8

Lightwave Associative Memory That Memorizes and Recalls Information Depending on Optical-Carrier Frequency

This chapter presents a lightwave neural network that learns behavior depending on optical frequency. In general, a neural network learns or self-organizes adaptively to environment. Because of this adaptability, we sometimes have difficulty in controlling the neural network at will. The lightwave network to be presented here learns a certain processing at a certain optical-carrier frequency, and another processing at another frequency, to possess both the adaptability and the controllability. From a different viewpoint, we can consider this compatibility as a mean of multiplexing of behavior in the optical frequency domain. The frequency-domain multiplexing provides the neural network, which is characterized by distributedness and parallelism, with a new dimension of massive parallelism utilizing the vast optical frequency bandwidth. In this sense, the network presents a novel direction in optical information processing hardware.

8.1 Utilization of Wide Frequency Bandwidth in Optical Neural Networks

The present state-of-the-art optical fiber communications involves frequency-domain multiplexing (FDM) in the optical frequency domain. The communications frequency (so-called 1.55μm-wavelength band) is first divided into L band (186.25‘190.95THz) and C band (191.60‘196.55THz). Then each band is further divided into channels of 100 or 200GHz bandwidth just like the channels in television and radio broadcasting systems. Each channel carries information independent of the information conveyed by other channels.

If we utilize the same technique in optical information processing, we can realize a massive parallelism in the vast optical frequency domain. Simultaneously, the fact that early-days' telephone-switching systems were equivalent to electronic computers suggests us that the realization of an optical FDM information processor should also be equivalent to the construction of an adaptive and intelligent all-optical FDM packet switch.

A. Hirose: Complex-Valued Neural Networks, SCI 400, pp. 133–142.
springerlink.com

Parallelism and distributedness are the most specific characteristics in the construction and dynamics of neural networks. Optical circuits have plasticity and three dimensionality in connections. At these points, we can say that optical circuits possess high spatial parallelism. In particular in short-wavelength circuits, the spatial parallelism can be enhanced. Therefore, conventional optical neural networks emphasized the advantage in the spatial parallelism.

On the other hand, electronic neural hardware has also been investigated for a long time. Owing to high locality of electrons, electronic circuits have been greatly minimized and integrated, and are nowadays integrated even three dimensionally. Consequently, the spatial density of electronic hardware is generally much higher than that of optical hardware.

Ordinary optical circuits don't work well in space smaller than its wavelength. Moreover, optical circuits have the so-called telescope-effect problem. That is, lightwave generally requires such a long propagation length that an optical circuit cannot be densely integrated in this direction. To solve such problems, we need to go further into optical nanotechnology, for instance, techniques to practically utilize evanescent light.

However, the operation speed of electron devices is inevitably restricted by the electron charge. Contrarily, lightwave has an extremely higher operation speed (= wider bandwidth). The ultrawide bandwidth is a big advantage in optical circuits and should be explored much more. Additionally, conventional neural optical circuits have utilized only the power of lightwave ($\approx$ intensity and amplitude) among other parameters such as phase, frequency, and polarization. If we pay attention to frequency, we realize the vast optical frequency bandwidth mentioned above, which is a crucial point to be investigated in neural networks. To manipulate frequency precisely with high resolution, we have to deal with phase of lightwave. For this purpose, the complex-valued neural networks are essential and indispensable.

This chapter describes a coherent optical neural network that has carrier-frequency-dependent behavior by utilizing the wide optical bandwidth. Carrier frequency is the frequency of the pure sinusoidal wave without information modulation. When we use a semiconductor laser, we can easily modulate the carrier frequency by changing the injection current fed to the laser diode. This type of frequency modulation (FM) is called the direct frequency modulation. In this chapter, we construct a coherent lightwave associative memory that memorizes and recalls patterns dependent on the value of the optical-carrier frequency. Neural connection weights are determined by use of the complex-valued Hebbian rule expressed in a frequency-dependent form. Optical experiments demonstrate carrier-frequency-dependent memory and recall [204],[192].

The system to be presented here follows the preceding construction of a coherent lightwave associative memory that memorizes complex-valued vectors reported in Ref.[205]. Coherent neural networks inherently possess

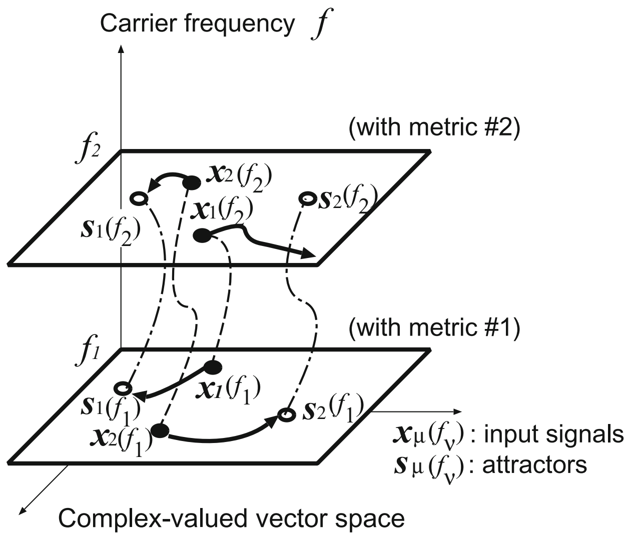

Fig. 8.1 Conceptual illustration of the carrier-frequency-dependent behavior in the associative memory system. (Reprinted from Fig.1 in [204]: Sotaro Kawata and Akira Hirose: Coherent lightwave associative memory system that posesesses a carrier-frequency-controlled behavior, *Opt. Eng.,* 42(9):2670–2675, 2003, with permission, SPIE.)

carrier-frequency-dependent behavior as mentioned in Section 4.3.8. The present system is based on this nature.

The first optical experiment of complex-valued associative memory was reported by Takeda & Kishigami in 1992 [82]. They constructed the epoch-making lightwave associative memory based on the mathematical analogy between the electromagnetic field in optical resonators formed with phase-conjugate mirrors and the dynamics of hermitian associative memories. The underlying idea is applicable to various situations where we deal with physical waves. However, the possible neural functions are limited within associative memories. On the other hand, the system to be presented here is not based on phase-conjugate physics, but using optical modulators having more flexible operation as an optical device. Therefore, in this sense, the bases described below have a higher potential to construct neural networks with a variety of functions.

Besides, we have several reports on the utilization of the optical-frequency domain so far, for instance, an FDM memory using volume hologram [206], a winner-take-all network based on the nonlinearity in semiconductor laser oscillation with external optical feedback and the spectral hole burning [207], and numerical experiments showing various generalization characteristics in the frequency domain realized by learning and self-organization [83].

8.2 Optical-Carrier-Frequency Dependent Associative Memory: The Dynamics

Figure 8.1 is an illustration presenting the concept of the optical-carrier-frequency-dependent behavior in the associative memory. The carrier frequency means the frequency of the lightwave that conveys information, and is identical to the oscillation frequency of the semiconductor laser used.

If the optical carrier frequency is f_1, the associative memory system has an information metric corresponding to f_1. When the memory is fed with an input vector $\boldsymbol{x}_1(f_1)$, it recalls a memorized vector $\boldsymbol{s}_1(f_1)$ that is nearest to the input in the metric system determined by f_1. For a different input vector $\boldsymbol{x}_2(f_1)$, it recalls another one, $\boldsymbol{s}_2(f_1)$, nearest to $\boldsymbol{x}_2(f_1)$. However, on condition that the system has a carrier frequency of f_2, the metric is changed, giving the memory a different worldview. That is, for instance, it recalls $\boldsymbol{s}_1(f_2)$ for an input $\boldsymbol{x}_2(f_2)$ since $\boldsymbol{x}_2(f_2)$ is near to $\boldsymbol{s}_1(f_2)$ in this metric. An input $\boldsymbol{x}_1(f_2)$ may be too far from any memorized vectors, and yields none of them.

The above recalling story is just a possible example. The recall behavior is determined by learning that is dependent on the carrier frequency. The detail is given below.

8.2.1 Recalling Process

Given the amplitude is fixed at unity without amplitude modulation, the neurodynamics to recall a memorized vector is expressed as follows. As mentioned in Chapter 4, the amplitude can be variable in general. However, in the present system, it is constant for simplicity.

$$\boldsymbol{x}(d+1) = A\left(|\mathbf{W}\boldsymbol{x}(d)|\right)\exp(i\arg\{\mathbf{W}\boldsymbol{x}(d)\}) \tag{8.1}$$

$$x_i = |x_i|\exp(i\alpha_i) \tag{8.2}$$

$$w_{ji} = |w_{ji}|\exp(i2\pi f\tau_{ji}) \tag{8.3}$$

where $A(\cdot)$ and $\mathbf{W} = [w_{ji}]$ are amplitude nonlinear function and connection weights, respectively, and d is discrete time, f is carrier frequency, α_i is input signal phase, and τ_{ji} is delay time of the connections.

We realize phase modulation by using a phase-modulation-type (more precisely, delay-time modulation type) spatial light modulator (SLM), which is named parallel-aligned liquid-crystal spatial light modulator (PAL-SLM). In the experiment below, we use two SLMs, i.e., one (SLM#1) is to generate an input vector, and the other (SLM#2) works as connection weights. The details will be given later in Fig.8.2. In the present system, we modulate only the phase (or actually, delay time), and keep the amplitude unchanged.

8.2.2 Memorizing Process

The learning dynamics to realize an associative-memory network is explained as follows. We apply sequential correlation learning to the weights

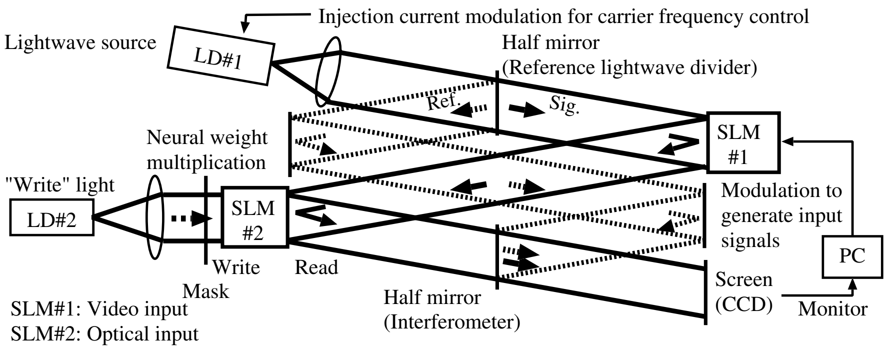

Fig. 8.2 Construction of the lightwave associative memory whose behavior is controllable by carrier-frequency modulation. (Reprinted from Fig.2 in [204] in figure caption of Fig.8.1 with permission.)

$w_{ji} = |w_{ji}| \exp(i2\pi f_\nu \tau_{ji})$. That is to say, we present the vector to be memorized ($\boldsymbol{s}_{\mu,\nu}$) to the output of the network instead of $\boldsymbol{y}$ in the ordinary Hebbian learning, while the conjugate $(\boldsymbol{s}_{\mu,\nu})^*$ to the input as $\boldsymbol{x}$, to make the network learn their correlation one by one. The process is, so to speak, a supervised complex-valued Hebbian learning. The updates of the amplitude and delay, $|w_{ji}|$ and τ_{ji}, of the weight w_{ji} is conducted in the manner described in Section 4.2 as

$$\tau \frac{d|w_{ji}|}{dt} = -|w_{ji}| + |y_j||x_i| \cos(\beta_j - \alpha_i - 2\pi f \tau_{ji}) \tag{8.4}$$

$$\tau \frac{d\tau_{ji}}{dt} = \frac{1}{2\pi f} \frac{|y_j||x_i|}{|w_{ji}|} \sin(\beta_j - \alpha_i - 2\pi f \tau_{ji}) \tag{8.5}$$

where $x_i = |x_i| \exp(i\alpha_i)$ and $y_j = |y_j| \exp(i\beta_j)$ are input and output signals, respectively, and τ is learning time constant [208].

In the case that all the memorized vectors are given at once, we can construct the weight matrix directly as the autocorrelation matrix.

8.3 Optical Setup

Figure 8.2 shows the system construction. We can control the carrier frequency f by choosing appropriate injection current of the light source (semiconductor laser: LD#1). Emitted lightwave is divided into signal and reference beams. The former is modulated by SLM#1, and becomes an input signal vector $\boldsymbol{x}$. Then the signal beams $\boldsymbol{x}$ are incident on SLM#2 and multiplied by the weights $\mathbf{W}$ whose values are determined by the backlight (LD#2) with an optical mask placed behind SLM#2. Figure 8.3 illustrates the SLM-surface assignment when the neuron number is 3 and the synapse number is 9, as well as the CCD surface to yield the summation.

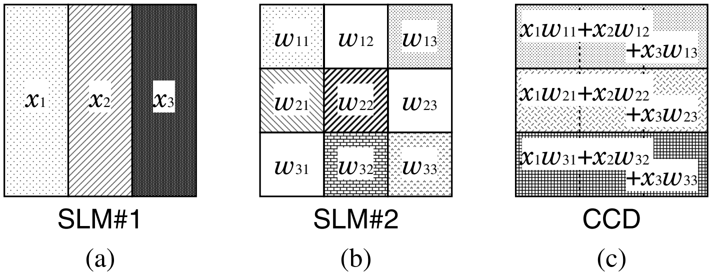

Fig. 8.3 Modulation-surface assignment of (a)input-signal generating SLM and (b)connection-weight multiplexing SLM, and (c)CCD detection-surface assignment in a 3-neuron and 9-synapse case for example. (Reprinted from Fig.3 in [204] in figure caption of Fig.8.1 with permission.)

The yielded neural output signals are mixed with the reference light beam at the half mirror to be homodyne-detected. To obtain two components orthogonal to each other, we modulate the input signal phase additionally by 0, 90, 180, and 270 degrees so that they yield four different interferences, resulting in extraction of the amplitude and phase information. The captured image is fed to the personal computer (PC) to generate the sum. (If we use a striped-surfaced detector, we can optically obtain the sum.) The PC generates the next signal vector to be sent to SLM#1 recurrently. The setup photograph is shown in Fig.8.4.

Figure 8.5 shows schematically the frequency-dependent behavior expressed as the modulation signals on the SLM surfaces. Even though the delay-time values τ_{ji} on SLM#2 are fixed, the resulting phase values of the weights $2\pi f \tau_{ji}$ are variable depending on the carrier frequency f. Hence, for an identical input vector, the system yields different output vectors dependent on f. In other words, the association behavior is dependent on the carrier frequency. This characteristic is effectively used, for both the learning and recall are consistently dependent on the frequency.

8.4 Frequency-Dependent Learning

The dependence of the homodyne output signals on the carrier frequency f is determined by the optical-path difference between the signal and reference light beams shown in Fig.8.2. When f varies, the interference fringe changes periodically with an interval of $c/\Delta L$.

The initial delay time $\tau_{ji0}^{\mathrm{Hebb}}$ is related to an arbitrary basis frequency f_0 as

$$\tau_{ji0}^{\mathrm{Hebb}} = \frac{\theta_0^{\mathrm{SLM}}}{2\pi f_0} + \frac{\Delta L}{c} \tag{8.6}$$

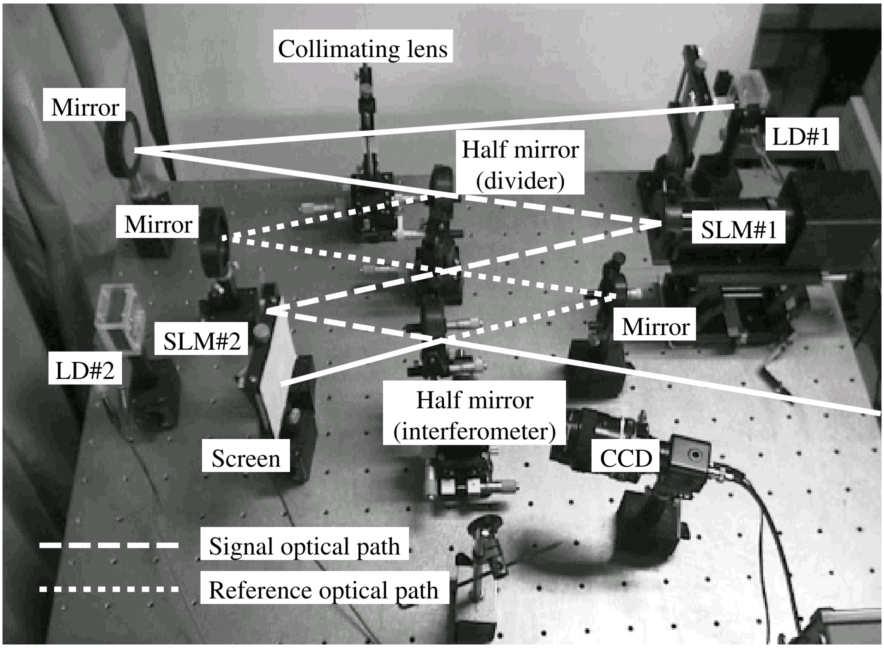

Fig. 8.4 Experimental setup.

where θ_0^{SLM} is an equivalent initial phase value of the SLM modulation, and is chosen at random in $[0, 2\pi]$. Generally, a large delay yields a large phase change even against a small frequency deviation. In the present system, the network learns an appropriate delay time $(\theta_0^{\mathrm{SLM}}/2\pi f_0) + (\Delta L/c)$ to have a suitable sensitivity to the frequency change.

In the experiment shown below, first we conduct learning numerically in the PC on the basis of the above-mentioned complex-valued Hebbian rule. Then, after learning is finished, we move to optical recall experiment. We named the μ th vector to be memorized at ν th frequency as $\boldsymbol{s}_{\mu,\nu}$. We show the vectors $\boldsymbol{s}_{\mu,\nu}$ by adjusting the carrier frequency at f_ν, one by one, to the neural network, and embed the vectors $\boldsymbol{s}_{\mu,\nu}$ in the memory in a frequency-dependent way. The relationship among the modulation phase $\theta_{ji}^{\mathrm{SLM\#2}}(f_0)$ at SLM#2, the corresponding delay time $\tau_{ji}^{\mathrm{SLM\#2}}$, and the optical-path difference ΔL are expressed as

$$\theta_{ji}^{\mathrm{SLM\#2}}(f_0) \equiv 2\pi f_0 \tau_{ji}^{\mathrm{SLM\#2}} = 2\pi f_0 (\tau_{ji}^{\mathrm{Hebb}} - \frac{\Delta L}{c}) \tag{8.7}$$

where $\tau_{ji}^{\mathrm{SLM\#2}}$ and $\tau_{ji}^{\mathrm{Hebb}}$ are the SLM delay time corresponding to $\theta_{ji}^{\mathrm{SLM\#2}}(f_0)$ and the total connection delay in (8.5), respectively. When the carrier

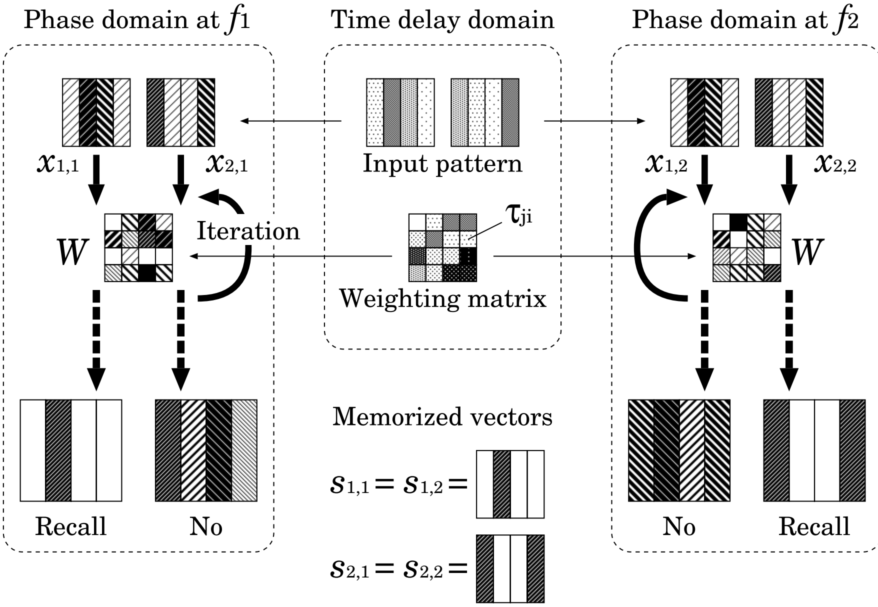

Fig. 8.5 Schematic illustration showing the frequency-dependent behavior expressed as the modulation signals on the SLM surfaces and results on the CCD. (Reprinted from Fig.4 in [204] in figure caption of Fig.8.1 with permission.)

frequency is f, the equivalent phase of the weight $\theta_{ji}(f)$ is determined frequency dependently as

$$\begin{aligned}\theta_{ji}(f) &= 2\pi f(\tau_{ji}^{\mathrm{SLM}\#2} + \frac{\Delta L}{c}) \\ &= \frac{f}{f_0}\theta_{ji}^{\mathrm{SLM}\#2}(f_0) + 2\pi f\frac{\Delta L}{c} \ .\end{aligned} \tag{8.8}$$

The behavior shown in Fig.8.5 is just an example of the intended operation of the system. The system has a certain function at a chosen frequency. From another viewpoint, we can regard the system as a fully parallel processor in the frequency domain when we use multiple optical sources simultaneously, just like the frequency-domain multiplexing (FDM) in the communications. On the other hand, if the system finds out an optimal frequency on its own with a feedback mechanism, we can recognize a type of *volition* in the system, resulting in a self-organizing context-dependent information processor [208]. The volition is described in Chapter 10 together with the concept of developmental learning.

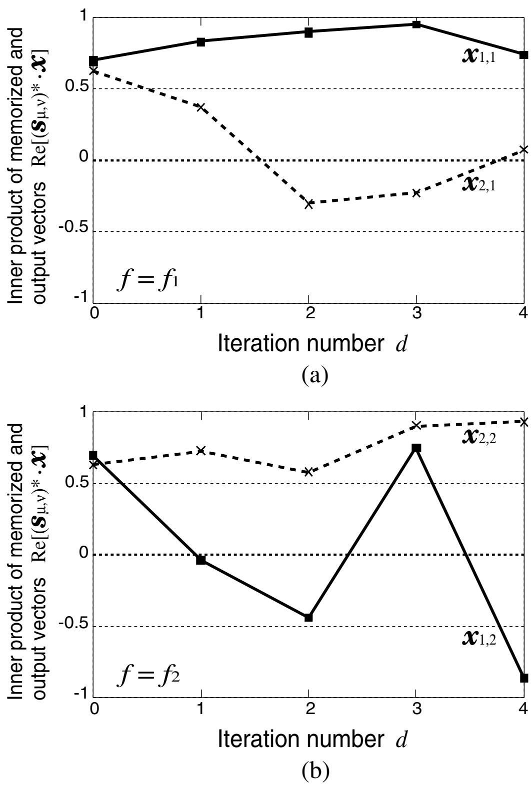

Fig. 8.6 Real part of inner products Re[$(\boldsymbol{s}_{\mu,\nu})^* \cdot \boldsymbol{x}$] versus recall iteration number. A unity inner product means a successful recall. Optical carrier frequency is (a)$f = f_1$, while (b)$f = f_2$. (Reprinted from Fig.5 in [204] in figure caption of Fig.8.1 with permission.)

8.5 Frequency-Dependent Recall Experiment

In Fig.8.2, the frequency of the light source LD#1 (wavelength $\cong$ 635[nm] ; $f_0 \cong 472$[THz]) is changed by injection current control. The frequency sensitivity is 10.5[GHz/mA]. The temperature is stabilized electrically. The resulting frequency controllability (resolution) is better than 0.016[GHz]. SLM#1 (Hamamatsu Photonics X6345) is modulated by a video input, while SLM#2 (X7665) is done by a spatially parallel optical input. However, basically we can use any modulation types of PAL-SLMs.

The optical path-length difference is about ΔL=6.8[mm], resulting in the frequency period of $c/\Delta L$ = 44.1[GHz]. Numbers of neurons and synaptic connections are 9 and 81, respectively. We choose two signal vectors $\boldsymbol{s}_{\mu,\nu}(\equiv \boldsymbol{s}_\mu(f_\nu))$ and two carrier frequencies f_ν ($\mu = 1$ only, while $\nu = 1, 2$) where we intend that $\boldsymbol{s}_{1,1}$ be recalled at f_1, and $\boldsymbol{s}_{1,2}$ at f_2, respectively. The frequencies are chosen as $f_1 = f_0$ and $f_2 = f_0 + c/(4\ \Delta L)$ so that the network behavior becomes independent even in a simple connection case.

We select two almost-orthogonal vectors to be memorized, $\boldsymbol{s}_{1,1}$ and $\boldsymbol{s}_{1,2}$. The amplitudes of all the vector elements are fixed to unity. We generate input vectors $\boldsymbol{x}_{1,1}$ and $\boldsymbol{x}_{1,2}$ by adding phase noise whose distribution is homogeneous within 30 % range of $\pm\pi$. The elements' amplitudes of the both input vectors are unity again. The learning process generates the weights w_{ji} on a PC. First, we initialize the delays τ_{ji} by choosing the initial phase θ_0 in (8.6) at random. Then the learning process based on (8.5) is iterated 1000 times, with which the weights settle at a sufficiently steady state. The learning gain K is 0.5.

Figure 8.6 shows the recall result for noisy input vectors $\boldsymbol{x}_{1,1}$ (near to $\boldsymbol{s}_{1,1}$ at f_1) and $\boldsymbol{x}_{1,2}$ (near to $\boldsymbol{s}_{1,2}$ at f_2) when one of the carrier frequencies f_1 or f_2 is chosen. In the case of f_1 shown in Fig.8.6(a), the real part of the inner product $\mathrm{Re}[(\boldsymbol{s}_{1,1})^* \cdot \boldsymbol{x}_{1,1}]$ converges almost at unity, which means that the system recalls the corresponding vector $\boldsymbol{s}_{1,1}$. On the other hand, the inner product $\mathrm{Re}[(\boldsymbol{s}_{1,2})^* \cdot \boldsymbol{x}_{1,2}]$ presents oscillatory behavior, which means failure in the recall of $\boldsymbol{s}_{1,2}$. In Fig.8.6(b), contrarily, the system recalls $\boldsymbol{s}_{1,2}$, while it does not recall $\boldsymbol{s}_{1,1}$. In this way, the frequency-dependent associative recall is realized.

8.6 Summary

We have presented a coherent optical associative memory whose behavior can be changed by the modulation of the carrier frequency. The system has a homodyne-detector structure where the signal and reference optical-path lengths are slightly different from each other. The path-length difference resulted in the frequency-dependent behavior. The carrier-frequency dependent complex-valued Hebbian rule functioned well to realize the frequency-dependent recall consistently.

This basic idea leads to future frequency-domain parallelism utilizing the vast optical-frequency bandwidth in optical neural networks. It is also applicable to future FDM systems and wavelength / wave number division multiplexed systems related to various wave phenomena. In addition, there have been several exploring ideas and analyses reporting, for example, optical frequency-multiplexed learning logic circuit [51] [52].