

Filter bank transceiver design for ultra wideband

Christian Ibars, Mònica Navarro, Carles Fernández-Prades,
Xavier Artiga, Ana Moragrega and Ciprian George-Gavrincea
Centre Tecnològic de Telecomunicacions de Catalunya - CTTC
Spain

Antonio Mollfulleda
Gigle Networks
Spain

Montse Nájjar
Universitat Politècnica de Catalunya
Spain

1. Introduction

Ultra Wideband (UWB) communications are steadily gaining acceptance worldwide, as the regulatory bodies of several countries define its frequency bands and rules of operation. Out of the many technologies available to generate an UWB signal, two of them have been favored, in recent standardization activities, namely Orthogonal Frequency Division Multiplexing (OFDM) and Impulse Radio (IR). On one hand the IEEE 802.15.3a standard defines an UWB signal based on OFDM, which consists of a large number of digitally-generated narrow sub-carriers. On the other hand, the IEEE 802.15.4a standard defines an UWB signal based on pulses of short duration, technology known as IR. In addition to these technologies included in standards, other modulation schemes have been proposed for UWB, such as frequency-modulated UWB, or IR transmitting even shorter, sub-nanosecond pulses. The latter has the advantage of providing very accurate ranging and localization estimates based on the time of arrival, which can be measured with high precision thanks to the large signal bandwidth.

In this chapter we will address the challenges encountered in the transmission and detection of an Impulse Radio Ultra Wideband (IR-UWB) signal in environments with multipath propagation. The main problem in such systems is that multipath produces time dispersion resulting in a very large number of resolvable paths at the receiver. We will argue that it is considerably difficult to design receivers that have both high efficiency in capturing the received signal energy and low complexity. The traditional approaches, such as correlating the received signal with a local template, or the RAKE Receiver, suffer from poor performance since they have to lock to a single propagation path, or a small subset of them, respectively. Non-coherent receivers also fail to deliver the full potential of UWB; in particular, they do not provide sufficient ranging accuracy.

Motivated by this challenge, in this chapter we develop a receiver architecture based on a frequency domain filter bank. The receiver consists mainly of a stage of bandpass filters fol-

lowed by intermediate speed analog-to-digital converters (ADCs) that sample each subband at Nyquist rate. This architecture presents two important advantages: it allows a flexible design where the number of branches and the ADC speed can be selected depending on the desired performance; and it can potentially capture all the received signal energy regardless of the severity of multipath propagation. We analyze the Filter Bank Receiver by first developing the underlying theory of time-frequency signal representations, and then deriving its performance in terms of mean square error (MSE) and bit error rate (BER), comparing it to other receiver types.

Another important performance aspect of UWB receivers is their capability to provide accurate time of arrival estimates, and thus accurate ranging and localization. We address this aspect by first analyzing synchronization schemes for UWB and then presenting a time of arrival algorithm in the frequency domain, suitable for the filter bank receiver architecture. Subsequently, we describe algorithms for ranging and localization based on time of arrival.

The last part of this chapter is devoted to describing a real implementation of the filter bank receiver. The architecture consists of an UWB RF front end followed by a digital back end which processes the received samples and implements the signal detection and time of arrival estimation algorithms. Wideband RF design is tackled using wideband low noise amplifiers (LNAs) and downconversion stages to provide lowpass signals to the ADC stage. Gaussian bandpass filters are used, which have very good time and frequency concentration properties. The digital section is implemented with a hybrid architecture consisting of a programmable logic device (PLD) to handle the high aggregate sampling rate, and a digital signal processor (DSP) to implement signal processing algorithms in a flexible and easily re-programmable fashion.

2. Ultra Wideband Signals and Channel

In impulse radio technology, the transmitted signal consists of a series of low energy, wide bandwidth pulses $p(t)$, with a duration in the order of hundreds of picoseconds. Transmitted pulses are typically shaped so that their power spectrum $|P(f)|^2$ satisfies the power spectral density (PSD) constraints set by regulatory bodies (*Generic Harmonized European Standard for UWB Communications (ETSI EN 302 065)*, 2008; *Revision of part 15 of the Commission's Rules Regarding Ultra-Wideband Transmission Systems*, 2002). Several pulses are sent for each transmission symbol in order to achieve the necessary energy per bit. At the transmitter, bits are grouped in data symbols and transmitted at symbol rate $R_s = 1/T_s$ where T_s is the symbol period. Each symbol interval is divided in N_f equally sized intervals of length T_f , known as frame intervals. In turn, each frame interval is divided in N_c equally spaced time intervals of length T_c known as chip intervals. In each frame interval only one pulse is transmitted, which is located in one of the N_c chip intervals. In case that Pulse Position Modulation of order M (M-PPM) is used, each chip interval is further divided in M modulation intervals of length T_Δ ¹. Typically, time hopping (TH) or direct sequence (DS) spreading is used to determine the pulse position and polarity, respectively, in each frame interval, with the purpose of satisfying the PSD mask. A general expression for a waveform carrying a block of N_b consecutive

¹ Although we define here the division in modulation intervals within the chip interval, it may take place at higher levels. For example, the symbol period may be divided into M modulation intervals and subsequently divided in frame and chip intervals, as in the IEEE 802.15.4a Standard.

symbols can be written as

$$s(t) = \sum_{l=0}^{N_b} \sum_{h=0}^{N_l-1} c_{h,l}^{DS} a_l p(t - lT_s - b_l T_\Delta - c_{h,l}^{TH} T_c - hT_f) \tag{1}$$

where $c_{h,l}^{DS} \in \{-1, 1\}$ is the DS code chip amplitude for the h -th frame of the l -th symbol, $c_{h,l}^{TH} \in \{0, \dots, N_c - 1\}$ selects the chip interval of the h -th frame of the l -th symbol, a_l is the amplitude of the l -th symbol and b_l is the time position of the l -th symbol. An example of the resulting transmitted signal is shown in Fig. 1.

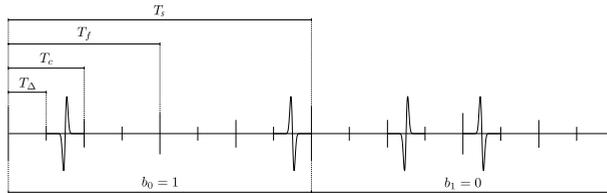


Fig. 1. Example of IR-UWB signal with TH and DS codes. $N_b = 2, N_s = 2, N_c = 2$, BPSK. Bit sequence $\{b_0 = 0, b_1 = 1\}$. TH Code sequence $\{0, 1, 1, 0\}$. DS Code sequence $\{1, -1, 1, -1\}$.

In order to fully understand the design principles of IR-UWB receivers, it is important to understand the channel propagation effects on the transmitted signal. In order to develop a common channel model, the IEEE 802.15.4a Standards Group considered several possibilities and established a modification of the traditional Saleh-Valenzuela model (Saleh & Valenzuela, 1987) as a statistical channel model for UWB. According to this model, signal paths are grouped into clusters, containing several rays with different gains ($\beta_{q,r}$) and propagation delay ($\tau_{q,r}$) where q and r are cluster index and path index inside each cluster respectively. The resulting baseband channel model is given by

$$h(t) = \sum_q \sum_r \beta_{q,r} \delta(t - \tau_r) \tag{2}$$

The UWB channel has been characterized in IEEE 802.15.4a (Molisch et al., 2004), where several types of channels are described. Among other environments, channel types 3 and 4 model an indoor environment in line of sight (LOS) and non-line of sight (NLOS) configurations respectively. The main challenge in UWB propagation stems from the fact that the receiver is very resolvable in time. Since the transmitted pulse is very short, paths separated by one nanosecond or less may be resolvable by the receiver, resulting in hundreds of paths for typical propagation environments. According to values provided in (Molisch et al., 2004), Table 1 shows the number of paths containing the 85% of the energy and the r.m.s. delay spread of channel models 3 and 4. The first parameter ranges from 22 to 45, while the latter can be up to 13 ns. Therefore, channel dispersion is very challenging for the receiver, both in terms of capturing the received signal energy and in terms of channel estimation, as it is argued in the following sections.

For the sake of simplicity and without loss of generality we will consider a simplified channel model that avoids the differentiation of paths in clusters. In this model each ray is represented by its amplitude (β_r) and delay τ_r , where r is the path index. Then, the baseband channel model can be expressed as

$$h(t) = \sum_r \beta_r \delta(t - \tau_r) \tag{3}$$

Channel Model	Environment	r.m.s. delay spread	Avg. no. paths for > 85% energy	Avg. no. paths over -10 dB
3	Indoor (office) LOS	10 ns	22.4	14.4
4	Indoor (office) NLOS	13 ns	45.5	30.4
5	Outdoor LOS	29 ns	35.8	17.3
6	Outdoor NLOS	74 ns	65.1	24.5

Table 1. Time Dispersion of IEEE 802.15.4a Channel Models.

where index r includes information of the number of cluster and the number of ray in each cluster. According to this channel model and assuming the presence of additive white Gaussian noise (AWGN) $\eta(t)$ with variance σ_η^2 the received signal is given by

$$r(t) = \sum_r \beta_r s(t - \tau_r) + \eta(t) \quad (4)$$

In general the maximum excess delay of the channel, $\tau_{max} = \max\{\tau_r\}$, can be longer than the modulation interval T_Δ , and even longer than the symbol interval T_s , causing inter-symbol interference (ISI). However, to evaluate the capabilities of capturing the received signal energy on different receiver approaches it is assumed that $\tau_{max} < T_\Delta$. Note that with this assumption all bit energy lies in one of the available modulation intervals. In this case, defining

$$h_p(t) = p(t) \otimes h(t) \quad (5)$$

the received signal can be modeled as

$$r(t) = \sum_l \sqrt{E} h_p(t - lT_s - b_l T_\Delta - c_{h,l}^{TH} T_c - hT_f) + \eta(t) \quad (6)$$

3. Filter Bank Receiver Architecture

As it can be inferred from the previous section, designing a good receiver for UWB in dispersive channels is a challenging task given the large number of resolvable propagation paths. We begin this section with a brief overview of UWB receivers. As we will see, these receivers, which provide very good performance in traditional narrowband transmission, present several shortcomings when dealing with UWB signals in dispersive channels. This motivates the filter bank architecture, which consists in splitting the UWB signal in several subbands and then sampling them at Nyquist rate. This architecture aims at maximizing the received signal energy while keeping complexity at bay. Moreover, we will see that the presented receivers may be obtained through a generalized filter bank interpretation.

3.1 Overview of Ultra Wideband Receivers

A first approach, referred to as *stored reference (SR) Receiver*, consists in correlating the received signal with a locally generated template $s_{tmp}(t)$ and then sampling the output at rate $F = \frac{1}{T}$, where T is the integration time of the correlation. The stored reference receiver is shown in Fig. 2. The signal template $s_{tmp}(t)$ is generated with the same structure used at the transmitter, that is,

$$s_{tmp}^{sr}(t) = \sum_l \sum_h \sum_{k=0}^{N_{sr}} c_{h,l}^{DS} p(t - lT_s - c_{h,l}^{TH} T_c - hT_f - kT - \tau_s) \quad (7)$$

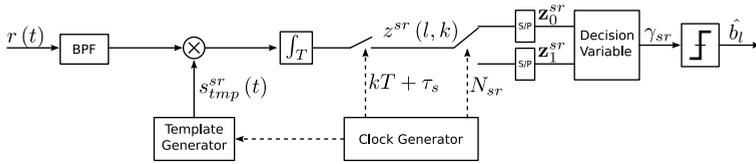


Fig. 2. Stored Reference Receiver.

where $p(t)$ is the transmitted pulse with unit energy, l represents the symbol index and N_{sr} is the number of samples taken in each modulation interval, which is given by $N_{sr} = \frac{T}{T_\Delta}$. The term τ_s represents the sampling offset, which is intentionally modeled to represent the ability of the receiver to modify τ_s to maximize the captured energy, particularly on multipath channels. After the analog correlator, the signal is sampled by an ADC at rate N_{sr}/T , and fed to a bank of correlators operating on decision variables \mathbf{z}_0^{sr} and \mathbf{z}_1^{sr} , containing all samples taken at each respective modulation interval. Finally, a decision variable is obtained as

$$\gamma_l = \mathbf{w}^T (\mathbf{z}_1^{sr} - \mathbf{z}_0^{sr}) \tag{8}$$

where the weights \mathbf{w} correspond to a whitened matched filter for reception with Gaussian noise (Kay, 1998).

Another common approach to receive spread spectrum signals in a multipath propagation environment is the RAKE Receiver, shown in Fig. 3. The RAKE receiver uses multiple correlators that lock at different multipath replicas (Zhu et al., 2008). The output of the correlators is sampled and combined before symbol detection. RAKE receivers require accurate estimation of the delay, amplitude, phase and shape (distortion) of the pulses at each individual arrival. When the number of fingers N_{rk} is equal to the number of resolvable paths, this receiver constitutes in fact the matched filter receiver. In practice, given the large number of resolvable UWB paths, the receiver will capture only a fraction of the received signal energy, since increasing N_{rk} has a relevant impact on its complexity.

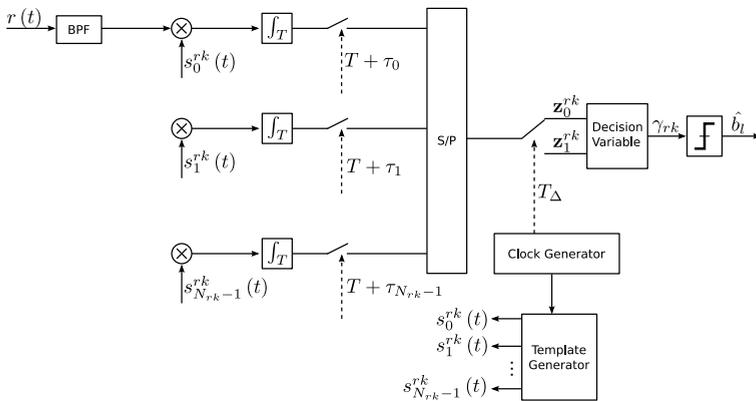


Fig. 3. RAKE Receiver.

Since the received signal is correlated by locally generated replicas of the effective channel impulse response, the analysis of the RAKE receiver is similar to that of the SR receiver. The

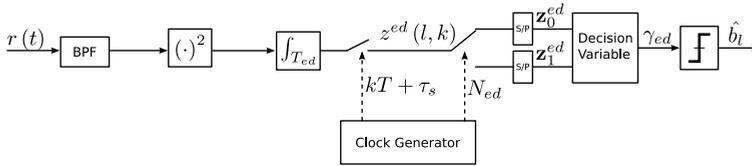


Fig. 4. Energy Detector Receiver.

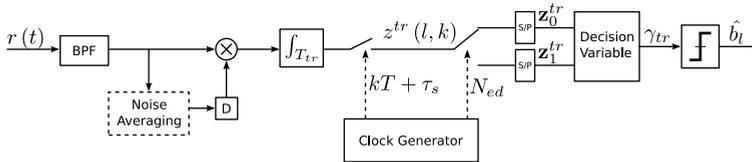


Fig. 5. Transmitted Reference Receiver.

signal template of the k -th finger is constructed in the same way as the transmitted signal and is given by

$$s_k^{rk}(t) = \sum_l \sum_h p(t - lT_s - c_l^{TH}T_c - hT_f - \tau_k) \tag{9}$$

where $k \in \{0, \dots, N_{rk} - 1\}$ and τ_k represents the k -th propagation path with largest energy. Note that the RAKE receiver requires N_{rk} ADC converters sampling at symbol rate while SR receiver requires just one converter and an oversampling factor of N_{sr} with respect to the RAKE approach. In both cases the number of samples taken in a modulation interval is a critical parameter.

In the following we discuss two non-coherent approaches, namely the *Energy Detector* and the *Transmitted Reference Receiver*. In contrast to RAKE and SR approaches the energy detector (ED) receiver does not correlate the received signal with a local template. Instead, the analog front-end adds up the received energy in each modulation interval to create the decision variables. Fig. 4 shows the main blocks that constitute the ED receiver.

The Transmitted Reference Receiver was originally created to avoid the local knowledge of the correlation template, which requires a priori knowledge of the carrier to interference ratio (CIR). To this end, the transmitted signal generates a reference pulse for each transmitted pulse. The two pulses are separated an interval that is longer than the maximum excess delay of the channel. The first pulse is used as a correlation template while the second pulse is used for signal detection. The receiver functional block diagram is depicted in Fig. 5. The signal is correlated by a delayed version of itself, integrated and then sampled at rate T . For each modulation interval N_{tr} samples are collected, where $N_{tr} = \frac{T_s}{T}$, and then processed.

3.2 Time-frequency Representations

The coherent receivers described in the previous sections, namely stored reference and RAKE receivers, use analog components to generate a signal template and correlate it with the received signal. This approach places serious complexity constraints on these structures, such as the number of fingers of the RAKE receiver. An interesting alternative is to perform all these operations (template generation and correlation) in the digital domain, where more complex structures can be addressed if enough computing power is available. However, implementing an all-digital receiver (see (Blazquez et al., 2003) for example) requires sampling the received

signal at Nyquist rate, which can be very challenging for signals with up to 10 GHz bandwidth. An alternative approach is considered in this section, where sampling is performed in a time-frequency grid rather than in the time domain only.

Representing a signal by a discrete set of samples is a fundamental problem in signal processing. Signal projection onto a set of base functions and series expansion of signals provides an elegant way to represent a broad range of possibilities for signal acquisition. An important step in a series expansion is the choice of an appropriate basis to represent the signal of interest. That is, given a signal $s(t)$ choosing a set of functions $\gamma_k(t)$ and $g_k(t)$ such that the basis coefficients are obtained as

$$s_k = \langle s(t), \gamma_k(t) \rangle = \int_{-\infty}^{+\infty} s(\tau) \gamma_k(\tau) d\tau \quad (10)$$

while the original signal is reconstructed as

$$s(t) = \sum_k s_k g_k^*(t) \quad (11)$$

where (10) is known as the analysis equation while the linear expansion in (11) is known as the synthesis equation. If $g_k(t) = \gamma_k(t)$, then (11) is known as the orthogonal series expansion of $s(t)$. Otherwise, the functions $\gamma_k(t)$ and $g_k(t)$ are a set of biorthogonal functions with the property

$$\langle g_j(t), \gamma_i(t) \rangle = \delta_{i-j}(t)$$

where $\delta_{i-j}(t)$ is defined such that $\delta_{i-j}(t) = 0$ unless $i = j$, in which case $\delta_0(t) = \delta(t)$. The idea to take signal samples from its representation in the time-frequency plane was suggested by (Gabor, 1946). Instead of processing the entire signal at once, Gabor proposed to divide the signal into equally sized segments and then perform the Fourier transform of each segment. The result provides local information about the frequency content on each time interval. To perform the two-dimensional sampling on $s(t)$ the time axis is divided into N equally sized intervals of length T , which represent the time-domain sampling period. Each segment of $s(t)$ is labeled as $s_n(t)$ and is given by

$$s_n(t) = s(t) \Pi\left(\frac{t-nT}{T}\right) \quad (12)$$

where n indexes the number of segments and $\Pi(t)$ represents a square pulse of unit length and amplitude. Since $s_n(t)$ is a time-limited signal, its Fourier transform $S_n(f)$ can be expressed as the sequence of discrete samples

$$S_n(f) = \sum_k S(kB) \text{Sinc}\left(\frac{f}{B} - k\right) \quad (13)$$

where $B = \frac{1}{T}$ is the sampling interval in the frequency domain. Fig. 6 shows an example of the two-dimensional sampling grid. The product BT defines the density of samples taken from the time-frequency energy distribution of the signal. Nyquist density, defined as the time-frequency product $BT = 1$, provides the minimum set of samples required for perfect signal reconstruction. Lower values of BT imply oversampling the signal, while larger values correspond to subsampling, in which case not all the signal energy is captured.

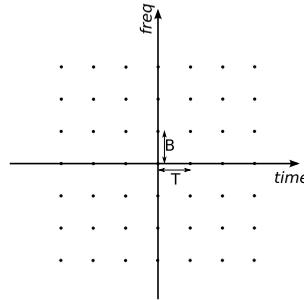


Fig. 6. Example of 2-dimensional sampling grid.

Let us denote by $s_{n,m}$ the two-dimensional samples of $s(t)$. Indexes n and m represent the time and frequency domain sampling, respectively. Denoting $\hat{s}_n(t)$ as the periodic signal obtained by the sampling of $S_n(f)$ it follows that

$$s_{n,m} = S_n(mB) = \int_{-T}^T \hat{s}_n(t) e^{-j2\pi mBt} dt \tag{14}$$

$$= \int_{-\infty}^{\infty} s(t) \Pi\left(\frac{t-nT}{T}\right) e^{-j2\pi mBt} dt \tag{15}$$

Therefore, the two-dimensional samples can be expressed as a function of the time-domain signal $s(t)$ as follows

$$s_{n,m} = \int_{-\infty}^{\infty} s(t) \gamma_{n,m}(t) dt \tag{16}$$

where

$$\gamma_{n,m}(t) = \Pi\left(\frac{t-nT}{T}\right) e^{-j2\pi mBt} \tag{17}$$

Using the synthesis equation of the series expansion, the reconstruction of the signal $s(t)$ is given by

$$\hat{s}(t) = \sum_n \sum_m s_{n,m} \gamma_{n,m}^*(t)$$

Assuming that an uncounted number of frequency domain samples are taken on each interval and $BT = 1$, one would provide a perfect signal reconstruction. Limiting the number of samples on each interval to a finite number M leads to the M -th order representation of $s(t)$, which is defined as the truncated series expansion,

$$\tilde{s}(t) = \sum_n \sum_{m=0}^{M-1} s_{n,m} \gamma_{n,m}^*(t)$$

A figure of merit of the quality of signal acquisition can be given in terms of the MSE of the truncated series representation, which is given by

$$\begin{aligned} MSE_{2D} &= \int |s(t) - \sum_n \sum_{m=0}^{M-1} s_{n,m} \gamma_{n,m}^*(t)|^2 dt \\ &= \sum_n \sum_{m=M}^{\infty} |s_{n,m}|^2 \end{aligned}$$

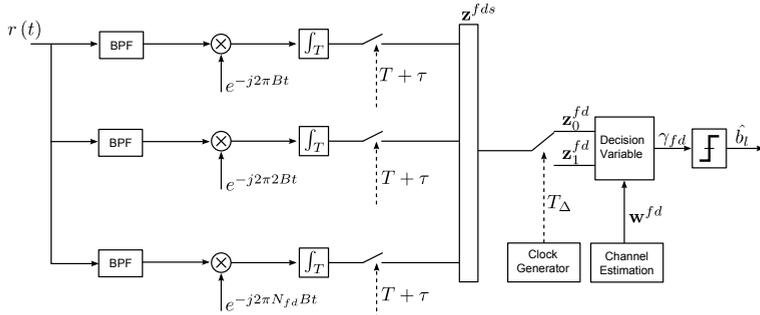


Fig. 7. Block diagram of a Filter Bank Receiver.

Therefore, having the a priori knowledge of the signal bandwidth one can define M to achieve a given MSE on the signal reconstruction.

3.3 Filter Bank Receiver

The fundamental idea behind the filter bank receiver is to provide an implementation that allows us to obtain a highly accurate time-frequency representation of the received UWB signal. Consider an IR signal as described in Section 2, in which 2-Pulse Position Modulation (PPM) modulation is used. The signal length conveying a block of N_p bits is $T_p = N_p T_s$. The channel model is the same one used in (3) where the maximum excess delay of the channel is considered lower than the modulation interval. The noise is assumed to be AWGN with variance σ_{η}^2 . Fig. 7 shows the block diagram of the filter bank receiver. The receiver implements a 2-dimensional sampling stage by splitting the input signal into M paths and correlating it with the basis function

$$\gamma_{n,m}(t) = \Pi\left(\frac{t - nT}{T}\right) e^{-j2\pi mB(t-nT)} \tag{18}$$

In practice, this operation may be carried out with orthogonal, or quasi-orthogonal, bandpass filters. The output of each correlator is sampled at a rate $\frac{1}{T}$, and all samples associated to each modulation interval are arranged in a vector. The captured samples can be expressed as

$$z_{n,m} = \langle r(t), \gamma_{n,m}(t) \rangle \tag{19}$$

where $m \in (1, \dots, M)$ indexes the samples in frequency domain. The collected samples for each symbol form a two-dimensional set of samples. Using (6) and (19) it follows that, when the signal is present on a given interval,

$$z_{n,m} = \langle h_p(t), \gamma_{n,m}(t) \rangle + \langle \eta(t), \gamma_{n,m}(t) \rangle \tag{20}$$

Let $\mathbf{z}_0^{fb}(l)$ be the vector gathering all samples in the modulation interval associated to $b_l = 0$ of the l -th symbol, and $\mathbf{z}_1^{fb}(l)$ the vector for the time interval corresponding to $b_l = 1$. The statistics of $\mathbf{z}_0^{fb}(l)$ and $\mathbf{z}_1^{fb}(l)$ are given by

$$\mathbf{z}_0^{fb}(l) = \begin{cases} \mathcal{N}\left(\mathbf{z}_p^{fb}(k), \mathbf{C}_{\eta}^2\right), & b_l = 0 \\ \mathcal{N}\left(\mathbf{0}, \mathbf{C}_{\eta}^2\right), & b_l = 1 \end{cases} \tag{21}$$

$$\mathbf{z}_1^{fb}(l) = \begin{cases} \mathcal{N}(\mathbf{0}, \mathbf{C}_\eta^2), & b_l = 0 \\ \mathcal{N}(\mathbf{z}_p^{fb}(k), \mathbf{C}_\eta^2), & b_l = 1 \end{cases} \tag{22}$$

where $\mathbf{z}_p^{fb} = [\{ \langle h_p(t), \gamma_{n,m}(t) \rangle \}_{n,m}]^T$ and \mathbf{C}_η is the noise covariance matrix. These vectors are combined to make the decision variable. Defining the combining weights as $\mathbf{w} = [w_0, w_1, \dots, w_{MN_{fb}-1}]^T$, the decision variable is given by

$$\gamma_l = 2\Re\{\mathbf{w}^H (\mathbf{z}_1^{fb} - \mathbf{z}_0^{fb})\} \tag{23}$$

The weights that lead to optimal solution for Gaussian noise correspond to the generalized matched filter solution (Kay, 1998), which defines the weights as $\mathbf{w} = [\mathbf{C}_{fb}^{-1} \mathbf{z}_p^{fb}]$. The statistics of the decision variable are given by

$$\gamma_l^{fb} = \begin{cases} \mathcal{N}(\mathbf{z}_p^{fbH} \mathbf{C}_{fb}^{-1} \mathbf{z}_p^{fb}, \mathbf{z}_p^{fbH} \mathbf{C}_{fb}^{-1} \mathbf{z}_p^{fb}), & b_l = 1 \\ \mathcal{N}(-\mathbf{z}_p^{fbH} \mathbf{C}_{fb}^{-1} \mathbf{z}_p^{fb}, \mathbf{z}_p^{fbH} \mathbf{C}_{fb}^{-1} \mathbf{z}_p^{fb}), & b_l = 0 \end{cases} \tag{24}$$

Using (24) it follows that the BER for the filter bank receiver is given by

$$BER_{fb} = Q\left(\sqrt{\mathbf{z}_p^{fbH} \mathbf{C}_{fb}^{-1} \mathbf{z}_p^{fb}}\right) \tag{25}$$

Provided that the set of functions $\gamma_{n,m}(t)$ constitute an orthogonal basis, and assuming that the noise vector is AWGN, the noise covariance matrix becomes $\mathbf{C}^{fb} = I\sigma_\eta^2$ and the BER is reduced to

$$BER_{fb} = Q\left(\sqrt{\frac{\mathbf{z}_p^{fbH} \mathbf{z}_p^{fb}}{2\sigma_\eta^2}}\right) \tag{26}$$

Note that the product $\mathbf{z}_p^{fbT} \mathbf{z}_p^{fb}$ can be expressed, in terms of the MSE of the M -th order representation of $s(t)$, as

$$\mathbf{z}_p^{fbT} \mathbf{z}_p^{fb} = \sum_{n=0}^N \sum_{k=1}^M \|H_n(kB)\|^2 = \tag{27}$$

$$= \sum_{n=0}^N \left(1 - \sum_{k=M+1}^\infty \|S_n(kB)\|^2\right) = \tag{28}$$

$$= \sum_{n=0}^N (1 - MSE_n) = 1 - MSE \tag{29}$$

where $H_n(f)$ is the Fourier transform of $h_p(t)\Pi\left(\frac{t-nT}{T}\right)$. Therefore, the BER of the filter bank receiver may also be expressed in terms of the reconstruction MSE.

At this point it is interesting to look back at the SR and RAKE receivers as implementations of a more general filter bank concept. Both receivers are based on sampling the received signal after correlating with a locally generated template. While the SR Receiver implements

a single correlator, the RAKE Receiver uses a multiple correlators in parallel. In both cases, the correlation can be expressed as a signal projection over a particular basis. Therefore, SR and RAKE receivers can be expressed as a generalized class of receiver based on signal projection over a given basis. The complexity of each receiver is determined by the following factors:

- Size of the basis, which determines the number of sampling units.
- Sampling frequency at each ADC.
- Sampling offset of each ADC.
- Requirement that sampling clock is synchronized with received signal.

This interpretation is useful to compare SR and RAKE receivers to the frequency domain filter bank receiver. In Table 2 we compare these receivers in terms of requirements and complexity. In the table, N stands for the number of samples captured in each modulation interval.

Receiver Type	Basis function	N	Number of sampling units	Synchronized clock
SR	$p(t)$	≥ 1	1	No
RAKE-M	$\beta_n p(t - \tau_n)$	1	$M_{rk} > 1$	Yes
Filter bank	$\Pi\left(\frac{t-nT}{T}\right) e^{-j2\pi m B t}$	$\geq M/B$	$\geq B/M$	No

Table 2. Comparison of UWB Receivers as Generalized Filter Banks.

3.4 Performance of the Filter Bank Receiver

We now compare the performance of the filter bank receiver with the other receivers described. In terms of performance, one may look at the reconstruction MSE or the BER, which are related by (29). Regarding the BER, the IEEE channel model type 3 is used, bounding the maximum excess delay to $\max\{\tau_r \leq 50ns\}$. The receivers are assumed to be accurately synchronized to the leading edge of the received signal. The receiver settings used in the simulations are summarized in Table 3, which also reports the mean square error for each receiver. It can be seen that a filter bank receiver achieves the lowest MSE if 4 or more filters are used, even compared to a RAKE receiver with 16 fingers. The average BER of the receivers in Table 3 is shown in Fig. 8. As it can be seen, the BER of the filter bank receiver approaches the matched filter bound as the number of filters increases, and is lower than for the other receivers for $M \geq 4$. In the simulations, an ideal basis function with square pulses was used. Such basis may not be used in practice, since it requires each filter to have unlimited bandwidth. The use of different basis functions may deteriorate the performance of the filter bank receiver. Moreover, if the basis functions do not form an orthogonal basis, further receiver processing is necessary. Results in (Mollfulleda et al., 2010) show, however, that the performance of a filter bank receiver using implementable Gaussian filters is still superior to other receiver types.

4. Time of Arrival Estimation and Localization

The wide bandwidth used by UWB radio systems provides two fundamental advantages for localization applications, especially in short-range wireless networks. On the one hand, diversity in frequency components increases the probability that some portion of the signal power can penetrate and go through obstacles. This allows applications such as through-the-wall

Receiver Type	Sampling Rate	No. of ADC	samples on each Modulated interval	MSE (dB)
SR-2G	2 GHz	1	$N_{sr} = 50$	-1.67
SR-4G	4 GHz	1	$N_{sr} = 100$	-2.52
RK-8	20 MHz	8	$N_{rk} = 8$	-3.92
RK-16	20 MHz	16	$N_{rk} = 16$	-5.92
Filter Bank M=2	2 GHz	4	$N_{fb} = 400$	-2.3
Filter Bank M=4	2 GHz	8	$N_{fb} = 800$	-10.59
Filter Bank M=6	2 GHz	12	$N_{fb} = 1200$	-13.8

Table 3. Receiver Settings for BER Simulation.

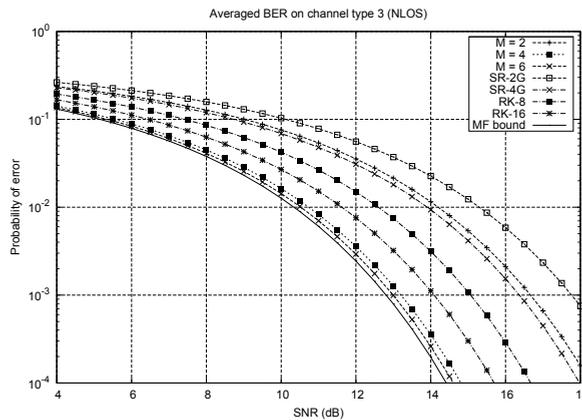


Fig. 8. Averaged BER curves in channel type 3 for the Filter Bank Receiver, SR receiver, and RAKE Receiver. The Matched filter bound is also plot for comparison.

radars, or vision through an opaque medium, where challenges are related to counteracting the distortion due to the dispersive properties of different materials. On the other hand, a large relative bandwidth improves ranging accuracy: it is well known that the Cramér-Rao Bound (the theoretical minimum variance achievable by an unbiased estimator) of time–delay estimator is inversely proportional to the RMS bandwidth, also known as Gabor bandwidth, defined as:

$$B_{RMS} = \sqrt{\frac{\int_{-\infty}^{\infty} f^2 |P(f)|^2 df}{\int_{-\infty}^{\infty} |P(f)|^2 df}} \tag{30}$$

where $P(f)$ is the Fourier transform of the pulse waveform. Inspecting equation (30), it is easily seen that we can increase the RMS bandwidth using modulations whose power spectrum concentrates a greater percentage of their power farther from the signal center frequency, because of the weighting term f^2 . In that sense, IR–UWB systems exhibit a very large RMS bandwidth due to the short duration of the pulses (less than one nanosecond), which have strong high–frequency components. This implies that the fundamental lower limit of the variance of an IR–UWB timing estimator is dramatically small, and hence this explains the potentiality of

this technology for accurate localization or ranging purposes. However, current technology can hardly cope with sampling rates at or above the Nyquist rate, and thus the Cramér–Rao Bound is difficult to attain, notably in dense–multipath indoor environments.

Indeed, one of the most attractive features of UWB systems is their intrinsic potential to allow for very accurate positioning. Due to their large signal bandwidth, UWB signals exhibit very high time resolution. This high time resolution allows receivers to resolve individual components in dense multipath propagation scenarios, which has strong implications to obtain accurate time of arrival (TOA) measurements. In order to measure the TOA between two nodes the receiver needs to identify the first arriving path, which is associated to LOS propagation. In a multipath scenario, the direct path may be masked in a multipath cluster, unless the receiver is able to resolve individual paths within that cluster. Therefore, time–based positioning is the most widely selected technique when implementing localization in UWB systems. This section addresses the estimation of timing measurements and their application to ranging and positioning. We start addressing the more general problem of synchronization for very broadband signal systems and continue with a review of the state–of–the–art on TOA estimation algorithms, which can be broadly classified in two main categories: energy detection and correlator–based.

Special attention is placed to the the filter bank receiver in order to derive suitable algorithms for TOA estimation of the LOS signal, emphasizing some of the main drawbacks linked to the receiver architecture, and arguing for a frequency domain processing for the estimation of the TOA, particularly suitable for the filter–bank receiver, as shown in (Navarro & Nájjar, 2007). The filter–bank architecture of the receiver falls very naturally into frequency domain processing since the filters output represents the DFT components of the received UWB signal. Also, well–known high–resolution spectral estimation methods can be applied directly to the frequency domain signal samples achieving very accurate timing estimation. Finally, the section addresses positioning by considering a particular approach based on data fusion algorithms. This technique is particularly attractive for developing global positioning systems that solve the outdoor/indoor transition problem efficiently.

4.1 Synchronization for Wideband Signals

Synchronization is an essential part in every communication system, but it is critical in UWB communication systems, given that small timing errors of the order of the pulse duration can considerably degrade system performance. In general, synchronization involves recovering information at several levels: carrier recovery, symbol timing recovery, and frame synchronization. However, in the case of carrier–less impulse radio UWB, based on PPM modulation, synchronization is reduced to timing synchronization.

The synchronization technique to be developed depends on the receiver architecture and capabilities. A broad classification falls into synchronization for coherent and non–coherent receivers. The coherent reception of IR–UWB signal indicates the recovery of polarity of UWB pulse as well as the position information, while the non–coherent reception can only recover the position information by collecting the energy of the pulses. Most of the research work on IR–UWB implementation focuses on a non–coherent receiver because it has the advantage of avoiding the use of a bank of correlators and pulse matched filter, which leads to simpler implementations.

Nevertheless, we shall review the coherent techniques suitable for PPM modulated IR–UWB signals to provide a benchmark for lower complexity techniques. The basis can be found on

the matched filter solutions, see (Oh & Kim, 2008; Oh et al., 2009; Wu et al., 2008), and (Kim et al., 2009). As introduced earlier, the advantages of coherent reception include:

- First, the coherent reception is important for frame acquisition and ranging during the preamble symbol interval. Timing or frame acquisition in a coherent receiver is very accurate compared with a non-coherent receiver, which in turn leads to accurate ranging.
- Second, coherent reception can achieve more channel coding gain since convolutional decoding can be performed during the header and payload intervals.

Research on coherent algorithms for UWB receivers was first conducted by (Lee & Scholtz, 2002), proposing a CLEAN (see (Högbom, 1974)) based iterative correlation algorithm; while high precision can indeed be achieved, the iterative amplitude estimating and adjusting process makes the algorithm extremely computationally burdensome. (Wu et al., 2007) optimized the process of amplitude adjusting, which greatly reduced the computation complexity without harming performance; however, in presence of severe multipath, the algorithm still has to carry out repeated correlations and amplitude estimations. Therefore, some researchers considered detecting the direct path directly from the match-filtering output of the received signal. In the work by (Chung & Ha, 2003), the peak of match-filtering output was considered as the location of the direct path, but this is only applicable for LOS situations with transceiver antennas being omnidirectional. Threshold detection was conducted in (Low et al., 2005) and (Lee & Yoo, 2006), but presenting some difficulties for a practical implementation. Traditional tracking loops were implemented based on phase-locked loops (PLLs) and baseband code tracking delay-locked loops (DLLs). However, because of the extremely low duty cycle and significantly large bandwidth in IR-UWB, it is no longer valid to use the traditional sampling and interpolation method, see (Gardner, 1986). This challenge motivates us to draw on conventional PLL, delay locked loop, and code tracking loop theories to design appropriate equivalent timing locked loop for tracking of ultra-short IR-UWB impulses.

4.1.1 Synchronization in UWB Systems with Dirty Templates

The Timing with Dirty Templates (TDT) approach applied to PPM was presented in (Yang, 2006). The derivation is as follows: first, we reformulate the integrate-and-dump operation for PPM signals:

$$x(k; t) = \int_0^{T_s} r_{2k+1}(t; \tau) \check{r}_{2k}(t; \tau) dt \quad \forall \tau \in [0, T_s), \tag{31}$$

where

$$\check{r}_k(t; \tau) = r_k(t + \Delta; \tau) - r_k(t - \Delta; \tau), \tag{32}$$

and $\Delta = b_l T_\Delta$ is the PPM modulation index in (1). To see how (31) enables TDT, let us first consider its noise-free part

$$\chi(k; \tau) = \int_0^{T_s} \rho_{2k+1}(t; \tau) \check{\rho}(t; \tau) dt, \tag{33}$$

where $\rho_k(t; \tau)$ and $\check{\rho}(t; \tau)$ represent the noise-free parts of $r_k(t; \tau)$ and $\check{r}(t; \tau)$, respectively. By definition, and since waveform $p_r(t)$ has a nonzero support upper bounded by the symbol duration T_s , we have

$$\rho_k(t; \tau) = \sqrt{E} \sum_{m=0}^1 p_r(t + mT_s - \check{\tau}_0 - s_{k-k_0-m}\Delta), \tag{34}$$

$$\check{\rho}_k(t; \tau) = \rho_k(t + \Delta; \tau) - \rho_k(t - \Delta; \tau), \quad \forall t, \tau \in [0, T_s) \tag{35}$$

where $\check{\tau}_0 = [\tau_0 - \tau] \bmod T_s$ and $k_{\tau_0} = \lfloor \frac{\tau_0 - \tau}{T_s} \rfloor \in \{0, -1\}$. Although k_{τ_0} can induce demodulation delay, it does not affect the τ_0 estimation. For notational brevity, we will henceforth omit k_{τ_0} . Using (34), Appendix I of (Yang, 2006) shows that, when the PPM modulation index satisfies $\Delta \ll T_f$, the noise-free part of $x(k; \tau)$ in (31) simplifies to

$$\chi(k; \tau) \approx (s_{2k-1} - s_{2k})E_A(\check{\tau}_0) + (s_{2k} - s_{2k+1})E_B(\check{\tau}_0) \tag{36}$$

where we have used the definitions

$$E_A(\check{\tau}_0) = E \int_{T_s - \check{\tau}_0}^{T_s} p_r^2(t) dt \tag{37}$$

$$E_B(\check{\tau}_0) = E \int_0^{T_s - \check{\tau}_0} p_r^2(t) dt \tag{38}$$

Non-data-aided TDT

Averaging with respect to the random symbols $\{s_k\}$, the mean-square of $\chi(k; \tau)$ is

$$E_s \{ \chi^2(k; \tau) \} \approx \frac{1}{2} \left(E_R^2 - 3E_A(\check{\tau}_0)E_B(\check{\tau}_0) \right) \tag{39}$$

which contains the energy product $E_A(\check{\tau}_0)E_B(\check{\tau}_0)$ and is uniquely maximized at $\check{\tau}_0 = 0$, that is, at the correct timing $\tau = \tau_0$. In (39), $E_R = E_A(\check{\tau}_0) + E_B(\check{\tau}_0) = E \int_0^{T_s} p_r^2(t) dt$ is the constant energy of the unknown aggregate template at the receiver.

When we take into account the bandpass-filtered zero-mean additive Gaussian noise $\eta(t)$, the symbol-rate samples obtained by integrating and dumping the products of adjacent “dirty templates” become

$$x(k; \tau) = (s_{2k-1} - s_{2k}) E_A(\check{\tau}_0) + (s_{2k} - s_{2k+1}) E_B(\check{\tau}_0) + \zeta(k; \tau) \tag{40}$$

where the noise term $\zeta(k; \tau)$ can be expressed as

$$\zeta(k; \tau) = \zeta_1(k; \tau) + \zeta_2(k; \tau) + \zeta_3(k; \tau) \tag{41}$$

$$\zeta_1(k; \tau) = \int_0^{T_s} \check{\rho}_{2k}(t; \tau) \eta_{2k+1}(t; \tau) dt \tag{42}$$

$$\zeta_2(k; \tau) = \int_0^{T_s} \rho_{2k+1}(t; \tau) (\eta_{2k}(t + \Delta; \tau) - \eta_{2k}(t - \Delta; \tau)) dt \tag{43}$$

$$\zeta_3(k; \tau) = \int_0^{T_s} \eta_{2k+1}(t; \tau) (\eta_{2k}(t + \Delta; \tau) - \eta_{2k}(t - \Delta; \tau)) dt \tag{44}$$

with $\eta_k(t; \tau) = \eta(t + kT_s + \tau), \forall t \in [0, T_s)$. Appendix II of (Yang, 2006) shows that the noise term $\zeta(k; \tau)$ can be well approximated as white Gaussian noise with zero mean and variance $\sigma_{\zeta}^2 \approx 2E_R N_0 + BT_s N_0^2$. Then, the mean square of the samples of $x(k; \tau)$ can be found as

$$E_{s, \zeta} \{ x^2(k; \tau) \} = E_s \{ \chi^2(k; \tau) \} + E_{xi} \{ \zeta^2(k; \tau) \} \tag{45}$$

$$\approx \frac{1}{2} \left(E_R^2 - 3E_A(\check{\tau}_0)E_B(\check{\tau}_0) + 2\sigma_{xi}^2 \right) \tag{46}$$

which is uniquely maximized when $\check{\tau} = 0$, that is, when $\tau = \tau_0$. Then, the non-data-aided TDT is:

$$\tau_0 = \arg \max_{\tau \in [0, T_s]} E_{s, \check{\zeta}} \{x^2(k; \tau)\} \quad (47)$$

Replacing the ensemble mean with its sample mean estimator, we have the following timing algorithm:

$$\hat{\tau}_0 = \arg \max_{\tau \in [0, T_s]} \frac{1}{K} \sum_{k=1}^K \left(\int_0^{T_s} r_{2k}(t; \tau) \check{r}_{2k-1}(t; \tau) dt \right)^2 \quad (48)$$

This estimator is a blind (non-data-aided) unbiased and consistent estimator in the mean-square sense, as can be seen from the mean and variance of the cost function:

$$m(K, \tau) = E \left\{ \frac{1}{K} \sum_{k=1}^K \left(\int_0^{T_s} r_{2k}(t; \tau) \check{r}_{2k-1}(t; \tau) dt \right)^2 \right\} = \quad (49)$$

$$= \frac{1}{2} \left(E_R^2 - 3E_A(\check{\tau}_0)E_B(\check{\tau}_0) + 2\sigma_{\check{\zeta}}^2 \right) \quad (50)$$

and

$$\sigma^2(K, \tau) = \text{var} \left\{ \frac{1}{K} \sum_{k=1}^K \left(\int_0^{T_s} r_{2k}(t; \tau) \check{r}_{2k-1}(t; \tau) dt \right)^2 \right\} = \quad (51)$$

$$= \frac{2\sigma_{\check{\zeta}}^2}{K} \left(E_R^2 - 3E_A E_B + \sigma_{\check{\zeta}}^2 \right) + \frac{1}{4K} \left(E_R^2 - 3E_A E_B \right)^2 \quad (52)$$

It is worth emphasizing that the basic idea behind our TDT estimator is that $E_A(\check{\tau}_0)E_B(\check{\tau}_0)$ is minimized when $\check{\tau}_0 = 0$ and thus $\tau = \tau_0$. Although terms E_R and $\sigma_{\check{\zeta}}$ are unknown because $p_r(t)$ is unknown, they remain constant $\forall \tau$, and thus do not affect the peak-picking operation in finding $\hat{\tau}_0$.

Data-aided TDT

From (40), we observe that: 1) when $\check{\tau}_0 \neq 0$, then $x(k; \tau)$ only contributes noise if $s_{2k-1} = s_{2k} = s_{2k+1}$ and 2) when $\check{\tau}_0 = 0$, then $x(k; \tau)$ only contributes noise if $s_{2k} = s_{2k+1}$. To avoid these noise-only contributions that do not contain any timing information, the training sequence $\{s_k\}$ should be designed such that no successive symbols are the same. Hence, the training sequence for data-aided TDT is designed to comprise a repeated pattern (1, 0); that is

$$s_k = \{k + 1\}_2 \quad (53)$$

It can be easily verified that this pattern simplifies (40) to

$$x(k; \tau) = E_B(\check{\tau}_0) - E_A(\check{\tau}_0) + \zeta(k; \tau) \quad (54)$$

Then, its mean-square becomes

$$E_{\check{\zeta}} \{x^2(k; \tau)\} = E_R^2 - 4E_A(\check{\tau}_0)E_B(\check{\tau}_0) + \sigma_{\check{\zeta}}^2 \quad (55)$$

Notice that the estimator (48) relies on three major steps: correlation, averaging, and squaring. The training sequence in (48) allows us to swap the order of these steps and alleviate the noise effects. Specifically, in the data-aided mode, it follows from (54) that

$$E_{\xi}^2 \{x(k; \tau)\} = E_R^2 - 4E_A(\check{\tau}_0)E_B(\check{\tau}_0) \quad (56)$$

In other words, by taking the *squared-mean* instead of *mean-square*, the noise variance term in (55) is eliminated. This observation leads us to the following result of a timing algorithm tailored for our carefully designed training sequence:

$$\hat{\tau}_0 = \arg \max_{\tau \in [0, T_s)} \left(\frac{1}{K} \sum_{k=1}^K \int_0^{T_s} r_{2k}(t; \tau) \check{r}_{2k-1}(t; \tau) dt \right)^2 \quad (57)$$

The mean and variance of this cost function can be obtained as

$$m(K; \tau) = E_R^2 - 4E_A(\check{\tau}_0)E_B(\check{\tau}_0) + \frac{\sigma_{\xi}^2}{K} \quad (58)$$

$$\sigma^2(K; \tau) = \frac{2\sigma_{\xi}^2}{K} \left(E_R^2 - 4E_A(\check{\tau}_0)E_B(\check{\tau}_0) + \frac{\sigma_{\xi}^2}{K} \right) \quad (59)$$

4.1.2 Time of Arrival Estimation

One can view the estimation of TOA as a particular case of the timing acquisition problem. Synchronization is seen as the timing information required for data demodulation while TOA is linked to identify the first arriving path, but in essence both require similar techniques. In the case of IR-UWB the maximum likelihood (ML) solution is known (Lottici & Mengali, 2003) but has strong practical limitations due to the requirement of very high sampling rates and complexity. Although different ML approaches have appeared in the literature that manage to reduce complexity considerably López-Salcedo & Vázquez (2005); Yang & Giannakis (2005) there still exist practical limitations for their use in positioning applications. Efforts have been steered towards near optimum less complex solutions, most of them based on time domain approaches.

The majority of practical solutions for TOA estimation found in the literature can be broadly classified in energy-based and correlation-based approaches. A simplified representation of the receiver block diagram is depicted in Fig.9 for an scheme based on a correlator/matched-filter and energy detection, respectively. Energy based TOA estimators received considerable attention as a viable alternative to correlation-based methods (Cheong et al., 2005; Rabbachin et al., 2005) which require the estimation of the pulse shape. Indeed, energy detectors do not require expensive pulse-shape estimation algorithms and represent a good solution for low-power and low-complexity systems. The estimation scheme proposed and analyzed in (Rabbachin et al., 2005) is representative of a large class of energy based TOA estimation algorithms, where the incoming signal is squared and integrated over time intervals comparable with the pulse width. The integration time intervals defines in fact the time resolution of the algorithm and is also strongly related to its latency. These schemes also consider additional processing to improve signal quality. For instance, in (Rabbachin et al., 2005) the energies within non-overlapping adjacent windows are collected over several symbol periods to reduce the effects of background noise. The location of the direct path is computed as the index of the first interval where the energy overcomes a suitable threshold.

Thank You for previewing this eBook

You can read the full version of this eBook in different formats:

- HTML (Free /Available to everyone)
- PDF / TXT (Available to V.I.P. members. Free Standard members can access up to 5 PDF/TXT eBooks per month each month)
- Epub & Mobipocket (Exclusive to V.I.P. members)

To download this full book, simply select the format you desire below

