MATCHED FILTERS

The history of the matched filter can be traced back to more than half a century ago. In 1940s, due to World War II, radar became a very important detecting device. To enhance the performance of radar, D. O. North proposed an optimum filter for picking up signal in the case of white-noise interference (1). A little bit later, this technique was called matched filter by Van Vleck and Middleton (2). Dwork (3) and George (4) also pursued similar work. The filter has a frequency response function given by the conjugate of the Fourier transform of a received pulse divided by the spectral density of noise. However, the Dwork–George filter is only optimum for the case of unlimited observation time. It is not optimum if observations are restricted to a finite time interval. In 1952, Zadeh and Ragazzini published the work"Optimum filters for the detection of signals in noise" (5), where they described a causal filter for maximizing the signal-to-noise ratio (SNR) with respect to noise with an arbitrary spectrum for the case of unlimited observation time, and second for the case of a finite observation interval. Since then, extensive works on matched filters were done in 1950s. A thorough tutorial review paper called "An introduction to matched filters" (6) was given by Turin.

In the 1960s, due to rapid developments of digital electronics and digital computers, the digital matched filter has appeared (7–9). Turin gave another very useful tutorial paper in 1976, entitled "An introduction to digital matched filters" (10), in which the class of noncoherent digital matched filters that were matched to AM signals was analyzed.

At this time, matched filters have become a standard technique for optimal detection of signals embedded in steady-state random Gaussian noise. The theory of matched filter can be found in many textbooks (11–13).

In this article, we will briefly discuss the theory and application of matched filters. We will start with a continuous input signal case. Then, we will look at the discrete input signal case. Finally, we will provide some major applications of matched filters.

THE MATCHED FILTER FOR CONTINUOUS-TIME INPUT SIGNALS

As mentioned previously, the matched filter is a linear filter that minimizes the effect of noise while maximizing the signal. Thus, a maximal SNR can be achieved in the output. A general block diagram of matched-filter system is described in Fig. 1. To obtain the matched filter, the following conditions and restrictions are required in the system:

1. The input signal consists of a known signal $s_i(t)$ and an additive random noise process $n_i(t)$ with continuous parameter t. The corresponding output signal and noise are $s_o(t)$ and $n_o(t)$, respectively.
2. The system is linear and time invariant.
3. The criterion of optimization is to maximize the output signal-to-noise power ratio. Since noise $n_o(t)$ is random, its mean squared value $E\{n_o^2(t)\}$ is used as the output noise power.

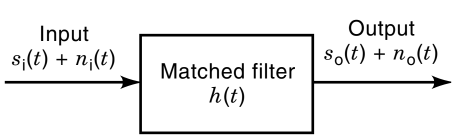

Figure 1. The block diagram of the matched filter in continuous time.

Mathematically, this criterion can be written as

$$\mathrm{SNR_o} = \frac{s_o^2(t)}{E\{n_o^2(t)\}} = \text{maximum} \tag{1}$$

at a given time t. The form of the matched filter can be derived by finding the linear time-invariant system impulse response function $h(t)$ that achieves the maximization of Eq. (1). The mathematical derivation process can be described as follows.

Since the system is assumed to be linear and time invariant, the relationship between input signal $s_i(t)$ and output signal $s_o(t)$ could be written as

$$s_o(t) = \int_{-\infty}^{t} s_i(\tau) h(t-\tau)\, d\tau \tag{2}$$

Similarly, the relationship between input noise $n_i(t)$ and output noise $n_o(t)$ could also be expressed as

$$n_o(t) = \int_{-\infty}^{t} n_i(\tau) h(t-\tau)\, d\tau \tag{3}$$

Substituting Eqs. (2) and (3) into Eq. (1), the output power SNR can be shown to be

$$\mathrm{SNR_o} = \frac{\left| \int_{-\infty}^{t} s_i(\tau) h(t-\tau)\, d\tau \right|^2}{\int_{-\infty}^{t}\int_{-\infty}^{t} R_n(\tau,\sigma) h(t-\tau) h(t-\sigma)\, d\tau\, d\sigma} \tag{4}$$

where $R_n(\tau, \sigma)$ is the autocorrelation function of the input noise $n_i(t)$ and is given by

$$R_n(\tau,\sigma) = E\{n_i(\tau) n_o(\sigma)\} \tag{5}$$

Now the unknown function $h(t)$ can be found by maximizing Eq. (4). To achieve this goal from Eqs. (4) and (5), one can see that the optimum $h(t)$ (i.e., the matched-filter case) will depend on the noise covariance $R_n(\tau, \sigma)$. Since $h(t)$ is required to be time invariance [i.e., $h(t - \tau)$ instead of $h(t, \tau)$], the noise at least has to be wide-sense stationary [i.e., $R_n(t, \tau) = R_n(t - \tau)$]. To obtain the optimum filter, based on the linear system theory (13), Eq. (4) can be rewritten as

$$\mathrm{SNR_o} = \frac{\left| \int_{-\infty}^{\infty} H(f) S(f) e^{i\omega t_0} df \right|^2}{\int_{-\infty}^{\infty} |H(f)|^2 P_n(f)\, df} \tag{6}$$

where $H(f) = \mathscr{F}[h(t)]$ is the Fourier transform of the impulse response function $h(t)$ (i.e., the transfer function of the system), $S(f) = \mathscr{F}[s(t)]$ is the Fourier transform of the known input signal $s(t)$, $\omega = 2\pi f$ is angular frequency, t_o is the sampling time when SNR_o is evaluated, and $P_n(f)$ is the noise power spectrum density function. To find the particular $H(f)$ that maximizes SNR_o, we can use the well-known *Schwarz inequality,* which is

$$\left|\int_{-\infty}^{\infty} A(f)B(f)\,df\right|^2 \leq \int_{-\infty}^{\infty} |A(f)|^2\,df \int_{-\infty}^{\infty} |B(f)|^2\,df \qquad (7)$$

where $A(f)$ and $B(f)$ may be complex functions of the real variable f. Furthermore, equality is obtained only when

$$A(f) = kB^*(f) \qquad (8)$$

where k is any arbitrary constant and $B^*(f)$ is the complex conjugate of $B(f)$. By using the Schwarz inequality to replace the numerator on the right-hand side of Eq. (6) and letting $A(f) = H(f)\sqrt{P_n(f)}$ and $B(f) = S(f)e^{i\omega t_0}/\sqrt{P_n(f)}$, Eq. (6) becomes

$$\mathrm{SNR}_o \leq \frac{\displaystyle\int_{-\infty}^{\infty} |H(f)|^2 P_n(f)\,df \int_{-\infty}^{\infty} \frac{|S(f)|^2}{P_n(f)}\,df}{\displaystyle\int_{-\infty}^{\infty} |H(f)|^2 P_n(f)\,df} \qquad (9)$$

In addition, because $P_n(f)$ is a non-negative real function, Eq. (9) can be further simplified into

$$\mathrm{SNR}_o \leq \int_{-\infty}^{\infty} \frac{|S(f)|^2}{P_n(f)}\,df \qquad (10)$$

The maximum SNR_o is achieved when $H(f)$ is chosen such that equality is attained. This occurs when $A(f) = kB^*(f)$, that is,

$$H(f)\sqrt{P_n(f)} = \frac{kS^*(f)e^{-i\omega t_0}}{\sqrt{P_n(f)}} \qquad (11)$$

Based on Eq. (11), the transfer function of the matched filter $H(f)$ can be derived as

$$H(f) = k\frac{S^*(f)}{P_n(f)}e^{-i\omega t_0} \qquad (12)$$

The corresponding impulse response function $h(t)$ can be easily obtained by taking the inverse Fourier transform of Eq. (12), that is,

$$h(t) = \int_{-\infty}^{\infty} H(f)e^{i\omega t}\,df = \int_{-\infty}^{\infty} k\frac{S^*(f)}{P_n(f)}e^{iw(t-t_0)}\,df \qquad (13)$$

In the matched-filter case, the output SNR_o is simply expressed as

$$\max\{\mathrm{SNR}_o\} = \int_{-\infty}^{\infty} \frac{|S(f)|^2}{P_n(f)}\,df \qquad (14)$$

For the case of *white noise*, the $P_n(f) = N_0/2$ becomes a constant. Substituting this constant into Eq. (13), the impulse response of the matched filter has a very simple form

$$h(t) = Cs_i(t_0 - t), \qquad (15)$$

where C is an arbitrary real positive constant, t_0 is the time of the peak signal output. This last result is one of the reasons why $h(t)$ is called a matched filter since *the impulse response is "matched" to the input signal in the white-noise case.*

Based on the preceding discussion, the matched filter theorem can be summarized as follows: The *matched filter* is the linear filter that maximizes the output signal-to-noise power ratio and has a transfer function given by Eq. (12).

In the previous discussion, the problem of the physical realizability is ignored. To make this issue easier, we will start with the white-noise case. In this case, the matched filter is physically realizable if its impulse response vanishes for negative time. In terms of Eq. (15), this condition becomes

$$h(t) = \begin{cases} 0, & t < 0 \\ s_i(t_0 - t), & t \geq 0 \end{cases} \qquad (16)$$

where t_0 indicates the filter delay, or the time between when the filter begins receiving the input signal and when the maximum response occurs. Equation 16 also implies that $s(t) = 0$, $t > t_0$, i.e., the filter delay must be greater than the duration of the input signal. As an example, let us consider the following signal corrupted by additive white noise. The known input signal has the form

$$s_i(t) = \begin{cases} Be^{bt}, & t < 0,\ B, b > 0 \\ 0, & t \geq 0 \end{cases} \qquad (17)$$

Substituting Eq. (17) into Eq. (16), the impulse response of matched filter $h(t)$ is

$$h(t) = \begin{cases} Be^{b(t_0-t)}, & t \geq t_0 \\ 0, & t < t_0 \end{cases} \qquad (18)$$

The physical realizability requirement can be simply satisfied by letting $t_0 \geq 0$. The simplest choice is $t_0 = 0$ so that $h(t)$ has a very simple form

$$h(t) = \begin{cases} Be^{-bt}, & t \geq 0 \\ 0, & t < 0 \end{cases} \qquad (19)$$

The output signal $s_o(t)$ of the system can be obtained by substituting Eqs. (17) and (19) into Eq. (2). The calculated result of $s_o(t)$ is

$$s_o(t) = \begin{cases} \dfrac{B^2}{2b}e^{bt}, & t < 0 \\ \dfrac{B^2}{2b}e^{-bt}, & t > 0 \end{cases} \qquad (20)$$

To give an intuitive feeling about the results above, Figs. 2(a)–2(c) illustrate the input signal $s_i(t)$, matched filter $h(t)$, and output signal $s_o(t)$. From Fig. 2(c), indeed, one can get the maximum signal at time $t = t_0 = 0$. Note that, in Fig. 2, we have assumed the following parameters: $B = b = 1$. The physical implementation of this simple matched filter can be achieved by using a simple RC circuit as illustrated in Fig. 3, in which the time constant of the RC circuit is $RC = 1$.

In many real cases, the input noise may not be white noise and the designed matched filter may be physically unrealizable. Now, let us look at another example with color noise (11). We assume that the input signal $s_i(t)$ has a form of

$$s_{\rm i}(t) = \begin{cases} e^{-t/2} - e^{-3t/2}, & t > 0 \\ 0, & t < 0 \end{cases} \tag{21}$$

and the input noise is wide-sense stationary with power spectral density

$$P_{\rm n}(f) = \frac{4}{1 + 4(2\pi f)^2} \tag{22}$$

To obtain the matched filter, first, we take the Fourier transform of input signal $s_i(t)$. Based on Eq. (21), the spectrum of

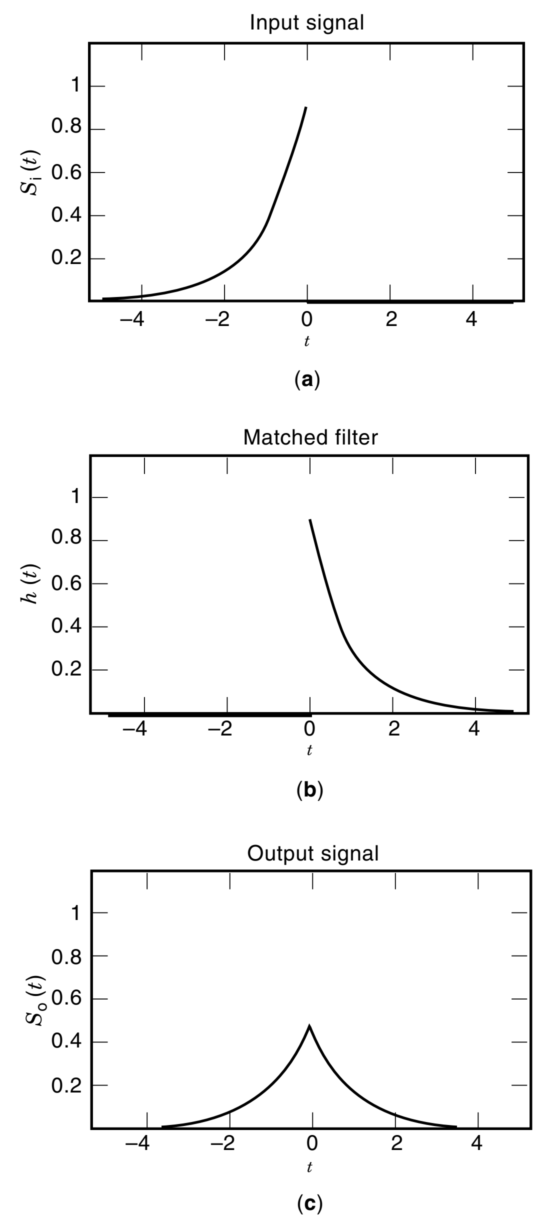

Figure 2. A simple matched-filter example for white noise and continuous time. (a) Input signal. (b) Matched filter. (c) Output signal. This figure provides an intuitive feeling about using matched filter for continuous time signal processing.

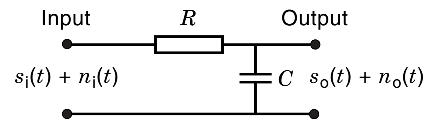

Figure 3. Implementation of the discussed example in text using a *RC* circuit. This figure shows that the continuous time matched filter can be physically implemented by using a simple RC circuit.

input signal can be shown to be

$$S_{\rm i}(2\pi f) = \frac{4}{(1 + i4\pi f)(3 + i4\pi f)} \tag{23}$$

Substituting Eqs. (22) and (23) into Eq. (12), the transfer function of matched filter $H(f)$ can be derived as

$$\begin{aligned} H(f) &= k\frac{S^*(f)}{P_{\rm n}(f)} e^{-i\omega t_0} \\ &= k\frac{1 + i4\pi f}{3 - i4\pi f} e^{-i\omega t_0} \end{aligned} \tag{24}$$

To simplify the expression, we let the arbitrary constant $k = 1$ for the later derivations. By taking the inverse Fourier transform of Eq. (24), the impulse response of the matched filter is

$$h(t) = -\delta(t - t_0) + 2e^{(t-t_0)3/2}u(t_0 - t) \tag{25}$$

where $u(t)$ is the unit step function. Note that this filter is not physically realizable because it has a nonzero value for $t < 0$. To solve this problem, one method is to take a realizable approximation by letting $h(t) = 0$ for $t < 0$. In this case, the approximated matched filter $h_a(t)$ can be expressed as

$$\begin{aligned} h_{\rm a}(t) &= h(t)u(t) \\ &= -\delta(t - t_0) + 2e^{(t-t_0)3/2}u(t_0 - t)u(t) \end{aligned} \tag{26}$$

Then, the output spectrum $S_o(f)$ of the output signal $s_o(t)$ due to this approximated matched filter is

$$S_{\rm o}(f) = S_{\rm i}(f)H_{\rm a}(f) \tag{27}$$

where $H_a(f)$ is the Fourier transform of $h_a(t)$. Again, by taking the inverse Fourier transform of Eq. (27), the output signal $s_o(t)$ can be derived as

$$\begin{aligned} s_{\rm o}(t) = &- e^{-3t_0/2}e^{-t/2}u(t) + \tfrac{2}{3}e^{-3(t_0+t)/2}u(t) \\ &- \tfrac{1}{3}e^{-3(t_0-t)/2}u(-t) + \tfrac{1}{3}e^{-3(t_0-t)/2} \end{aligned} \tag{28}$$

Again, to have an intuitive feeling about this example, Fig. 4 illustrates the input signal $s_i(t)$, the ideal physically unrealizable matched filter $h(t)$, approximated realizable matched filter $h_a(t)$, and the output signal $s_o(t)$ obtained with the approximated filter.

For the purpose of convenience, we assume that $t_0 = 1$ for these plots. From Fig. 4(d), one can see that, indeed, the output signal has a maximum value at $t = t_0 = 1$. However, there is no guarantee that this approximated filter is the optimum filter. In fact, it is shown that, a better output SNR can be

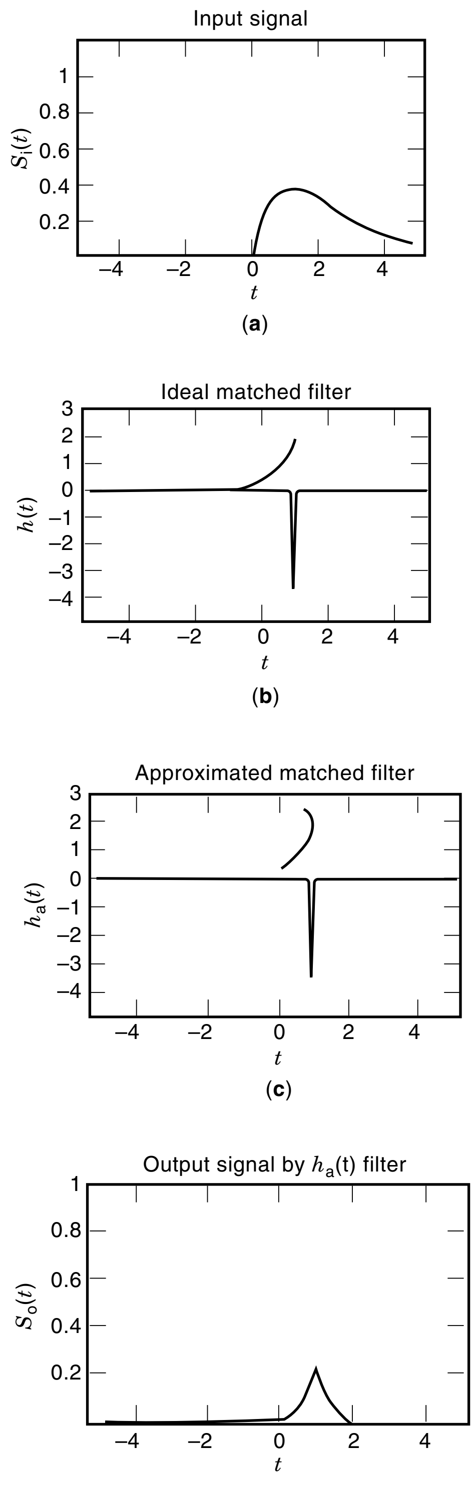

Figure 4. Example of a matched filter with color noise for a continuous-time signal. (a) Input signal. (b) Ideal matched filter. (c) Approximated matched filter. (d) Output signal with approximated matched filter. This figure illustrates how to deal with color noise with matched filter.

achieved for this problem if the prewhitening technique is employed for the signal detection (11).

Before the end of this section, we want to point out that, in practical terms, it is impossible to design an optimal matched filter for any signal which has an infinite time duration because it requires infinite delay time. However, the above examples are very fast exponential decaying signal, for which one can make the delay time long enough so that optimality can be approached to any desired degree. In other words, the practically "realizable" matched filter only exists for a time limited function. From this point of view, mathematically speaking, the above two examples both have infinite time duration. Thus, even for the second example, it becomes unrealizable. However, since there are extremely fast exponentially decaying signals, optimality can be achieved to any desired degree. In this sense, example 2 can be treated as "realizable" matched filter. Finally, since t_0 represents the delay time of the filter in the above examples, in practice, it must be selected longer than the time duration of the target signal. For the sole purpose of simplicity, in the above examples, the simple values (that are not strict in the mathematical sense) of t_0 are selected.

THE MATCHED FILTER FOR DISCRETE-TIME INPUT SIGNALS

In recent years, with rapid developments of the digital computers, digital signal processing becomes more and more powerful. Some major advantages of using digital signals as compared to their analog forms are the high accuracy, high flexibility, and high robustness. Right now, the matched filter can be easily implemented with the digital computer in real time. To implement the filter with digital computer, one has to deal with the discrete signal instead of continuous signal. In this case, for the same linear time invariant system as described in Fig. 1, the relationship between the output signal $s_o(t)$ and input signal $s_i(t)$ has changed from the continuous-time form Eq. (2) to the following discrete time form (11):

$$s_{oj} = \sum_{k=-\infty}^{j} h_{j-k} s_{ik} \tag{29}$$

where s_{ik} represents the input signal at time k ($k = 0, \pm 1, \pm 2, \ldots$), h_k is the discrete impulse response function of the linear, time-invariant matched filter, and s_{oj} is the corresponding discrete output signal at time j. In other words, the integration in Eq. (2) has been replaced by the summation in Eq. (29). Similarly, in the discrete-time case, the Eq. (3) is rewritten as

$$n_{oj} = \sum_{k=-\infty}^{j} h_{j-k} n_{ik} \tag{30}$$

Again, our objective is to find the optimum form of matched filter so that the output signal-to-noise power ratio will be maximum at some time q. Mathematically, it can be written as

$$\mathrm{SNR_o} = \frac{s_{oq}^2}{E\{n_{oq}^2\}} = \text{maximum} \tag{31}$$

To find h_k, we let maximim SNR, symbolized as $\mathrm{SNR}_{\mathrm{omax}}$, equal a constant $1/\alpha$. Since $\mathrm{SNR}_{\mathrm{omax}}$ represents the maximum *power* ratio, it has to be larger than 0, i.e., $\alpha > 0$. Substituting this assumption into Eq. (31), one can obtain

$$\mathrm{SNR}_{\mathrm{o}} = \frac{s_{\mathrm{o}q}^2}{E\{n_{\mathrm{o}q}^2\}} \leq \mathrm{SNR}_{\mathrm{omax}} = \frac{1}{\alpha} \tag{32}$$

Equation (32) can be rewritten as

$$E\{n_{\mathrm{o}q}^2\} - \alpha s_{\mathrm{o}q}^2 = C \geq 0 \tag{33}$$

where C is a positive real constant and the equality holds only for the optimum matched filter. To find this matched filter, one can substitute Eqs. (29) and (30) into Eq. (33). Then, one can get

$$\sum_{k=-\infty}^{q} \sum_{j=-\infty}^{q} R_n(k-j)h_{q-k}h_{q-j} - \alpha \left| \sum_{k=-\infty}^{q} s_{ik}h_{q-k} \right|^2 = C' \geq 0 \tag{34}$$

where $R_n(k-j)$ is the autocorrelation function of the input noise n_{i} and C' is another positive constant. Note that, in the process of deriving Eq. (34), we already assume that the input noise is at least wide-sense stationary. Under this assumption, the following condition holds:

$$R_n(k-j) = R_n(j-k) = E\{n_k n_j\} \tag{35}$$

Since the equality holds in Eq. (34) when h_k is an optimum matched filter regardless of the detail forms of input signal and noise, it can be shown that the following equation can be derived under this condition (11):

$$\sum_{j=-\infty}^{q} R_n(k-j)h_{q-j} = s_{ik} \tag{36}$$

To make our discussion easy to be understood, we start with the simple white-noise case. In this case, the autocorrelation function can be simply written as

$$R_n(k) = \begin{cases} N_{\mathrm{o}}/2, & k = 0 \\ 0, & k \neq 0 \end{cases} \tag{37}$$

Substituting Eq. (37) into Eq. (36), we obtain

$$\frac{N_0}{2} h_{q-k} = s_{ik} \tag{38}$$

To get a simpler expression of h_k, we let $\mathrm{l} = q - k$. Then, Eq. (38) can be rewritten as

$$h_{\mathrm{l}} = \frac{2}{N_{\mathrm{o}}} s_{\mathrm{i}(q-\mathrm{l})} \tag{39}$$

Comparing Eq. (39) with Eq. (15), one can see that Eq. (39) is exactly the discrete form of Eq. (19).

As an example, let us consider a discrete input signal s_{ik} to be given by

$$s_{ik} = \begin{cases} e^k, & k \leq 0 \\ 0, & k > 0 \end{cases} \tag{40}$$

The input noise is additive white noise with autocorrelation function

$$R_n(k) = \begin{cases} \dfrac{N_{\mathrm{o}}}{2} = 1, & k = 0 \\ 0, & k \neq 0 \end{cases} \tag{41}$$

Substituting Eqs. (40) and (41) into Eq. (39), we have

$$h_k = s_{i(q-k)} = \begin{cases} e^{q-k}, & k \geq q \\ 0, & k < q \end{cases} \tag{42}$$

Substituting Eqs. (40) and (42) into Eq. (29), the output signal $s_{\mathrm{o}j}$ can be derived as

$$s_{\mathrm{o}j} = \frac{1}{1-e^2} e^{-|j|}, \quad j = 0, \pm 1, \pm 2, \ldots \tag{43}$$

Again, to have an intuitive feeling about this result, Figs. 5(a), 5(b), and 5(c) illustrate the discrete input signal s_{ik}, the discrete matched filter h_k, and the discrete output signal $s_{\mathrm{o}j}$. To get the simplest form in Fig. 5, we assumed $q = 0$.

Equation (36) only deals with the physically realizable case. In general, Eq. (36) will be written as (11)

$$\sum_{j=-\infty}^{\infty} R_n(k-j)h_{q-j} = s_{ik} \tag{44}$$

In Eq. (44), we have replaced the limit q by ∞. To get the general form of a discrete matched filter for nonwhite noise, we take the z transform on both size of Eq. (44) and use the convolution theorem for z transforms. Then, Eq. (44) can be shown to be (14)

$$z^{-q} P_n(z) H(1/z) = S_{\mathrm{i}}(z) \tag{45}$$

where

$$\begin{aligned} P_n(z) &= \sum_{k=-\infty}^{\infty} R_n(k) z^{-k} \\ H(z) &= \sum_{k=-\infty}^{\infty} h_k z^{-k} \\ S_{\mathrm{i}}(z) &= \sum_{k=-\infty}^{\infty} s_{\mathrm{i}k} z^{-k} \end{aligned} \tag{46}$$

represent the power density spectrum, z transform of a discrete matched filter, and z transform of discrete input signal. To obtain the z transform of a discrete matched filter $H(z)$, Eq. (45) is rewritten as

$$H(z) = \frac{S_{\mathrm{i}}(1/z)}{P_n(z)} z^{-q} \tag{47}$$

In deriving Eq. (47), we have used a property of power density spectrum, that is, $P_n(z) = P_n(1/z)$. Theoretically speaking, the discrete matched filter in the time domain (that is, the impulse response function of discrete matched filter) can be obtained by taking the inverse z transform of Eq. (47) (14), that

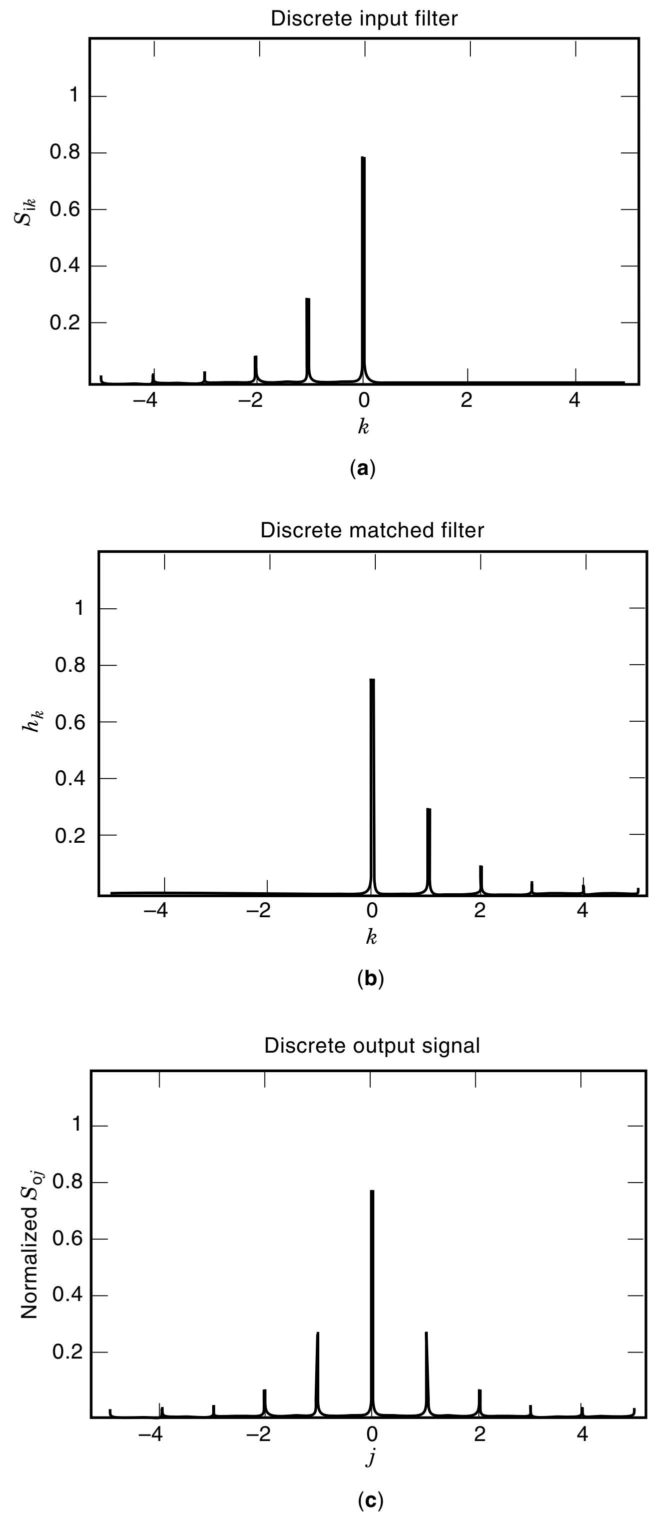

Figure 5. An example of matched filter for the white noise in discrete time. (a) Discrete input signal. (b) Discrete matched filter. (c) Discrete output signal. This figure gives an intuitive feeling about using matched filter for discrete time signal processing.

is

$$h_k = \frac{1}{2\pi i}\oint_\Gamma H(z)z^{k-1}\,dz = \frac{1}{2\pi i}\oint_\Gamma \frac{S_i(1/z)}{P_n(z)}z^{-q}z^{k-1}\,dz \qquad (48)$$

where Γ represents a counterclockwise contour in the region of convergence of $H(z)$ enclosing the origin. Note that, similar to the continuous-time case, the discrete matched filter defined by Eqs. (47) and (48) may not be realizable for arbitrary input signal and noise because h_k will not vanish for negative values of the index k.

To implement the color noise effectively, the prewhitening technique is used (11). In this approach, the input power density spectrum $P_n(z)$ is written as the multiplication of two facts $P_n^+(z)$ and $P_n^-(z)$, that is,

$$P_n(z) = P_n^+(z)P_n^-(z) \qquad (49)$$

where $P_n^+(z)$ has all of the poles and zeros of $P_n(z)$ that are inside the unit circle and $P_n^-(z)$ has all the poles and zeroes of $P_n(z)$ that are outside the unit circle. By this definition, it is easy to show that

$$P_n^+(z) = P_n^-(1/z) \qquad (50)$$

Note that, in the time domain, $P_n^+(z)$ corresponds to a discrete-time input signal that vanishes for all times $t < 0$. Similarly, $P_n^-(z)$ corresponds to a discrete-time input signal that vanishes for all time $t > 0$. This property can be easily proven in the following way. Assume that n_k is a discrete-time function that vanishes on the negative half-line; that is,

$$n_k = 0, \quad k < 0 \qquad (51)$$

If n_k is absolutely summable, that is, if

$$\sum_{k=-\infty}^{\infty} |f_k| = \sum_{k=0}^{\infty} |f_k| < \infty \qquad (52)$$

then, the z transform of this discrete function n_k becomes

$$N(z) = \sum_{k=-\infty}^{\infty} n_k z^{-k} = \sum_{k=0}^{\infty} n_k z^{-k} \qquad (53)$$

From Eq. (53), one can see that the function $N(z)$ exists everywhere when $|z| \geq 1$. Hence, the poles of $N(z)$ will all be inside the unit circle. Thus, $P_n^+(z)$ corresponds to a discrete-time input signal that vanishes for all time $t < 0$. Similarly, it can be shown that $P_n^-(z)$ corresponds to a discrete-time input signal that vanishes for all time $t > 0$. Assume $H_{\rm pw}(z)$ is the prewhitening filter. Based on the definition of prewhitening filter, $H_{\rm pw}(z)$ for the noise power spectrum $P_n(z)$ must satisfy (11)

$$\left[P_{\rm n}^+(z)H_{\rm pw}(z)\right]\left[P_{\rm n}^-(z)H_{\rm pw}(1/z)\right] = 1 \qquad (54)$$

From Eq. (54), one can conclude that the prewhitening filter is

$$H_{\rm pw}(z) = \frac{1}{P_{\rm n}^+(z)} \qquad (55)$$

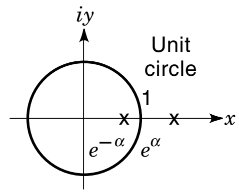

Figure 6. Pole locations of the power spectrum density.

Because $P_n^+(z)$ corresponds to a discrete-time input signal that vanishes for all time $t < 0$, the impulse response h_{pwk} of this prewhitening filter will vanish for $k < 0$. Hence, the prewhitening filter $H_{pw}(z)$ is physically realizable. For example, let us consider a color noise with power density spectrum

$$P_n(z) = \frac{N_0}{2}\frac{e^{2\alpha}}{(e^{\alpha}-z^{-1})(e^{\alpha}-z)}, \quad \alpha > 0 \tag{56}$$

Equation (56) shows that $P_n(z)$ contains poles both inside and outside the unit circle. As discussed in the early part of this section, this $P_n(z)$ can be written as the multiplication of $P_n^+(z)$ and $P_n^-(z)$. For the purpose of convenience and symmetry, we let

$$\begin{aligned} P_n^+(z) &= \sqrt{\frac{N_o}{2}}\frac{e^{\alpha}}{e^{\alpha}-z^{-1}} \\ P_n^-(z) &= \sqrt{\frac{N_o}{2}}\frac{e^{\alpha}}{e^{\alpha}-z} \end{aligned} \tag{57}$$

Based on Eq. (57), it is easy to show that $P_n^+(z)$ has a pole at $z = e^{-\alpha}$ (11). Since $\alpha > 0$, $z = e^{-\alpha} < 1$. In other words, this pole is inside the unit circle. Similarly, $P_n^-(z)$ has a pole at $z = e^{\alpha}$ that is a real number greater than unity. Figure 6 illustrates these pole locations of above power spectral density $P_n(z)$ in the complex plane. In the figure, we assume that $z = x + iy$. For this particular example, the poles are on the real axis.

When applying this prewhitening technique to the discrete matched filter, Eq. (47) will be rewritten as

$$H(z) = \frac{1}{P_n^+(z)}\left(\frac{S_i(1/z)}{P_n^-(z)}z^{-q}\right) \tag{58}$$

Equation (58) is the multiplication of two terms. The first term is the prewhitening filter and the second term is the remainder of the unrealizable matched filter. Note that this multiplication is equivalent to put two linear systems in tandem. Similar to the continuous-time case, this remaining unrealizable filter can be made realizable by throwing away the part that does not vanish for negative time.

APPLICATIONS OF A MATCHED FILTER

As mentioned in the first part of this article, the major application of the matched filter is to pick up the signal in a noisy background. As long as the noise is additive, wide-sense stationary, and the system is linear and time invariant, the matched filter can provide a maximum output signal-to-noise power ratio. The signal can be a time signal (e.g., radar signal) or spatial signals (e.g., images). To have an intuitive feeling about the time signal detection by a matched filter, let us consider the following simple example. For the purpose of convenience, the ideal input signal is assumed to be a normalized sinc function, that is, $\text{sinc}(t) = \sin(\pi t)/\pi t$, as shown in Fig. 7(a). This ideal signal is embedded into an additive broadband white noise. The corrupted signal is shown in Fig. 7(b). Figure 7(c) shows the system output when this corrupted

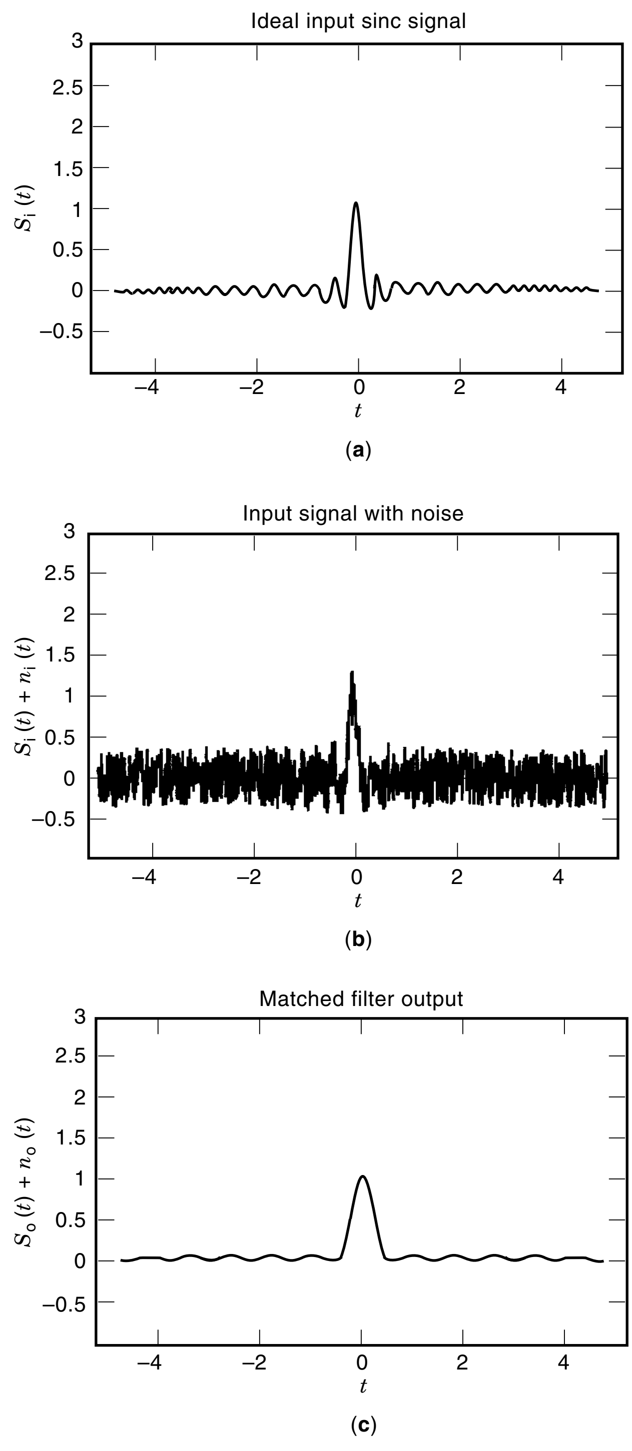

Figure 7. Results of the matched filter acting on an input signal with sinc function embedded in white noise. (a) Ideal input signal. (b) Signal with noise. (c) Matched-filter output.

signal passes through the matched filter. From Fig. 7(c), one can see that the much better signal-to-noise power ratio can be achieved by applying matched filter for the signal detection as long as the noise is additive at least wide-sense stationary noise.

Besides applying matched filters for the time-signal detection (such as the radar signal previously mentioned), they can also be used for spatial signal detection (15–17). In other words, we can use a matched filter to identify specific targets under the noisy background. Thousands of papers have been published in this field. To save space, here, we just want to provide some basic principles and simple examples of it. Since spatial targets, in general, are two-dimensional signals, the equations developed for the one-dimensional time signal needs to be extended into the two-dimensional spatial signal. Note that when matched filter is applied to the 2-D spatial (or image) identification, this filtering process can be described simply as a cross-correlation of a larger target image (including the noisy background) with a smaller filter kernel. To keep the consistency of the mathematical description, a similar derivation process (used for the 1-D time-signal case) is employed for the 2-D spatial signal. Assume that the target image is a two-dimensional function $s(x, y)$ and this target image is embedded into a noisy background with noise distribution $n(x, y)$. Thus, the total detected signal $f(x, y)$ is

$$f(x,y) = s(x,y) + n(x,y) \tag{59}$$

Similar to the one-dimensional time signal case, if $f(x, y)$ is a Fourier-transformable function of space coordinates (x, y) and $n(x, y)$ is an additive wide-sense stationary noise, the matched filter exists. It can be shown that $H(p, q)$ has a form of (15)

$$H(p,q) = k\frac{S^*(p,q)}{N(p,q)} \tag{60}$$

where $S^*(p, q)$ is the complex conjugate of the signal spectrum, $N(p, q)$ is the spectral density of the background noise, k is a complex constant, and (p, q) are corresponding spatial angular frequencies. Mathematically, $S(p, q)$ and $N(p, q)$ are expressed as

$$\begin{aligned} S(p,q) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{\infty} s(x,y)e^{-i(px+qy)}\,dx\,dy \\ N(p,q) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{\infty} n(x,y)e^{-i(px+qy)}\,dx\,dy \\ F(p,q) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{\infty} \big[s(x,y)+n(x,y)\big]e^{-i(px+qy)}\,dx\,dy \\ &= S(p,q) + N(p,q) \end{aligned} \tag{61}$$

For the purpose of simplicity, we assume that the input noise $n(x, y)$ is white noise. In this case, Eq. (60) is reduced to the simpler form

$$H(p,q) = k'S^*(p,q) \tag{62}$$

where k' is another constant. Now, assume that there is an input unknown target $t(x, y)$. Then, the corresponding spectrum is $T(p, q)$. When this input target passes through the matched filter $H(p, q)$ described by Eq. (62), the system output in the spectrum domain, that is, (p, q) domain, becomes

$$T(p,q)H(p,q) = K'T(p,q)S^*(p,q) \tag{63}$$

Assume that the final system output is $g(x', y')$, where (x', y') are the spatial coordinates in the output spatial domain. Based on the discussion in the section titled "The Matched Filter for Continuous-Time Input Signals," $g(x', y')$ can be obtained by taking the inverse Fourier transform of Eq. (63), that is,

$$g(x',y') = \int_{-\infty}^{+\infty}\int_{-\infty}^{\infty} T(p,q)S^*(p,q)e^{i(px'+qy')}\,dp\,dq \tag{64}$$

In Eq. (64), if the input unknown function $t(x, y)$ is the same as the prestored function $s(x, y)$, Eq. (64) becomes

$$\begin{aligned} g(x',y') &= \int_{-\infty}^{+\infty}\int_{-\infty}^{\infty} S(p,q)S^*(p,q)e^{i(px'+qy')}\,dp\,dq \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{\infty} |S(p,q)|^2 e^{i(px'+qy')}\,dp\,dq \end{aligned} \tag{65}$$

In this case, the system output $g(x', y')$ is the Fourier transform of the power spectrum $|S(p, q)|^2$, which is an entirely positive real number so that it will generate a big output at the original point (0, 0). Notice that, in recent years, due to the rapid development of digital computers, most 2-D filtering can be carried out digitally at relatively fast speed. However, to have an intuitive feeling about 2-D filtering, an optical description about this filtering process that was widely used in the earlier stage of image identification (15) is provided. Optically speaking, the result described by Eq. (65) can be explained in the following way. When the input target $t(x, y)$ is same as the stored target $s(x, y)$, all the curvatures of the incident target wave are exactly canceled by the matched filter. Thus, the transmitted field, that is, $T(p, q)S^*(p, q)$, in the frequency domain, is a plane wave (generally of nonuniform intensity). In the final output spatial domain, this plane wave is brought to a bright focus spot $g(0, 0)$ by the inverse Fourier transform as described in Eq. (65). However, when the input signal $t(x, y)$ is not $s(x, y)$, the wavefront curvature will in general not be canceled by the matched filter $H(p, q)$ in the frequency domain. Thus, the transmitted light will not be brought to a bright focus spot in the final output spatial domain. *Thus, the presence of the signal $s(x, y)$ can conceivably be detected by measuring the intensity of the light at the focal point of the output plane.* If the input target $s(x, y)$ is not located at the center, the output bright spot simply shifts by a distance equal to the distance shifted by $s(x, y)$. Note that this is the shift-invariant property of the matched filter. The preceding description can also mathematically be shown by Schwarz's inequality. Based on the cross-correlation theorem of Fourier transform (17), Eq. (64) can also be written in the spatial domain as

$$g(x',y') = \int_{-\infty}^{+\infty}\int_{-\infty}^{\infty} t(x,y)s^*(x-x',y-y')\,dx\,dy \tag{66}$$

which is recognized to be the cross-correlation between the stored target $s(x, y)$ and the unknown input target $t(x, y)$. By

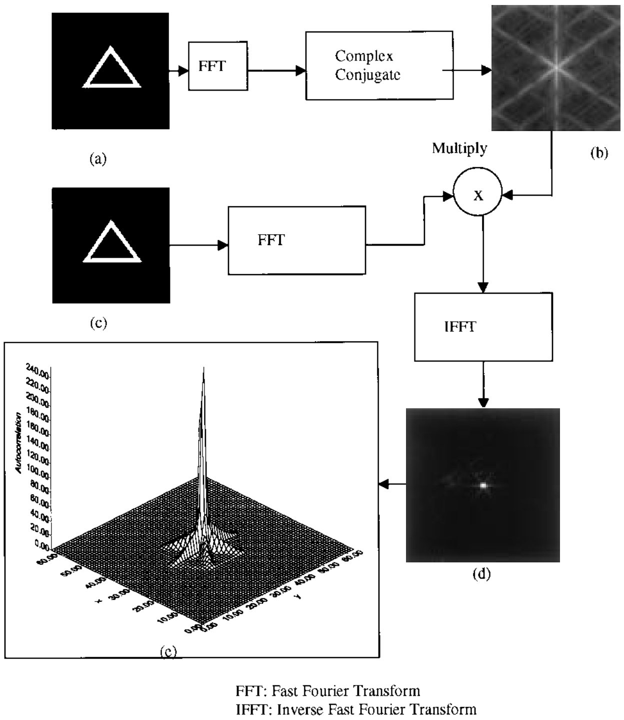

Figure 8. Autocorrelation results of the matched filter application to pattern recognition. (a) Stored training image. (b) Absolute value of the matched filter. (c) Unknown input target. (d) Autocorrelation intensity distribution. (e) Three-dimensional surface profile of autocorrelation intensity distribution. This figure shows that there is a sharp correlation peak for autocorrelation.

applying Schwarz's inequality into Eq. (66), we have

$$\left|\int_{-\infty}^{+\infty}\int_{-\infty}^{\infty} t(x,y)s^*(x-x',y-y')\,dx\,dy\right|^2 \leq \int_{-\infty}^{+\infty}\int_{-\infty}^{\infty} |t(x,y)|^2\,dx\,dy \int_{-\infty}^{+\infty}\int_{-\infty}^{\infty} |s(x-x',y-y')|^2\,dx\,dy \tag{67}$$

with the equality if and only if $t(x, y) = s(x, y)$. Because the integral limit is $\pm\infty$ in Eq. (67), by letting $x = x - x'$, $y = y - y'$, we have

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{\infty} |s(x-x',y-y')|^2\,dx\,dy = \int_{-\infty}^{+\infty}\int_{-\infty}^{\infty} |s(x,y)|^2\,dx\,dy \tag{68}$$

Substituting Eq. (68) into Eq. (67), we have

$$\left|\int_{-\infty}^{+\infty}\int_{-\infty}^{\infty} t(x,y)s^*(x-x',y-y')\,dx\,dy\right|^2 \leq \int_{-\infty}^{+\infty}\int_{-\infty}^{\infty} |t(x,y)|^2\,dx\,dy| \int_{-\infty}^{+\infty}\int_{-\infty}^{\infty} |s(x,y)|^2\,dx\,dy \tag{69}$$

with the equality if and only if $t(x, y) = s(x, y)$. To recognize the input target, we can use the normalized correlation intensity function as the *similarity criterion* between the unknown input target and the stored target. The normalized correlation intensity function is defined as

$$\frac{\left|\int_{-\infty}^{+\infty}\int_{-\infty}^{\infty} t(x,y)s^*(x-x',y-y')\,dx\,dy\right|^2}{\int_{-\infty}^{+\infty}\int_{-\infty}^{\infty} |t(x,y)|^2\,dx\,dy \int_{-\infty}^{+\infty}\int_{-\infty}^{\infty} |s(x,y)|^2\,dx\,dy} \tag{70}$$

Based on Eq. (69), we obtain

$$\frac{\left|\int_{-\infty}^{+\infty}\int_{-\infty}^{\infty} t(x,y)s^*(x-x',y-y')\,dx\,dy\right|^2}{\int_{-\infty}^{+\infty}\int_{-\infty}^{\infty} |t(x,y)|^2\,dx\,dy \int_{-\infty}^{+\infty}\int_{-\infty}^{\infty} |s(x,y)|^2\,dx\,dy} \leq 1 \tag{71}$$

with the equality if and only if $s(x, y) = t(x, y)$. Thus, one can conclude that the normalized correlation intensity function has a maximum value 1 when the unknown input target $t(x, y)$ is same as the stored target $s(x, y)$. In other words, if there is a 1 detected in the normalized correlation intensity function, we know that the unknown input target is just our stored target. Therefore, this unknown target is recognized.

Again, to have an intuitive feeling about the pattern recognition with matched filter, let us look at the following example. Figure 8(a) shows a triangle image that is used to construct the matched filter. Mathematically speaking, this image is $s(x, y)$. Then, the matched filter $S^*(p, q)$ is synthe-

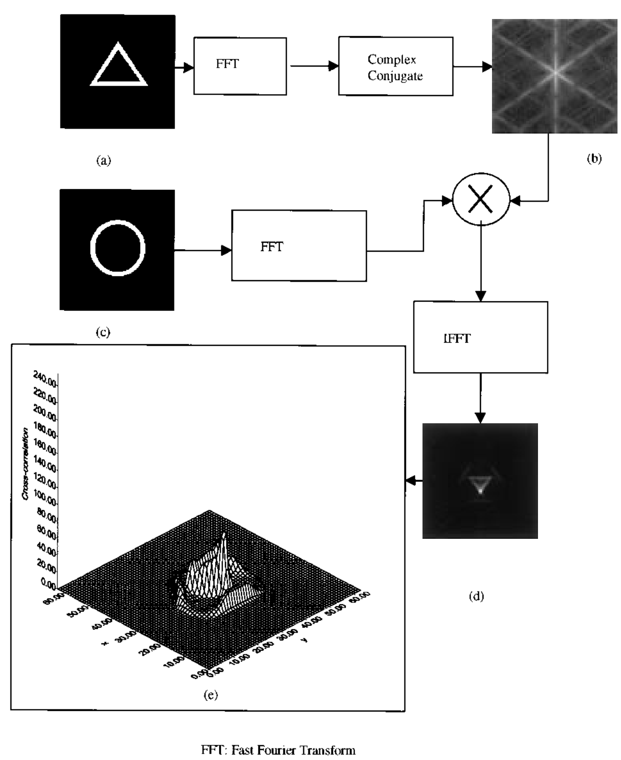

Figure 9. Cross-correlation results of the matched filter applied to pattern recognition. (a) Stored training image. (b) Absolute value of the matched filter. (c) Unknown input target. (d) Cross-correlation intensity distribution. (e) Three-dimensional surface profile of cross-correlation intensity distribution. This figure shows that there is no sharp correlation peak for cross correlation.

sized based on this image. Figure 8(b) shows the absolute value of this matched filter. When the unknown input target $t(x, y)$ is the same triangle image as shown in Fig. 8(c), Fig. 8(d) shows the corresponding autocorrelation intensity distribution on the output plane. Figure 8(e) depicts the corresponding three-dimensional surface profile of the autocorrelation intensity distribution. From this figure, one can see that, indeed, there is a sharp correlation peak in the correlation plane. However, if the unknown input target $t(x, y)$ is not the same image used for the matched-filter construction, the correlation result is totally different. As an example, Figs. 9(a) and 9(b) show the same stored image and matched filter. Figure 9(c) shows a circular image used as the unknown input target. Figures 9(d) and 8(e) illustrate the cross-correlation intensity distribution and corresponding three-dimensional surface profile. In this case, there is no sharp correlation peak. Therefore, from the correlation peak intensity, one can recognize the input targets. In other words, one can tell whether the unknown input target is the stored image or not. Before the end of this section, we would like to point out that, besides the 2-D matched filter, in recent years, 3-D (spatial-spectral) matched filters were also developed. Due to space limitations, we can not provide a detail description about this work. Interested readers are directed to papers such as the one written by Yu et al. (19).

CONCLUSION

In this article, we have briefly introduced some basic concepts of the matched filter. We started our discussion with the continuous-time matched filter. Then we extended our discussion to the discrete-time input signals. After that, some major applications of the matched filters such as the signal detection and pattern recognition were addressed.

BIBLIOGRAPHY

1. D. O. North, Analysis of the factors which determines signal/noise discrimination in radar, *Technical Report No. PtR-6C,* Princeton, NJ: RCA Laboratories, 1943.
2. J. H. Van Vleck and D. Middleton, A theoretical comparison of visual, aural and meter reception of pulsed signals in the presence of noise. *J. Appl. Phys.,* **17**: 940–971, 1946.
3. B. M. Dwork, Detection of a pulse superimposed on fluctuation noise, *Proc. IRE,* **38**: 771–774, 1950.
4. T. S. George, Fluctuations of ground clutter return in air-borne radar equipment, *J. IEE,* **99**: 92–99, 1952.
5. L. Z. Zadeh and J. R. Ragazzini, Optimum filters for the detection of signals in noise, *Proc. IRE,* **40**: 1123–1131, 1952.
6. G. L. Turin, An introduction to matched filters, *IRE Trans. Inf. Theory,* **IT-6**: 311–329, 1960.
7. C. R. Cahn, Performance of digital matched filter correlation with unknown interference, *IEEE Trans. Commun. Technol.* **COM-19**: Part II, 1163–1172, 1971.
8. D. J. Gooding, *A digital matched filter for signals of large time-bandwidth product. Technical Report No. 16,* Waltham, MA: Sylvania Communications System Laboratory, February 1969.
9. K. Y. Chang and A. D. Moore, Modified digital correlator and its estimation errors, *IEEE Trans. Inf. Theory,* **IT-16**: 699–700, 1970.
10. G. L. Turin, An introduction to digital matched filters, *Proc. IEEE,* **64**: 1092–1112, 1976.
11. J. B. Thomas, *An Introduction To Communication Theory and System,* New York: Springer-Verlag, 1988, p. 202.
12. J. H. Karl, *An Introduction to Digital Signal Processing,* New York: Academic Press, 1989, p. 217.
13. L. W. Couch, *Digital and Analog Communication Systems,* New York: Macmillan, 1990, p. 497.
14. L. B. Jackson, *Digital Filters and Signal Processing,* Boston: Kluwer Academic Publishers, 1989, p. 34.
15. A. Vander Lugt, Signal detection by complex spatial filtering, *IEEE Trans. Inf. Theory,* **IT-10**, 139–145, 1964.
16. *SPIE Milestone Series on Coherent Optical Processing,* edited by F. T. S. Yu and S. Yin (eds.), Bellingham, WA: SPIE Optical Engineering Press, 1992.
17. S. Yin, et al., Design of a bipolar composite filter using simulated annealing algorithm, *Opt. Lett.* **20**: 1409–1411, 1995.
18. F. T. S. Yu, *Optical Information Processing,* New York: Wiley-Interscience, 1983, p. 10.
19. X. Yu, I. Reed, and A. Stocker, Comparative performance analysis of adaptive multispectral detectors, *IEEE Trans. Signal Process.,* **41**, 2639, 1993.

SHIZHUO YIN
FRANCIS T. S. YU
University Park, PA

MATCHING. See BROADBAND NETWORKS.

MATERIALS. See FUNCTIONAL AND SMART MATERIALS.

MATERIALS, CONDUCTIVE. See CONDUCTING MATERIALS.

MATERIALS EVALUATION. See EDDY CURRENT NONDESTRUCTIVE EVALUATION.

MATERIALS, FUNCTIONAL. See FUNCTIONAL AND SMART MATERIALS.

MATHEMATICAL LINGUISTICS. See COMPUTATIONAL LINGUISTICS.

MATHEMATICAL OPTIMIZATION. See MATHEMATICAL PROGRAMMING.