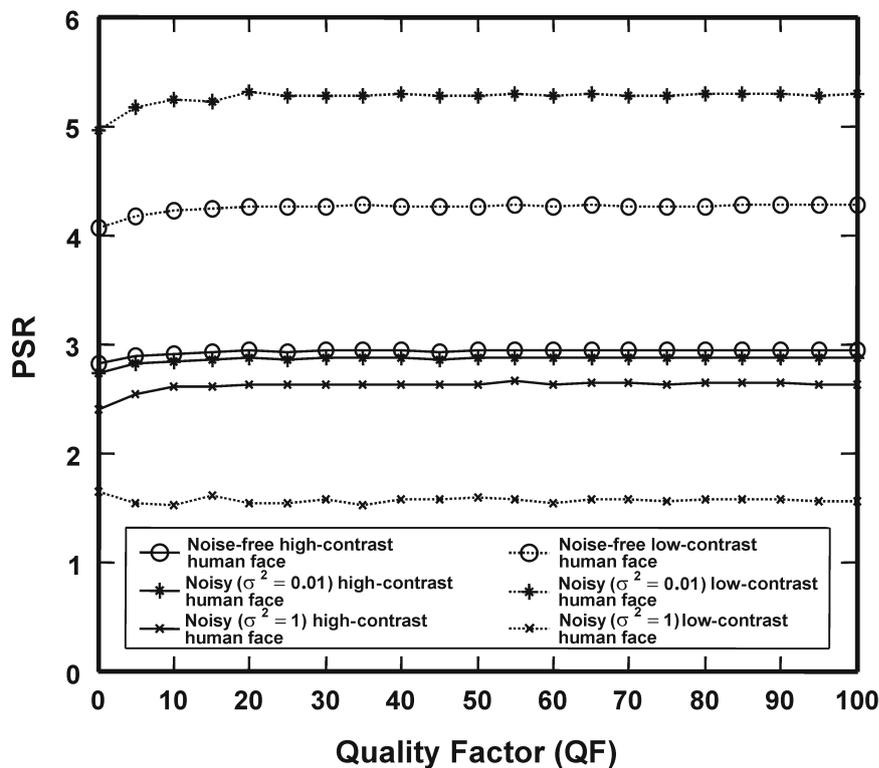
Figure 5.7 The PSR-based measurement of the detection performance of the joint transform correlator as a function of the QF of the compressed high-contrast human face reference.

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

5.2.4   Compressed Low-Contrast Human Face as the Reference Images

Figures 5.8(a), (b), (c), and (d) show the correlation outputs of the multiple-human face detection by using joint transform correlator with the low-contrast human face as the reference image. Figure 5.8(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. 5.8(b) further decreases, and their correlation profiles broaden. Figures 5.8(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

(a)

(b)

(c)

(d)
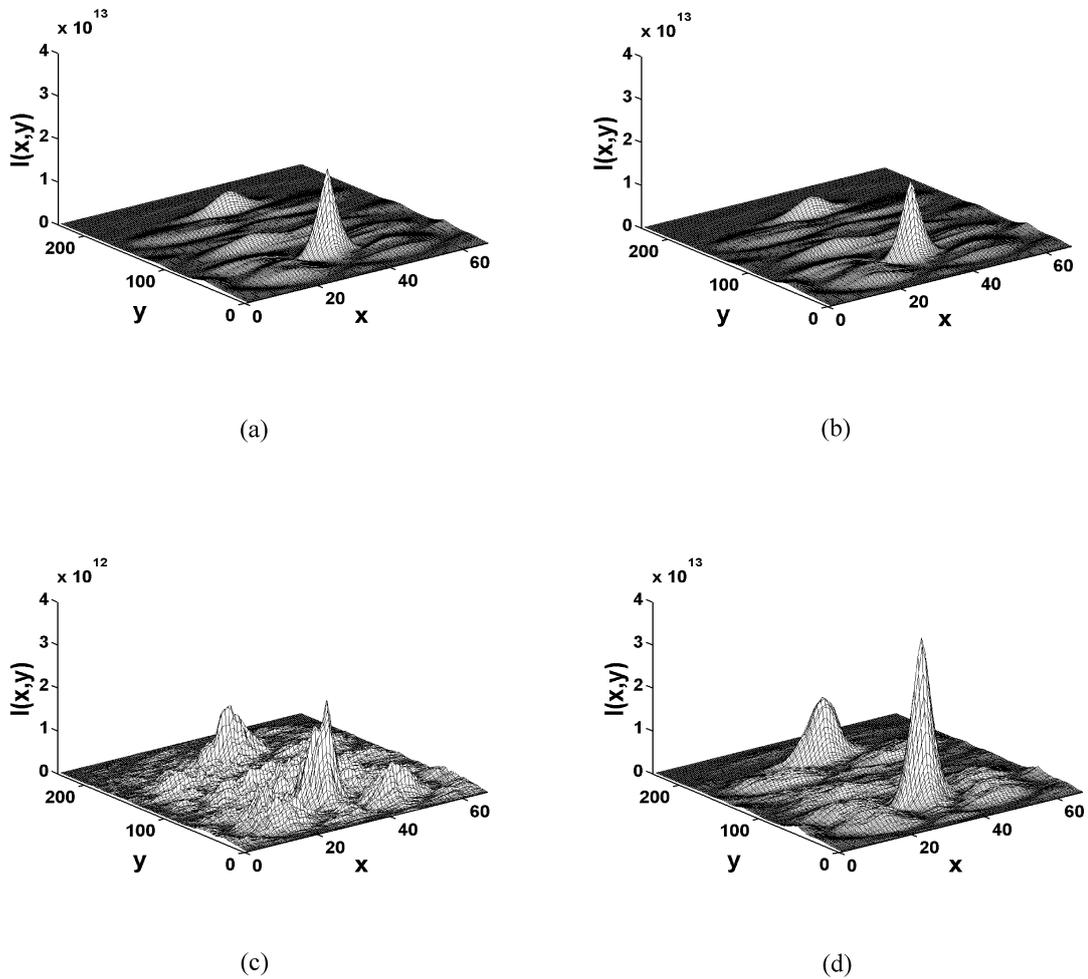
Figure 5.8   Simulation results of the multiple-target joint transform correlator. (a) autocorrelation of the uncompressed low-contrast human face; and cross-correlation outputs by using the compressed low-contrast human face reference (QF = 10) under a situation that the multiple-target scene is: (b) noise-free low-contrast human face, (c) noisy low-contrast human face ($\sigma^2 = 1$), and (d) noisy high-contrast human face ($\sigma^2 = 1$).

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. 5.8(a). This is because the image contrast is higher.

Figure 5.9 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.
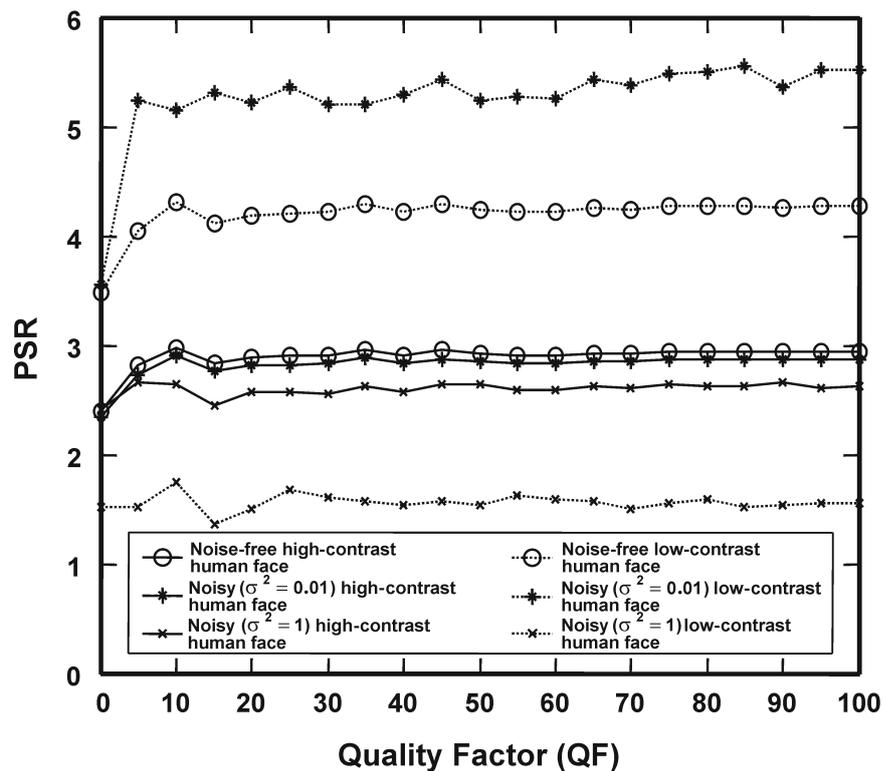
Figure 5.9   The PSR-based measurement of the detection performance of the joint transform correlator as a function of the QF of the compressed low-contrast human face reference.

5.3   Experimental Verifications

The experimental verifications of the multiple-object detection by using the joint transform correlator with compressed reference images were done by using the same optical setup employed in Chapter IV. As discussed in the preceding chapter, because of the limitation of the EASLM in displaying the low-contrast image, the experimental verifications of the joint transform correlator with compressed low-contrast reference image cannot be successfully performed.

The preliminary study of multiple-target detection was firstly done by using the high-contrast multiple-fingerprint target and the uncompressed high-contrast fingerprint as the reference. The target scene had a size of $124P_{ex} \times 558P_{ey}$, while the size of the compressed reference was $124P_{ex} \times 186P_{ey}$. To capture the generated joint power spectrum shown in Fig. 5.10(a) by the CCD sensor, the ND filter with density of 2.0 was used to reduce the light intensity. Three replications of the joint power spectrum that are the $(-1,0)^{th}$, the $(0,0)^{th}$, and the $(+1,0)^{th}$ orders can be observed from Fig. 5.10(a). The higher orders are separated from the zero order by the distance $\dfrac{\lambda f}{P_{ex}}$ in x direction. The enlarged zero order of the joint power spectrum illustrated in Fig. 5.10(b) shows the inclined fringes generated by the interference between the desired target and the reference beams. After redisplaying the zero order of the joint power spectrum onto the EASLM, the second optical Fourier transform produced the correlation output. The correlation output was then captured by the CCD sensor placed at the back focal plane of the Fourier transform lens L_1. The density of the ND filter used in this recording was 3.69. Figure 5.11 shows the resultant correlation

$(-1,0)^{th}$-order $(0,0)^{th}$-order $(+1,0)^{th}$-order

$$\frac{-\lambda f}{P_{ex}} \qquad \frac{\lambda f}{P_{ex}}$$

(a)

(b)
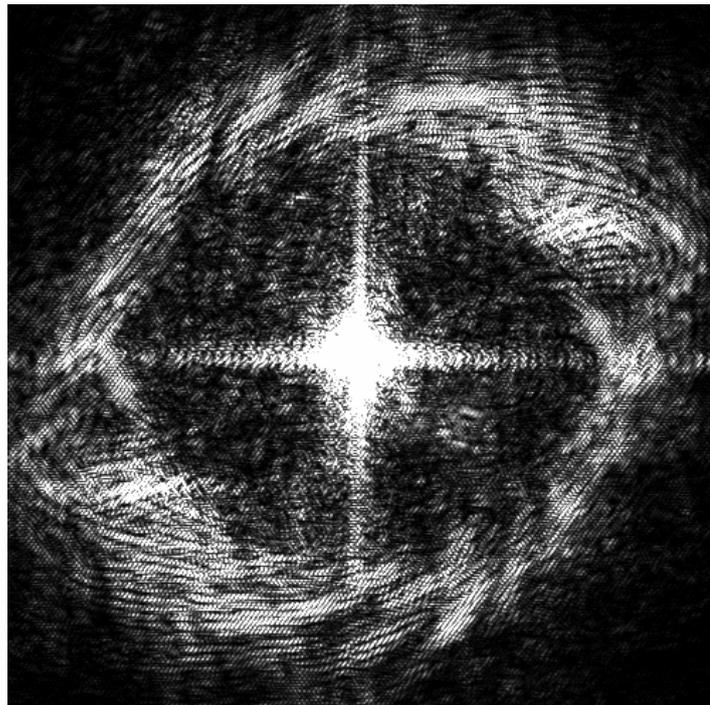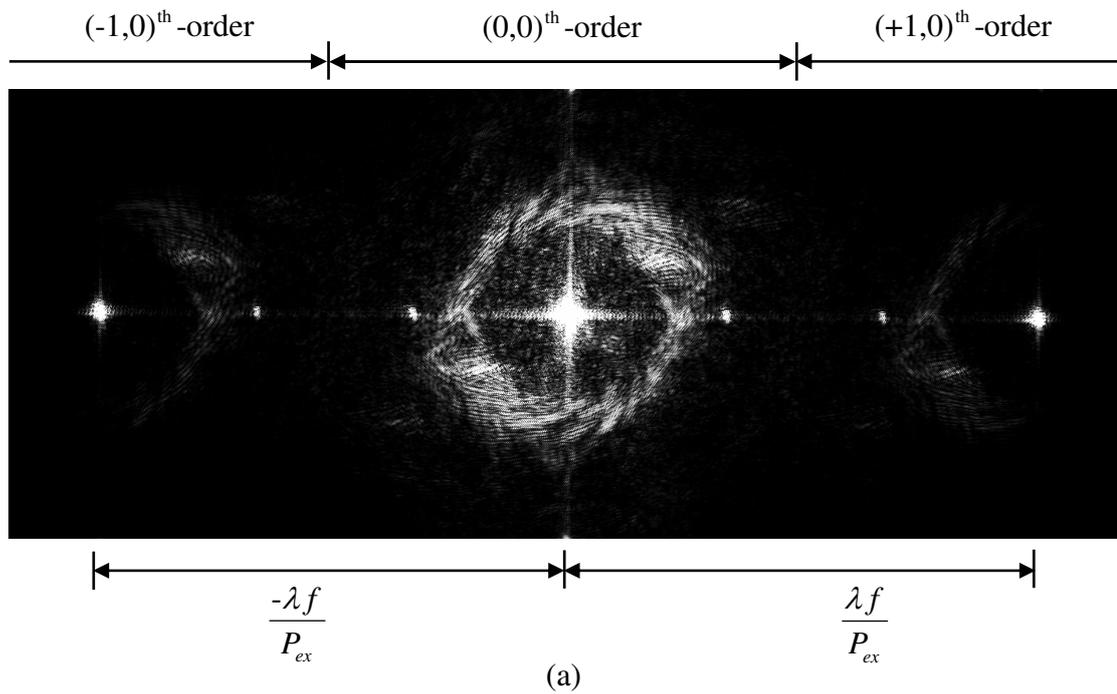
Figure 5.10   (a) joint power spectrum of the high-contrast multiple-fingerprint detection by using uncompressed reference, (b) its enlarged zero order of the joint power spectrum.

output of the multiple-fingerprint detection. The zero order of the correlation output contains the following components: three autocorrelations of the reference and the two input targets which appear at the origin, two cross-correlations between the reference and the desired input target $r(x_3, y_3) * t_1(x_3, y_3)$, two cross-correlations between the reference and non input target $r(x_3, y_3) * t_2(x_3, y_3)$, the two cross-correlations between the different targets $t_1(x_3, y_3) * t_2(x_3, y_3)$.
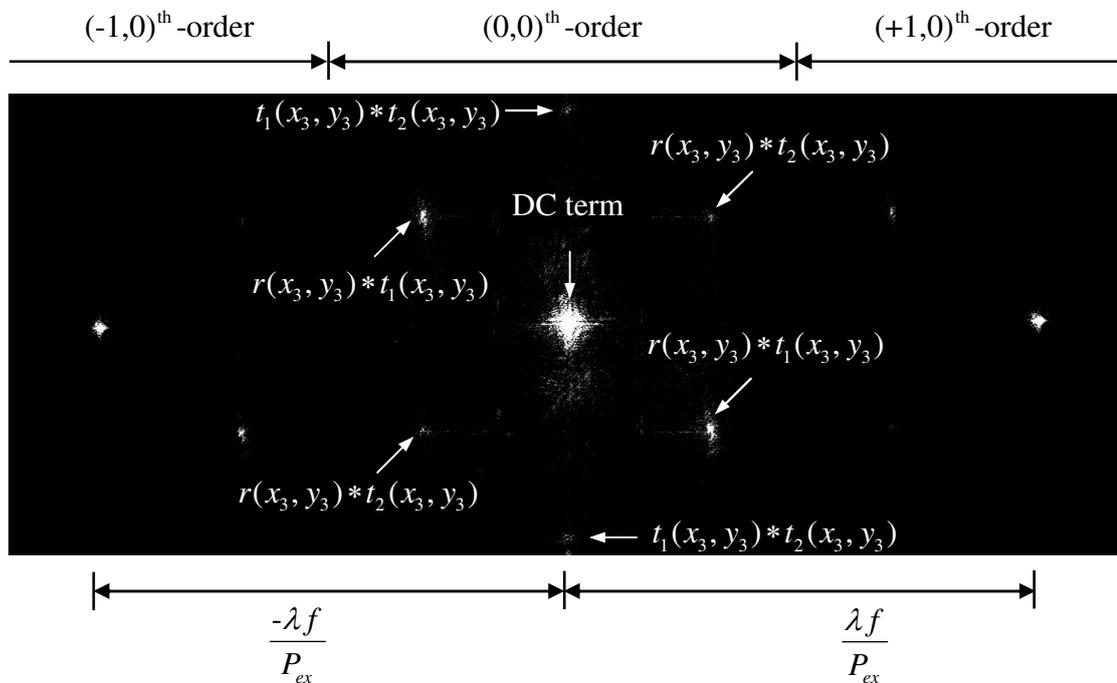
Figure 5.11 Correlation output of the high-contrast multiple-fingerprint detection.

In the case of the multiple-human face detection, the preliminary study was conducted by using uncompressed high-contrast human face as the reference. The generated joint power spectrum captured by the CCD sensor and its enlarged zero order are shown in Figs. 5.12(a) and (b), respectively. Figure 5.12(b) confirms that the joint power spectrum contains an inclined cosines fringes necessary to generate the desired correlation output. Since the displayed joint input image consisted of multiple

(a)

(b)
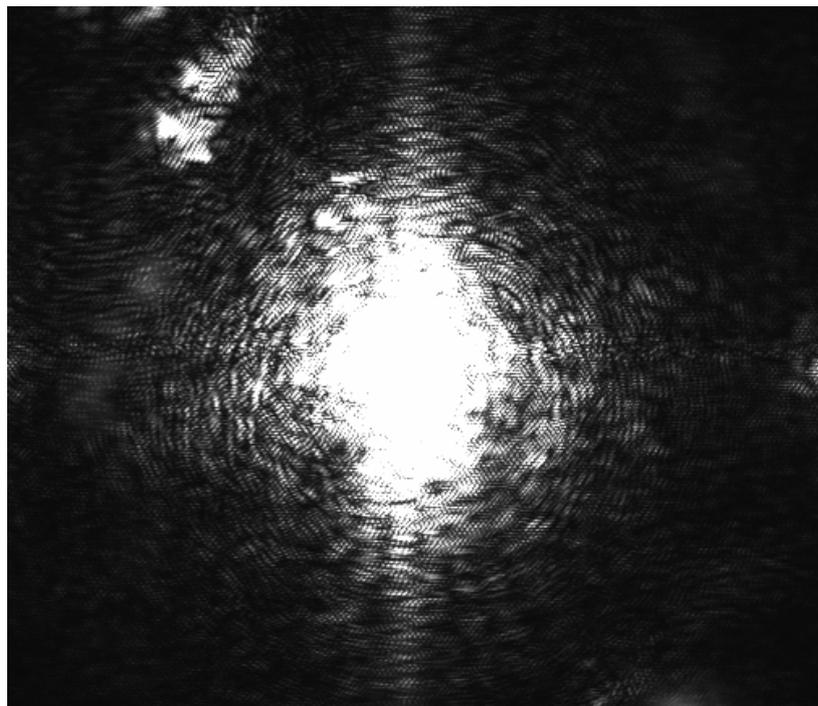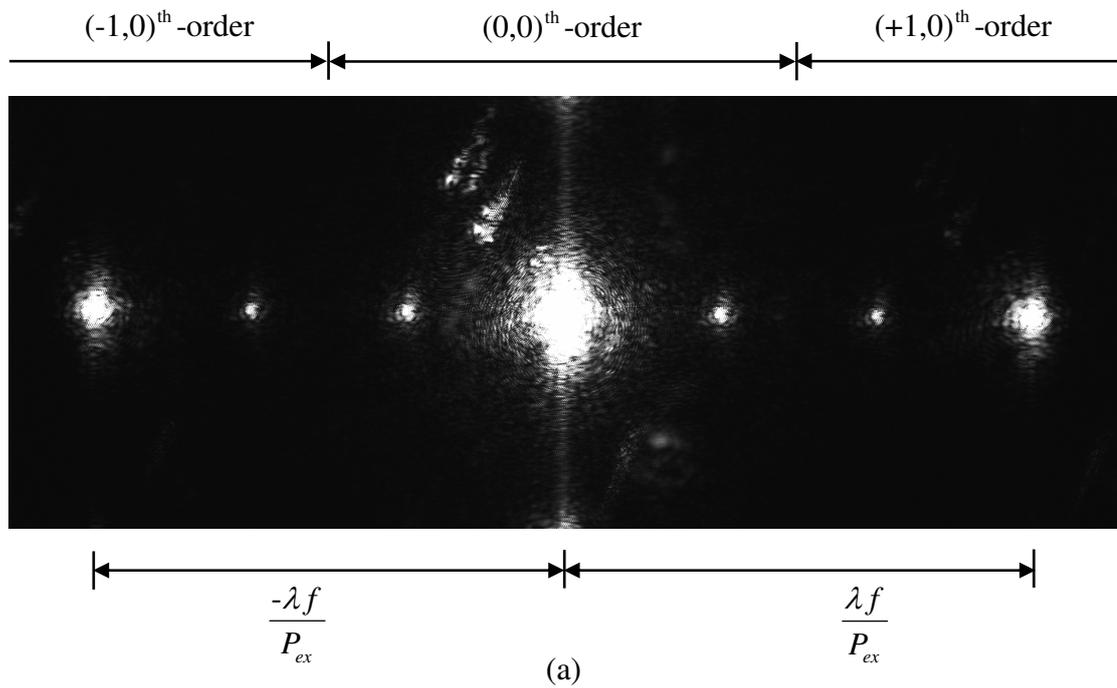
Figure 5.12   (a) joint power spectrum of the high-contrast multiple-human face detection by using uncompressed reference, (b) its enlarged zero order of the joint power spectrum.

images, large area of the EASLM transmitted light. In comparison to the single human face detection, the joint power spectrum had higher intensity of light. Thus the intensity of the joint power spectrum was reduced by using the ND filter with density of 1.0. Figure 5.13 shows the correlation output obtained by redisplaying the joint power spectrum shown in Fig. 5.12(b) onto the EASLM and taking the optical Fourier transform. This correlation output was captured by using the ND filter with the density of 3.0. Since the densities of the ND filters used in the recordings of the joint power spectrum and of the correlation output were smaller than that of the fingerprint detection, the correlation planes appear to be more noisy than that of Fig. 5.11.
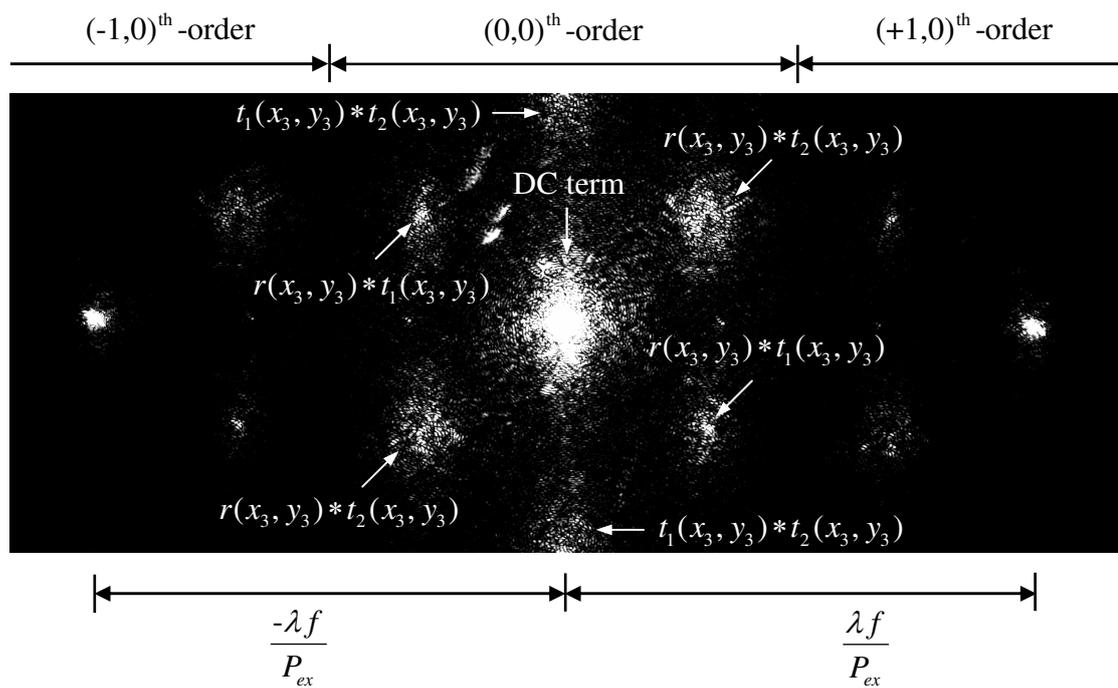
Figure 5.13   Correlation output of the high-contrast multiple-human face detection.

5.3.1   Compressed High-Contrast Fingerprint as the Reference Images

Figures 5.14(a), (b), (c), and (d) show the 3-D correlation outputs obtained from the experimental detections of the multiple-fingerprint by using the joint

(a)

(b)

(c)

(d)
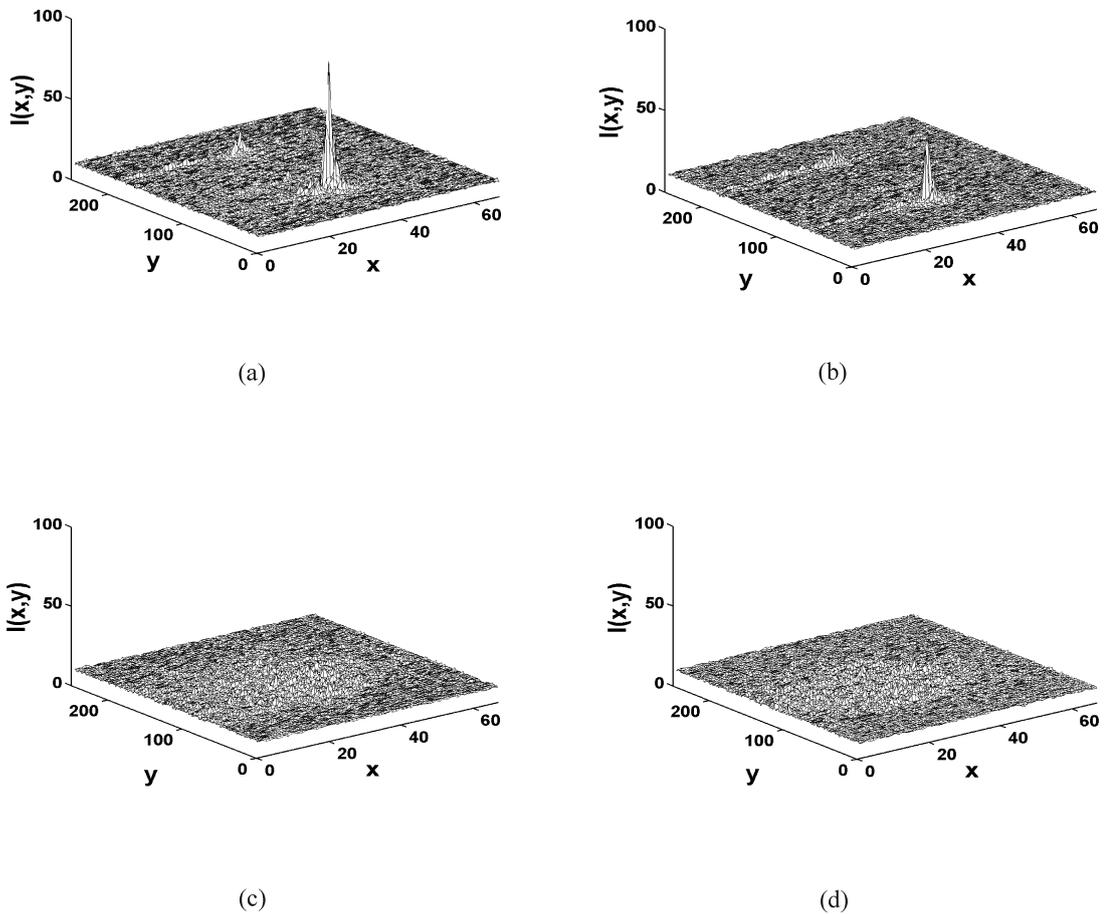
Figure 5.14 Experimental results of the multiple-target joint transform correlator. (a) autocorrelation of the uncompressed high-contrast fingerprint; and cross-correlation outputs by using the compressed high-contrast fingerprint reference (QF = 10) under a situation that the multiple target scene is: (b) noise-free high-contrast fingerprint image, (c) noisy high-contrast fingerprint image ($\sigma^2 = 1$), and noisy low-contrast fingerprint image ($\sigma^2 = 1$).

transform correlator with the compressed high-contrast fingerprint as the reference image. The autocorrelation peak of the uncompressed high-contrast fingerprint image shown in Fig. 5.14(a) is higher than the undesired secondary peak. This is in agreement with the simulation result shown in Fig. 5.2(a). When the reference is the compressed high-contrast fingerprint with QF = 10 and the multiple target is noise-free high-contrast fingerprint, both the correlation peaks shown in Fig. 5.14(b) become lower. Since the use of the uncompressed reference produces the secondary peak intensity that is smaller than the primary peak, the degradation caused by the compression lowers significantly the secondary correlation peak. Figures 5.14(c) and (d) show the correlation outputs of the joint transform correlator by using the compressed high-contrast fingerprint reference with QF = 10 for the noisy high- and low-contrast multiple-fingerprint targets with variance $\sigma^2 = 1$, respectively. It is clear that the correlation peaks are hardly observed. In particular, the secondary correlation peak is minimized. For the case of the high-contrast target, this is because the cosine fringes which encodes the desired correlation information were corrupted by the spectrum of the strong noise. However for the case of the low-contrast target, this is caused by the limitation of the EASLM in displaying low contrast image. This explanation is confirmed by Fig. 5.15(a) and (b) which show the corrupted joint power spectrum of the detections of the high- and low-contrast target, respectively.

Figures 5.16(a) and (b) show the PSRs as a function of the QF of the compressed high-contrast fingerprint reference for different multiple-target scenes obtained without and with averaging of intensities over 5×5 pixels neighborhood of the correlation peaks. The averaging process provides smooth variation of the PSRs.

<div align="center">(a)   (b)</div>
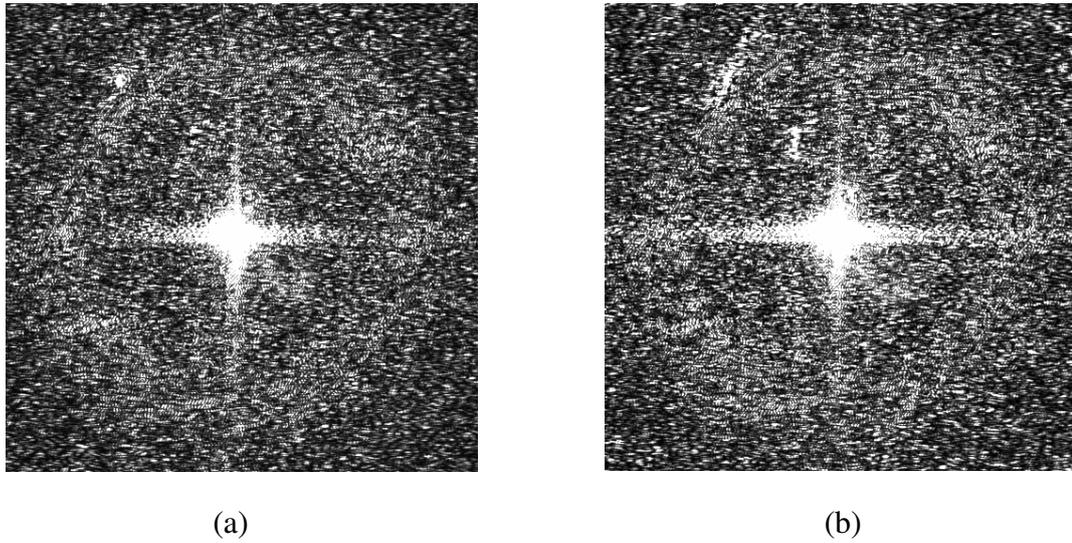
Figure 5.15   Joint power spectrum of the detection of (a) the noisy high-contrast multiple fingerprint target with variance $\sigma^2 = 1$ and (b) the noisy low-contrast multiple fingerprint target with variance $\sigma^2 = 1$ by using the compressed high-contrast fingerprint reference with QF = 10.

In the case of the high-contrast fingerprint detections, although their magnitude is different, the variation of the PSRs is similar to the simulation results in that they increase gradually as the QF becomes higher. When the target is the noise corrupted multiple-fingerprint with variance $\sigma^2 = 0.01$, its resultant PSR is higher than that of the noise-free target. This is because the secondary peak is minimized by the noise presence. Since this peak value appears in the denominator, the value of the PSR becomes higher. Furthermore, as the variance of the noise increases to unity, the joint power spectrum is strongly corrupted by the noise. Since the primary correlation peak is degraded, its PSR reduces sharply to less than 1.5. The detection of the low-contrast multiple-fingerprint targets gives different results compared to the simulation. The PSRs of the experimental results are always approximately equal to 1. This occurs
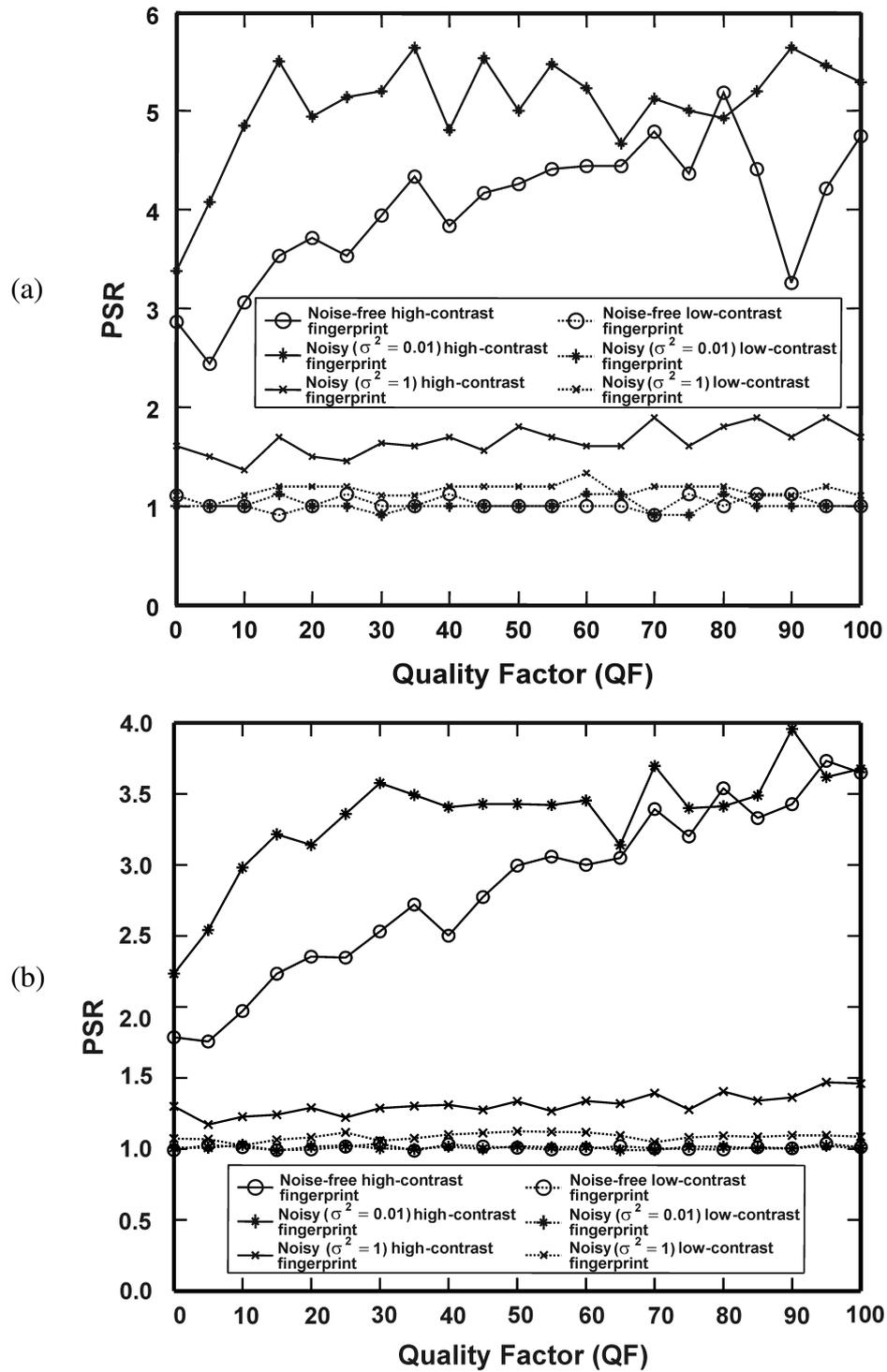
Figure 5.16 The PSR-based measurements as a function of the QF of the joint transform correlator with compressed high-contrast fingerprint reference: (a) without and (b) with intensities averaging over 5×5 pixels neighborhood of the original correlation peak.

because the EASLM cannot display effectively low-contrast images. Since the cosine

fringes were not generated, there were no primary and secondary correlation peaks.

As a result, the PSRs become approximately 1.

5.3.2 Compressed Low-Contrast Fingerprint as the Reference Images

Due to the limitation of the EASLM discussed in Chapter IV, the experimental

verifications of the multiple-fingerprint detection of using compressed low-contrast

fingerprint reference cannot be accomplished. The joint power spectrum of Fig.

5.17(b) which was experimentally generated does not contain the spectral information

as the one generated by the computer simulation shown in Fig. 5.17(a).
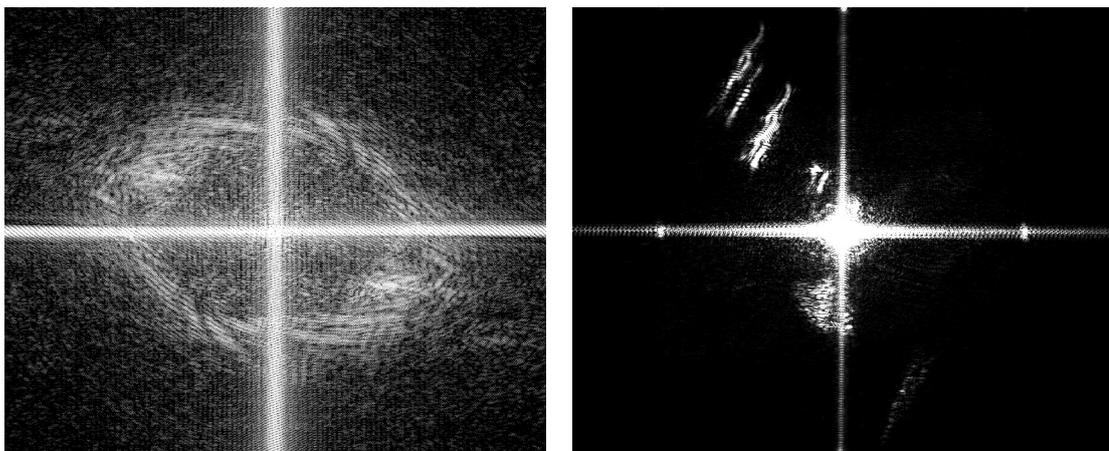
(a)   (b)

Figure 5.17   The joint power spectrum of the low-contrast multiple-fingerprint

obtained from (a) the simulation and (b) the experiment.

5.3.3 Compressed High-Contrast Human Face as the Reference Images

Figure 5.18(a) shows the 3-D plot of the correlation output of the multiple-

high-contrast human face detection by using the joint transform correlator with
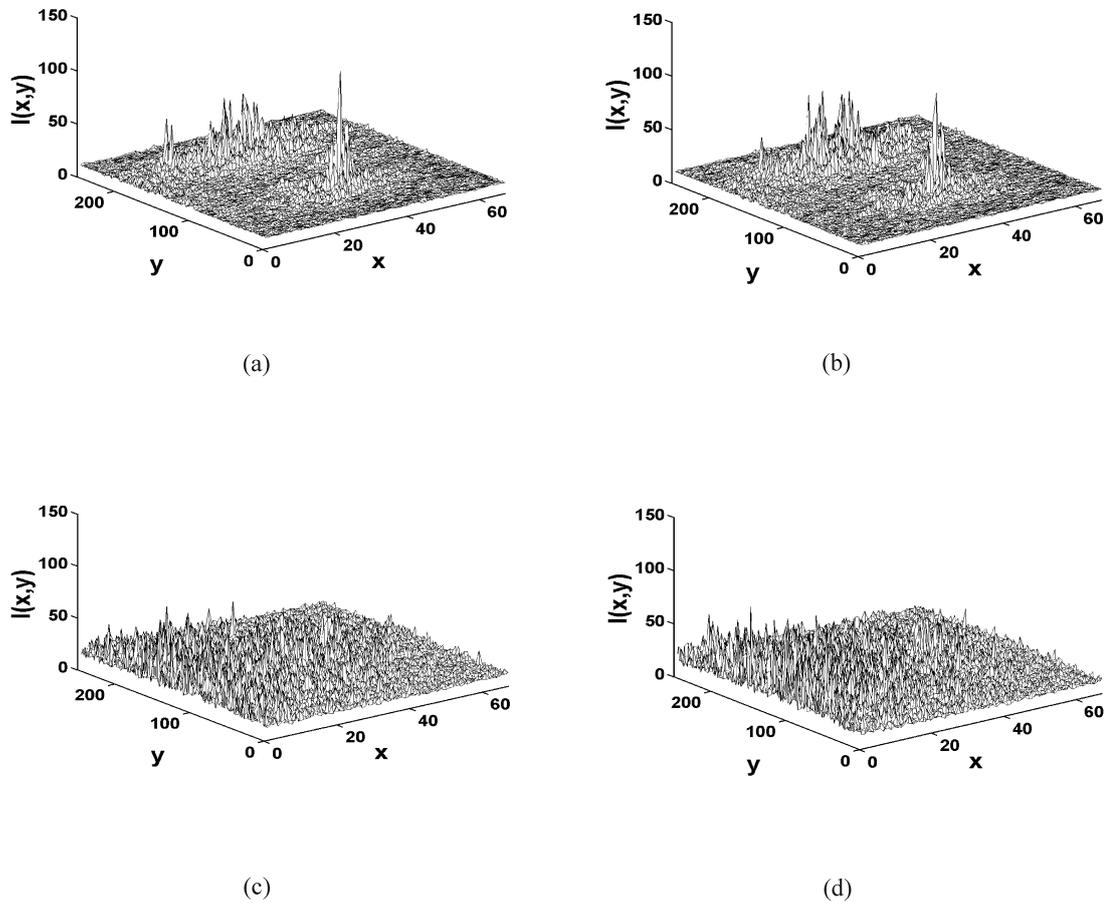
Figure 5.18 Experimental results of the multiple-target joint transform correlator. (a) autocorrelation of the uncompressed high-contrast human face; and cross-correlation outputs by using the compressed high-contrast human face reference (QF = 10) under a situation that the multiple target scene is: (b) noise-free high-contrast human face image, (c) noisy high-contrast human face image ($\sigma^2 = 1$), and noisy low-contrast human face image ($\sigma^2 = 1$).

uncompressed high-contrast human face reference, while Figs. 5.18(b), (c), and (d) illustrated the correlation outputs of the detection with the reference compressed at QF = 10. It can be seen from Fig. 5.18(a) that the autocorrelation peak intensity is higher and sharper than that of the secondary correlation peak generated by the non target. As the reference is compressed, the correlation peak intensities shown in Fig. 5.18(b) reduce slightly. This is in agreement with the computer simulation results shown in Fig. 5.6. However, when the human face targets are strongly corrupted by noise, Figs. 5.18(c) and (d) show that there are no observable correlation peaks. As discussed in Sect. 5.3.1, the presence of strong noise in the input and the limitation of the EASLM degrade significantly the correlation outputs.

Figures 5.19(a) and (b) show the variation of the PSRs as a function of the QF of the compressed high-contrast human face reference for different multiple-target scenes obtained without and with averaging of intensities over 5×5 pixels neighborhood of the correlation peak. By averaging correlation intensities over 5×5 pixels neighborhood of the correlation peak, the values of the resultant PSRs of the noise-free and the noisy high-contrast human face with variance $\sigma^2 = 0.01$ vary around 3. These are almost the same as the simulation results.

5.3.4   Compressed Low-Contrast Human Face as the Reference Images

The experimental verifications of the multiple-human face by using low-contrast human face reference cannot be successfully accomplished because the EASLM has a low contrast ratio. As evident by Figs. 5.20 (a) and (b) which shows the joint power spectrum of the multiple-human face detection of using low-contrast human face reference, the optically generated joint power spectrum is far from the ideal joint power spectrum generated by digitally computation.
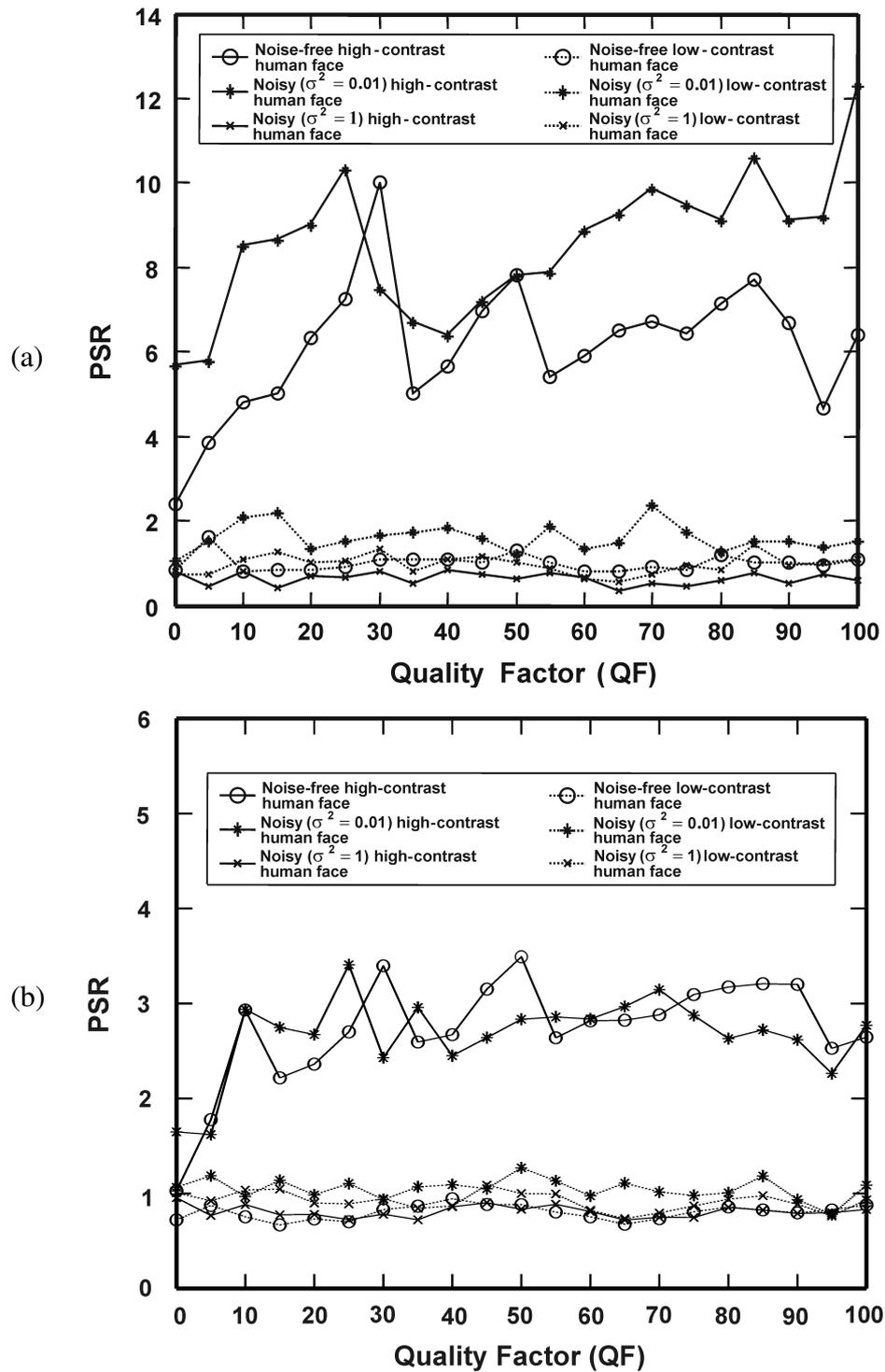
Figure 5.19 The PSR-based measurements as a function of the QF of the joint transform correlator with compressed high-contrast human face reference: (a) without and (b) with averaging of intensities over 5×5 pixels neighborhood of the original correlation peak.

(a) (b)
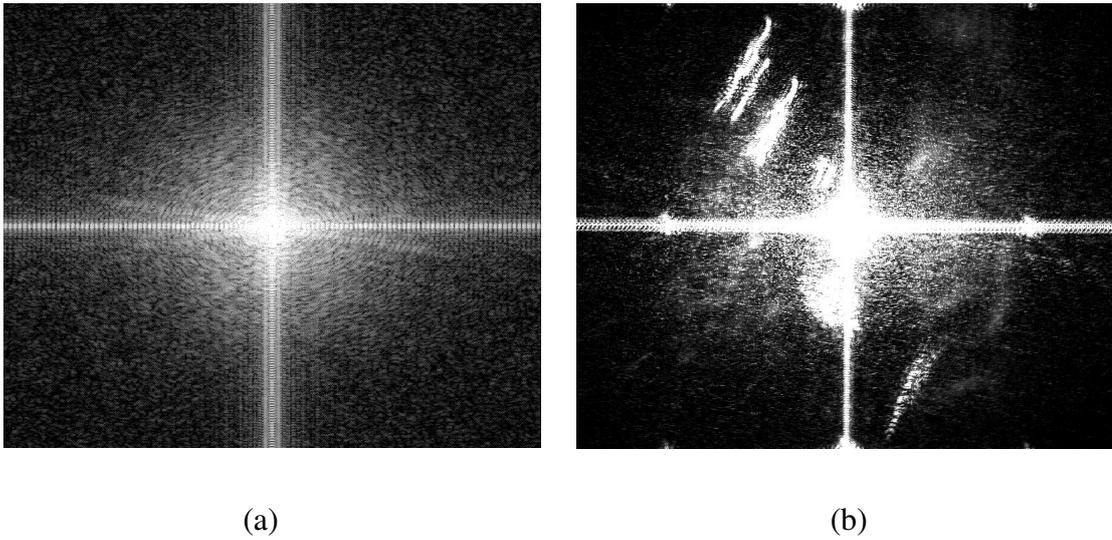
Figure 5.20 The joint power spectrum of the multiple-human face detection by using low-contrast human face reference obtained from (a) the simulation and (b) the experiment.

CHAPTER VI

CONCLUSIONS

In this dissertation, in order to solve storage problems and improve processing time of automatic target recognition systems, a real-time implementation of joint transform correlator by using JPEG-compressed reference images has been theoretically studied and experimentally verified. The storage problem arises from the fact that the joint transform correlator employs a large number of reference images to deal with all possible variations of targets, while the process of the target and the reference images introduces a time delay which depends on the size of images.

There were two primary objectives in this dissertation. The first objective was to study the effects of reference image compression on the recognition performance of the joint transform correlator. The second one was to obtain a guide line to optimize the recognition performance of the joint transform correlator on using compressed reference images. In order to achieve these objectives, studies of single- and multiple-target recognitions were performed by using fingerprint and human face images as test scenes with different spatial-frequency contents, and by taking into account the presence of noise in the input targets and of contrast difference between the target and the reference images. The JPEG-compressed reference images with different compression levels were generated by using ACDsee software. The computer simulations of the joint transform correlator with compressed reference images were performed by using MATLAB. The experimental verifications of the joint transform

correlator with compressed reference images were done by using one stage optical Fourier transformer where the twisted nematic liquid crystal display, the He-Ne laser, and the CCD sensor were used as the EASLM, the coherent light source, and the light detector, respectively. However, due to the limitation of the EASLM on displaying low contrast image and low dynamic range of the CCD sensor, not all experimental verifications could be performed.

In Chapter III, the effect of JPEG compression on the information contents of the compressed reference images was assessed by using the CR and the PSNR. From the four types of reference images, the compressed low-contrast human face and the high-contrast fingerprint have the highest and the lowest CRs, respectively. The CR of the high-contrast human face is slightly higher than the low-contrast fingerprint for the low QF. The objective evaluation by using the PSNR shows that PSNRs of the low-contrast image is higher than that of the high-contrast image. The subjective evaluation of the compressed image shows that at low QF, the quality of the image is reduced with distinctive blocking artifacts.

In Chapter IV, the performance degradation of the proposed joint transform correlator used for single-target detection is summarized in Table 6.1. The simulation results show that the effects of compression of the high-contrast human face reference on the correlation performance of the joint transform correlator is insignificant for all given target scenes. Although the detection performance of the joint transform correlator by using the compressed low-contrast human face decreases at the low QF, the degradation due to the noise presence and the contrast difference is small. Besides being sensitive to noise, the detection performance of the joint transform correlator by using the compressed fingerprint reference depends on the compression, and the

Table 6.1

Performance degradation of the single-target detection by using the joint transform correlator with compressed reference images

Simulation

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

Experiment

Compressed reference image	Condition of multiple-target images			
	Noise free		Noisy	
	High contrast	Low contrast	High contrast	Low contrast
High-contrast fingerprint	Gradual	Severe	Severe	Severe
Low-contrast fingerprint	Not available	Not available	Not available	Not available
High-contrast human face	Gradual	Severe	Severe	Severe
Low-contrast human face	Not available	Not available	Not available	Not available

degradation for the low-contrast fingerprint is more significant compared with the high-contrast one. In the experiment, only detections of target by using the compressed high-contrast references could be verified. This is because the used EASLM could not display efficiently low-contrast images and the dynamic range of the CCD sensor is low. As a result, the remaining experimental verifications by using low-contrast references could not be successfully performed. The experimental results show that the recognition performance of the high-contrast fingerprint target by using compressed high-contrast fingerprint is in agreement with the simulation results, while the recognition of the high-contrast human face target by using compressed high-contrast human face reference is different with the simulation, because the generated joint power spectrum is corrupted by the speckle noise. Furthermore, the performance degradation of the low-contrast fingerprint and human face target detections are severe because of the limitation of the EASLM.

Since the experimental verifications of the real-time joint transform correlator with compressed reference images are not complete, the guideline for optimizing the recognition performance of the system is based on the simulation results. In the case of single-target detection, the guideline is determined by considering that the normalized PCD obtained from the simulations must not decrease more than 10 percent below the maximum value that is unity. Expect for detection of the noisy low-contrast fingerprint target with variance $\sigma^2 = 1$, the high-contrast fingerprint reference can be compressed until QF = 30 that is equivalent to the compression of about 5.2 times smaller than the original size. Since the effect of compression on the low-contrast fingerprint image is greater than that of the high-contrast image, the maximum compression of the low-contrast fingerprint reference is QF = 85 that

corresponds to the file size ratio of about 5.3 times smaller than the original size. Therefore, regardless of the contrast, the maximum compression of the fingerprint reference is about 5.2 times. Based on the same guideline for the detection of the fingerprint, Fig. 4.8 shows that the joint transform correlator by using compressed high-contrast human face can be accomplished at any compression levels. In the case of the low-contrast human face reference, the smallest QF is 5 that corresponds to the CR = 36.5. Thus, the maximum compression of the human face reference is about 33.6 times.

In Chapter V, the multiple-target detection by using joint transform correlator with compressed reference images was investigated. Table 6.2 summarizes the performance degradation obtained from the computer simulation and the experiment. The simulation results showed that 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 joint transform correlator is not significant for all given target scenes regardless of the noise and the contrast difference. This is in agreement with study of the single target detection. 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. It is worth mentioning that the use of the low-contrast reference with high-spatial-frequency contents may yield false alarms. The experimental verifications of the detection of multiple-target give the same results as the single target. In the case of the multiple-target detection, the guideline for optimizing the detection performance is that the PSR must be greater than unity. By taking this into account, except for the

Table 6.2

Performance degradation of the multiple-target detection by using the joint transform correlator with compressed reference images

Simulation

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

Experiment

Compressed reference image	Condition of multiple-target images			
	Noise free		Noisy	
	High contrast	Low contrast	High contrast	Low contrast
High-contrast fingerprint	Severe	Severe	Severe	Severe
Low-contrast fingerprint	Not available	Not available	Not available	Not available
High-contrast human face	Gradual	Severe	Severe	Severe
Low-contrast human face	Not available	Not available	Not available	Not available

detection of the noisy multiple low-contrast fingerprint target with variance $\sigma^2 = 1$, the compression of the high-and the low-contrast fingerprint references can be done at any levels. The same compression requirement can also be applied to the joint transform correlator with compressed high-and low-contrast human face.

REFERENCES

REFERENCES

Alam, M. S. and Karim, M. A. (1993). Joint-transform correlation under varying illumination. **Applied Optics**. 32(23): 4351-4356.

Bhanu, B. and Jones, T. L. (1993). Image understanding research for automatic target recognition. **IEEE Aerospace and Electronic System Magazine**. October: 15-22.

Feitelson, D. G. (1988). **Optical Computing: A survey for Computer Scientists**. Massachusetts: The MIT press.

Goodman, J. W. (1996). **Introduction to Fourier Optics** (2nd ed.). New York: McGraw-Hill.

Heckbert, P. (1998). **Fourier transforms and the fast Fourier transform (FFT) algorithm** [On-line]. Available: http://www.cs.cmu.edu/afs/andrew/scs/cs/ 15-463/pub/www/notes/fourier/fourier.pdf

Hess, R. F., Bradley, A., and Piotrowski, L. (1983). Contrast coding in amblyopia. I. Differences in the neural basis of human amblyopia. **Proceedings of the Royal Society of London Series B**. February, 217(1208): 309-330.

Jutamulia, S., Storti, G. M., Gregory, D. A., and Kirsch, J. C. (1991). Illumination-independent high-efficiency joint transform correlator. **Applied Optics**. 30(29): 4173-4175.

Li, C. T., Yin, S., and Yu, F. T. S. (1998). Nonzero-order joint transform correlator. **Optical Engineering**. 37(1): 58-65.

Li, J., Zhang, Y., and Hu, J. (1996). Object recognition with a wavelet-transform-based joint transform correlator. **Optical Engineering**. 35(3): 775-777.

Lugt, A. V. (1964). Signal detection by complex spatial filtering. **IEEE Transactions on Information Theory**. IT-10: 139-145.

Oppenhein, A. V. and Schafer, R. W. (1989). **Digital signal processing**. Englewood Cliffs, N.J.: Prentice-Hall

Pati, G. S. and Singh, K. (1998). Illumination sensitivity of joint transform correlators using differential processing: computer simulation and experimental studies **Optics Communications**. 147: 26-32.

Pennebaker, W. B. and Mitchell, J. L. (1993). **JPEG Still Image Data Compression Standard**. New York:Van Nostrand Reinhold.

Ratches, J. A., Walters, C.P., Buser, R. G., and Guenther, B. D. (1997). Aided and automatic target recognition based upon sensory inputs from image forming systems. **IEEE Transactions on Pattern Analysis and Machine Intelligence**. 19(9): 1004-1019.

Roberge, D. and Sheng, Y. (1994). Optical wavelet matched filter. **Applied Optics**. 33(23): 5287-5293.

Salomon, D. (1998). **Data Compression: The Complete Reference**. New York: Springer-Verlag.

Tanone, A., Uang, C-M., Yu, F. T. S., Tam, E. C., and Gregory, D. A. (1992). Effect of thresholding in joint-transform correlation. **Applied Optics**. 31(23): 4816-4822.

Wang, R. K., Shang, L., and Chatwin, C. R. (1996). Modified fringe-adjusted joint transform correlation to accommodate noise in the input scene. **Applied**

Optics. 35(2): 286-296.

Widjaja, J. and Suripon, U. (2004). Real-time joint transform correlator using compressed reference images. **Optical Engineering**. 48(8): 1737-1745.

Widjaja, J. (2003). Effects of image compression on digital specklegrams. **Optics and Lasers in Engineering**. 39(4):501-506.

Yu, F. T. S., Lin, S. T. W., and Gregory, D. A. (1987). Adaptive real-time pattern recognition using a liquid crystal TV based joint transform correlator. **Applied Optics**. 26: 1370-1372.

Yu, F. T. S. and Lu, X. J. (1984). A real-time programmable joint transform correlator. **Optics Communications**. 52(1): 10-16.

Zhang, S. and Karim, M. A. (1999). Illumination-invariant pattern recognition with joint-transform-correlator-based morphological correlation. **Applied Optics**. 38(35): 7228-7237.

CURRICULUM VITAE

Miss Ubon Suripon was born on May 22, 1974 in Buriram. She received her B.Eng. degree in Telecommunication Engineering and her M.Sc. degree in Laser Technology from Suranaree University of Technology in 1997 and 2000, respectively. She pursued her Ph.D. degree in Laser Technology from the same university where she worked as a teaching assistant in the School of Mathematics and the School of Physics. Her research of interests includes optical and digital image processing.