

文章编号: 2095-4980(2023)01-0044-07

短波窄带莫尔斯信号自动识别技术

刘长远^{1,2}

(1. 中国电子科技集团公司第五十四研究所, 河北 石家庄 050081; 2. 河北省电磁频谱认知与管控重点实验室, 河北 石家庄 050081)

摘要: 短波环境中存在着大量令人感兴趣的莫尔斯信号, 为了降低人工侦收成本, 研究了莫尔斯信号在短波环境下的窄带信号自动识别技术。针对短波信道环境, 提出一种联合时频域特征的快速识别算法, 并给出短波环境下莫尔斯窄带识别算法的流程设计, 最后获得算法的实测结果。实验结果表明, 该算法下的莫尔斯识别具有较快的速度和较好的识别率, 以及较低的虚警率和漏警率。

关键词: 短波; 莫尔斯; 窄带; 信号识别

中图分类号: TN911

文献标志码: A

doi: 10.11805/TKYDA2020468

Automatic identification technique of Morse signal in shortwave narrowband channel

LIU Changyuan^{1,2}

(1.The 54th Research Institute of CETC, Shijiazhuang Hebei 050081, China;

2.Hebei Key Laboratory of Electromagnetic Spectrum Cognition and Control, Shijiazhuang 050081, China)

Abstract: There are large amount of Morse signals of interest in shortwave environment. In order to reduce the workload of the related personnel, an automatic identification technique of Morse signals in shortwave narrowband channel is put forward. Based on analyzing the problems of Morse signal identification in short wave narrowband channel, a fast identification method combining with the characteristics of time and frequency domains is proposed and a new Morse signal identification process is designed. By using the proposed method, Morse signal can be identified very quickly with high identification probability, low leak alarm rate and false alarm rate.

Keywords: shortwave; Morse; narrowband; signal identification

莫尔斯信号作为当今世界通信体制中重要的一种通信方式, 它通过一系列特定的“点”、“划”组合来表示对应的英文字母、阿拉伯数字等符号进行发报, 发报者与收报者之间只要拥有同样的报文规定, 就可以实现通信, 因此莫尔斯电码具有编码方式简单、易于实现等特点^[1]。但在短波段中, 信号环境往往十分复杂, 噪底起伏多变, 这就增加了接收信号的难度, 且由于抄报的人工成本过于高昂, 借助计算机实现莫尔斯信号的识别就显得意义重大^[2]。美国学者 Gold B^[3]曾在 1959 年公开发表过相关研究, 但时间久远, 方法陈旧。国内对莫尔斯的识别研究主要分为 3 种方法: 时域分析法、频域分析法和时频域分析法。包络检波法^[4]属于时域分析法, 该方法具有较强的实时性, 但抗干扰能力差, 无法适应复杂的短波环境; 锁相跟踪^[5]、频率模拟和卡尔曼滤波等^[2]频域分析技术容易受到单频信号的影响, 在短波环境中会存在很高的虚警; 文献[6]采用图像处理的时频分析方式, 处理所需数据量大, 且需要提前进行样本训练, 识别更新速度慢, 无法做到快速识别。本文提出了一种在短波环境下, 基于时频信息的莫尔斯信号窄带快速识别方法, 较好地克服了因噪底不平、信号衰落等原因无法识别信号的问题。

1 莫尔斯信号

常规莫尔斯信号在时间上分为传号和空号时段, 传号由点和划 2 种符号组成, 一般情况下, 标准的莫尔斯信号点划的长度比为 1:3。空号根据不同的编码规则, 与传号的点符号时间长度相比, 一般点空比为 1:1、1:3 和

1:5。莫尔斯信号可以用式(1)表示为：

$$S(t) = AS_d(t)\cos(w_0t) \tag{1}$$

式中 $S_d(t) = \begin{cases} 1 \\ 0 \end{cases}$ ， $S_d(t)$ 为 1 时，表示莫尔斯以概率 P 发传号信息，为 0 时，表示莫尔斯以 $1-P$ 的概率发空号信息。

2 短波窄带信号识别特点

当前，短波窄带信道内的信号识别仍有较大困难，由于短波环境的复杂，信道十分拥挤^[7]，各种信号混杂，信号本身带宽各异，有的信号带宽可以达到好几万赫兹，而莫尔斯信号可视为单一频率。同时信道间隔也并不固定，带宽较窄的信号频率间隔能达到百赫兹量级，以采样率 9.6 kHz 为例，带宽为 4.8 kHz 的窄带信号极有可能同时存在多个信号。

同时，由于短波远距离通信依靠电离层反射，电离层长期处于不稳定状态，这就导致当信号入射电离层时，一部分被电离层吸收，一部分被电离层反射发生极化变化，从而导致信号衰落^[8]。

短波传输环境也存在因路径不同而产生的多径干扰以及多普勒频移现象等。这些都会造成信号时强时弱，信号质量差，从而给识别带来困难。图 1 是实采的短波莫尔斯信号时域波形，信号的衰落十分严重^[9]。

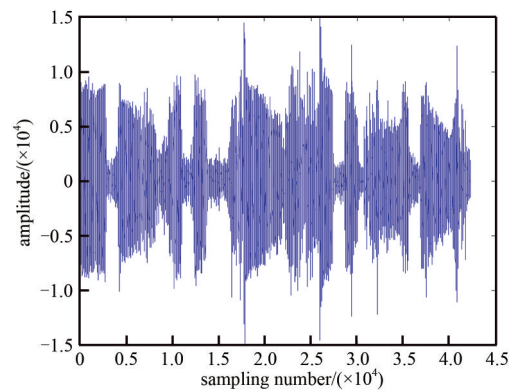


Fig.1 Time domain of the Morse signal in shortwave
图 1 短波莫尔斯信号时域图

3 莫尔斯识别算法

3.1 识别流程

本文提出的莫尔斯信号的识别流程如图 2 所示，将接收机的窄带数据送入滤波器中进行形态学滤波，消除信号带宽内噪底不平的影响。之后在平稳噪底上进行频域特征检测，由于莫尔斯信号的频域特征与单音一致，则可按照单音的信号检测准则进行。然后对符合频域检测准则的信号频点进行窄带滤波，获得只包含该信号的时域数据后进行时域包络变换，将变换后的数据进行莫尔斯点划比等特征匹配，输出满足特征匹配要求的数据及数据个数。若输出信号个数等于 1，则直接输出结果；若输出信号个数大于 1，将各信号之间进行时域互相关，根据互相关系数大小判断输出结果。

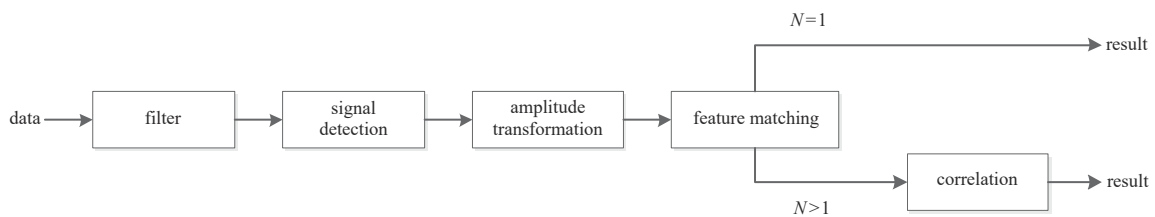


Fig.2 Algorithm processing
图 2 算法流程图

3.2 形态学滤波

本文研究的莫尔斯信号处在短波环境中，短波通信的主要问题如前文所述，则莫尔斯信号会随着噪底变化而高低起伏。当使用固定门限去判断是否存在信号时，如果莫尔斯信号正好位于频谱的低噪声位置，就会发生漏警。因此，在进行莫尔斯信号的识别时，需要先对数据进行形态学滤波的操作。

设定义域为 $F = \{0, 1, \dots, N-1\}$ ，信号 f 为该定义域上的离散函数，又有数据元素 a 是 $A = \{0, 1, \dots, M-1\}$ 上的离散函数，且满足 $N > M$ ，则一维的形态变换公式为：

$$\begin{cases} f(x) \text{关于} a \text{的膨胀: } (f \oplus a)(x) = \max_{m=0,1,\dots,M-1} \{f(n+m) + f(m)\}, & n=0, 1, \dots, N-M \\ f(x) \text{关于} a \text{的腐蚀: } (f \ominus a)(x) = \max_{m=0,1,\dots,M-1} \{f(n-m) - f(m)\}, & n=0, 1, \dots, N-M \end{cases} \tag{2}$$

根据式(2), f 关于 a 的开闭运算为:

$$\begin{cases} O_a(f) = (f \ominus a) \oplus a \\ C_a(f) = (f \oplus a) \ominus a \end{cases} \quad (3)$$

对信号的频域进行膨胀运算, 可以减小信号频谱的谷值, 拉宽峰值; 而频域的腐蚀运算则与膨胀运算正好相反, 会减小信号的谱峰, 加宽谷值。开运算可以将小于数据元素的信息“抹掉”, 像噪声引起的频域毛刺或者突发等, 以此来抑制信号的峰值信息; 而闭运算则能够将小于数据元素的信息进行填充, 能够实现对信号波谷区域的抑制^[10]。

当接收机采集的窄带数据送到处理器后, 先对它进行傅里叶变换求得频谱数据, 之后对频谱数据进行形态学滤波, 前后的效果如图3所示。

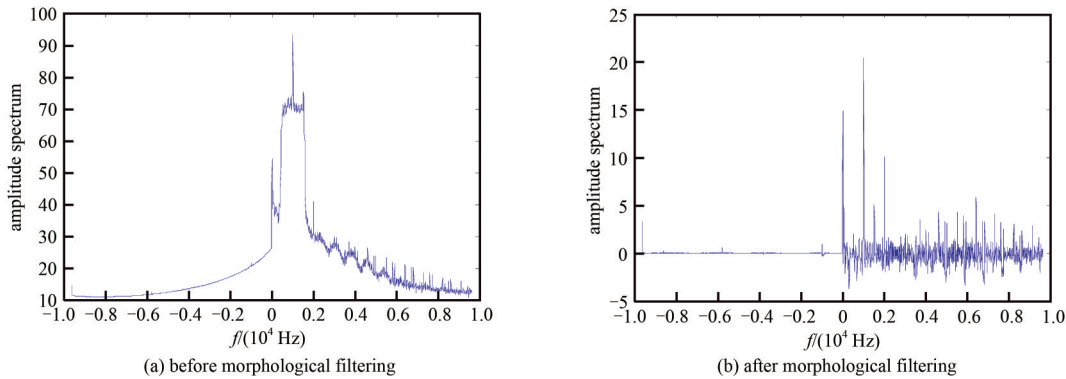


Fig.3 Morphological filtering effect
图3 形态学滤波效果图

3.3 时域包络变换

经过形态学滤波后, 可以找到 N 个单一频点, 对每个频点进行带通滤波, 并尽可能少地保留噪声分量, 得到数据 $S(t)$, 再求其时域包络。当其中含有莫尔斯信号时, 对 $S(t)$ 求时域包络, 推导为:

$$R(t) = S(t) \cdot \cos(w_0 t + \varphi) = A \cdot S_d(t) \cos(w_0 t) \cos(w_0 t + \varphi) \quad (4)$$

$$I(t) = S(t) \cdot \sin(w_0 t + \varphi) = A \cdot S_d(t) \cos(w_0 t) \sin(w_0 t + \varphi) \quad (5)$$

对 $R(t)$ 和 $I(t)$ 进行滤波, 可求得其包络为:

$$B(t) = \sqrt{R^2(t) + I^2(t)} = \sqrt{\left(\frac{A \cos \varphi S_d(t)}{2}\right)^2 + \left(\frac{A \sin \varphi S_d(t)}{2}\right)^2} = \frac{A}{2} S_d(t) \quad (6)$$

由于短波信道的复杂性, 直接求得的信号包络也会存在衰落, 需要进行适当的操作减少衰落对信号的影响。非线性变换是数字信号处理中一个常见的操作, 它能将感兴趣的信息处理得更具有针对性, 使后面的数据处理更容易, 此处可以使用非线性变化的方法对信号进行处理。

在对数据进行非线性变换时, 一般变换的基准为:

1) $f(x) \in [x_{\min}, x_{\max}]$, 当 $x \in [x_{\min}, x_{\max}]$;

2) 当 $x_1, x_2 \in [x_{\min}, x_{\max}]$ 时, 且 $x_1 > x_2$, 则有 $f(x_1) > f(x_2)$ 。

2个基准说明了一般非线性变换的特点, 一个是一般要求 $f(x)$ 与 x 的取值范围相同; 另一个要求非线性变换函数 $f(x)$ 是一个单调非减函数。目的是保证数据在变换前后分布范围相同, 且仅改变数据在空间中的距离大小, 而不改变数据的排列顺序^[11]。

常见的非线性变换函数有指数函数、多项式函数和分段函数。但当研究的对象是短波环境下的莫尔斯信号时, 信号不仅自身存在衰落, 还有可能在莫尔斯信号开始时有较高的冲击幅值, 如图4(a)所示。使用上述的非线性变换函数并不能很好地适应短波环境下莫尔斯识别, 本文设计的非线性变换函数是反正切函数, 该公式为:

$$f(x) = \frac{1}{\pi} \arctan(\alpha(x - 0.5)) + \beta \quad (7)$$

式中: α 为伸缩系数, 当 α 值越大, 函数 $f(x)$ 就越陡峭, 相反, 函数 $f(x)$ 变化就越平缓; β 为偏移系数, 该值依据

x 的取值范围和 α 值进行适当的设置，保证 $f(x)$ 在固定范围内变化。该非线性函数类似分段函数，在定义域上具有单调非减的性质，且能保证值域在 $[0, \pi]$ 之间变化。但不同于分段函数，该函数是平滑曲线，这保证了数据在后面进行包络的突发检测时具有连续性，不会发生信号突变影响突发检测。

图 4(a) 是直接求得的莫尔斯信号时域包络，图 4(b) 是经过该非线性变换函数后的信号时域包络。可以发现信号经过非线性变换后，信号部分更加明显，同时噪声部分也被抑制。

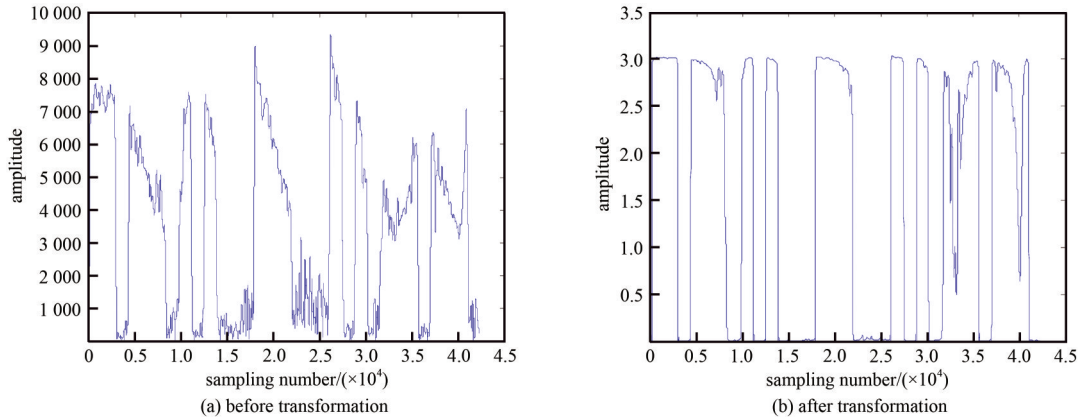


Fig.4 Temporal envelope
图 4 时域包络图

3.4 突发检测

当信号经过非线性变换，获得信号的时域包络之后，需要检测信号有能量的部分是否满足莫尔斯点划比和点空比。本文采用的突发能量检测方法为双滑动窗口法。通过设置 2 个时间长度一样的连续滑动窗口，并分别计算窗口中的数据能量，用能量比作为判决变量。双滑动窗口法的原理框图如图 5 所示。用 A_n 和 B_n 表示窗口 A、B 的数据能量和，公式如下：

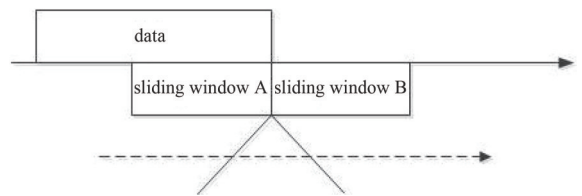


Fig.5 Schematic of double sliding window
图 5 双滑窗原理图

$$A_n = \sum_{l=0}^{L-1} r_{n+l} r_{n+l}^* = \sum_{l=0}^{L-1} |r_{n+l}|^2 \tag{8}$$

$$B_n = \sum_{l=0}^{L-1} r_{n+L+l} r_{n+L+l}^* = \sum_{l=0}^{L-1} |r_{n+L+l}|^2 \tag{9}$$

式中： L 为窗长； r_n 为第 n 个采样值，则判决量 $m_n = B_n/A_n$ 为滑窗 A 和 B 中的数据能量比。当 2 个滑窗都只包含噪声时，判决量 m_n 的比值趋近于 1。当窗口 A 中含有信号能量逐渐增多，窗口 B 中不含有能量时，判决量 m_n 则会逐渐变大；相反的，当窗口 A 中不含有信号能量，窗口 B 开始包含信号能量，随着窗口 B 中的能量逐渐增大， m_n 又开始逐渐下降。

由图 4 可知，信号经过非线性变化后，信号在时域幅度上进行了拉伸和压缩，利用双滑窗进行突发检测时会很大的起伏。工程上，一般将判决门限设置在 10~20 dB 之间。这样通过观察 m_n 取值与判决门限的关系，也就可以找到信号能量的突发和截止时刻，从而获得各对应频点数据的时域特征，这些特征包括信号的突发位置、突发持续时间、突发间的长度比值等，之后将这些特征信息和莫尔斯信号的点划比与点空比进行匹配，初步判定是否满足莫尔斯时域信号特征。

3.5 莫尔斯的互相关

对于 4.8 kHz 的信号带宽，经过前文叙述的判决方法得到的信号个数常常会大于 1，此时就需要甄别该信号究竟是频移键控(Frequency-Shift Keying, FSK)类信号造成的虚警，还是谐波莫尔斯信号，本文采用互相关法来进行区分^[12]。互相关法是用来观察 2 个信号波形的相似程度和到达最大相似度时 2 个信号的时间差的大小。而在窄带信号识别中，识别相近的 2 个频点是否是莫尔斯数据，可以利用互相关法对 2 个频点数据的相似程度来进行区分。

设有 2 个信号 $f(x)$ 和 $g(x)$, 则它们的相关系数被定义为它们相同时间内数据乘积的积分除以各自平方的几何平均值之积, 公式有:

$$r = \frac{\int f(x)g(x)dx}{\sqrt{\int [f(x)]^2 dx} \sqrt{\int [g(x)]^2 dx}} \quad (10)$$

在离散数据中, 则有:

$$r = \frac{\sum_{n=1}^N f(n)g(n)}{\sqrt{\sum_{n=1}^N [f(n)]^2} \sqrt{\sum_{n=1}^N [g(n)]^2}} \quad (11)$$

由上式可知, $|r| \leq 1$ 。当且仅当一段数据中 2 个频点信号随时间变化的规律相同时, 即为谐波莫尔斯信号时, 理论上 $r=1$; 对于 2FSK 信号, 2 个频点信号随时间变化的规律正好相反, 理论上 $r=-1$; 当只是 2 个不相干的莫尔斯信号时, 理论上 $r=0$ 。工程上, 可设定门限为 -0.5 , 当互相关系数大于门限时, 可认为是莫尔斯信号。

4 实验结果与分析

4.1 实际数据识别结果

通过多组实际窄带信号, 测试该算法的识别能力。算法设计的单次识别数据时间长度为 2.2 s, FFT 点数为 16 k。第 1 组为 2FSK 信号, 第 2 组为手摇发报莫尔斯信号, 第 3 组为莫尔斯与多音混合信号, 第 4 组为语音信号, 第 5 组为 2FSK 和莫尔斯信号, 第 6 组为扫频与莫尔斯信号。6 组数据的语图如图 6 所示。

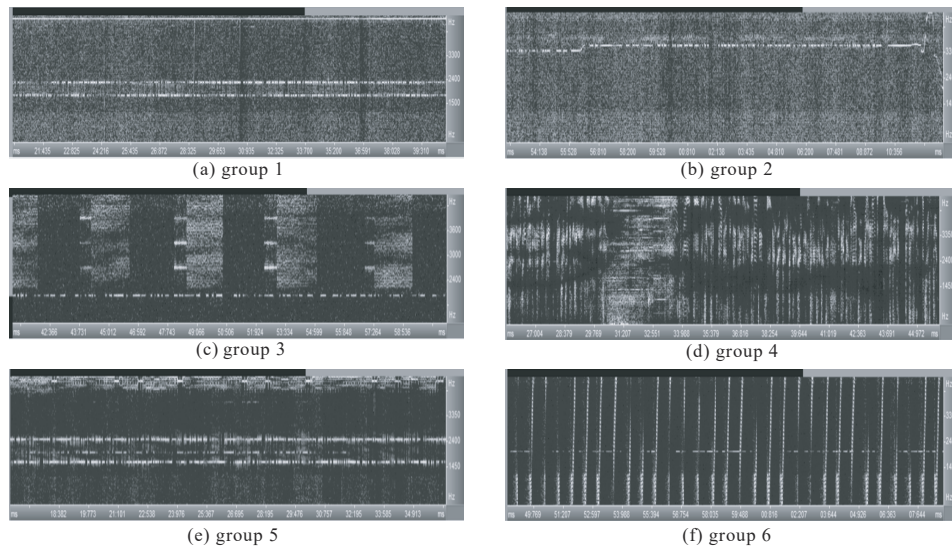


Fig.6 Signals spectrogram

图6 信号语图

将上述的实采数据进行该算法的实际识别测试, 统计莫尔斯信号的识别率和虚警率, 识别结果如表 1 所示。其中识别率(Accuracy)为正确识别的信号识别次数(Correct times)与实际应识别出的总次数(Correct times and Missing times)的比值, 虚警率(False alarm rate)为错误识别的信号识别次数(Error times)与识别出是莫尔斯信号的总次数(Identification times)的比值。

表1 识别结果

Table1 Identification results

signal samples	group 1	group 2	group 3	group 4	group 5	group 6
identification times	53	31	26	62	54	29
correct times	0	27	26	0	43	24
error times	1	1	0	2	1	1
missing times	0	1	0	0	1	0
accuracy/%	0	96.43	100	0	97.73	100
false alarm rate/%	1.89	3.23	0	3.22	1.85	3.45

由表 1 的结果可以发现，在确实存在莫尔斯信号的 4 组实采样本中，识别的正确率都在 95% 以上，说明该识别技术对莫尔斯信号的识别效果较好。

对样本添加不同的噪声，使信号的信噪比在 5~15 dB 之间，每组样本进行的蒙特卡洛次数为 20 次，统计莫尔斯信号识别的情况，结果如图 7 所示。

从图 7 可以看出，整体上，随着信噪比的增加，该算法对莫尔斯信号的识别正确率也不断上升，虚警率也在不断下降。当信噪比大于 9 dB 时，正确识别的概率在 95% 以上，虚警率在 5% 以下。

4.2 对比实验

文献[6]的算法是基于模式识别，通过图像处理的方法提取时频特征进行识别的。该算法同样适用于数据文件同时存在多种信号的识别环境，且在 FFT 点数在 128 k 以上，信噪比 10 dB 以上的条件下，识别准确率达到 95% 以上。

为了与文献[6]的算法对比，采用实采的数据对其进行训练，并让两种算法的 FFT 点数一致，均为 64 k，测试样本文件的莫尔斯信号信噪比为 15 dB，通过改变输入数据的时间长度，查看算法识别的准确率，识别结果如表 2 所示，其中算法 1 为本文算法，算法 2 为文献[6]算法。通过对比可以发现，本文算法可以实现短数据量的快速识别，且识别率均在 90% 以上，而

文献[6]的算法在数据量少时，识别正确率急速下降，要想获得 90% 以上的正确率，识别需要的数据时间长度要在 4 s 以上。可见本文提出的算法能更快地给出识别结果，且整体上准确率更高。

5 结论

本文提出了一种适用于短波环境下的莫尔斯识别算法。利用形态学滤波能将窄带带宽内位于谷域的莫尔斯信号检测出来，再利用非线性变换对信号的时域包络进行处理，将变换后的信号进行莫尔斯编码的时域特征匹配，可以将质量较差的莫尔斯信号识别出来。实验证明，在多种信号类型并存的短波环境下，本算法在 9 dB 以上的检测正确率可达 95% 以上，漏检率和虚警率分别低于 5% 和 3%。该算法运算复杂度低，运算速度快，单次识别需要的数据量小，迁移性强，工程易于实现，且已经应用于某短波实际工程中，大大减轻了业务员的人工识别工作量，获得了较好的效果。同时该算法的原理适用于多信号同时存在的识别，具备推广到宽带上的条件，与宽带谱峰检测相结合，为后续继续开展宽带上的莫尔斯快速识别提供了一种解决思路。

参考文献：

[1] 王志永,刁文静,李冬梅. 短波宽带信道莫尔斯信号自动识别技术[J]. 计算机与网络, 2011(8):57-59. (WANG Zhiyong,DIAO Wenjing,LI Dongmei. Research on automatic identification of Morse signal in shortwave wideband channel[J]. Computer and Network, 2011(8):57-59.)

[2] 林兴龙. 人工莫尔斯光信号自动识别技术的研究[D]. 北京:北京邮电大学, 2019. (LIN Xinglong. Research on automatic recognition technology of artificial Morse optical signal[D]. Beijing:Beijing University of Posts and Telecommunications, 2019.)

[3] GOLD B.Machine recognition of hand-sent Morse code[J]. Information Theory Ire Transactions, 1959,5(1):17-24.

[4] 陈莹莹,彭仕勇,高勇. 莫尔斯报务信号的检测和识别[J]. 计算机应用与软件, 2009,26(6):241-243. (CHEN Yingying, PENG Shiyong, GAO Yong. Detection and recognition of Morse telegraph signals[J].Computer Applications and Software, 2009,26(6):241-243.)

[5] 马威,张敬修,王虎帮. Morse 电码自动译码系统[J]. 兵工自动化, 2007,26(6):51-52. (MA Wei,ZHANG Jingxiu,WANG Hubang. Automatic decoding system of Morse code[J]. Ordnance Industry Automation, 2007,26(6):51-52.)

[6] 龚智贞,贾克斌,孙中华,等. 宽带环境下 Morse 信号自动检测[J]. 信号处理, 2017,33(8):1097-1102. (GONG Zhizhen,JIA Kebin,SUN Zhonghua,et al. Automatic detection of Morse signals in wide band environment[J]. Journal of Signal Processing, 2017,33(8):1097-1102.)

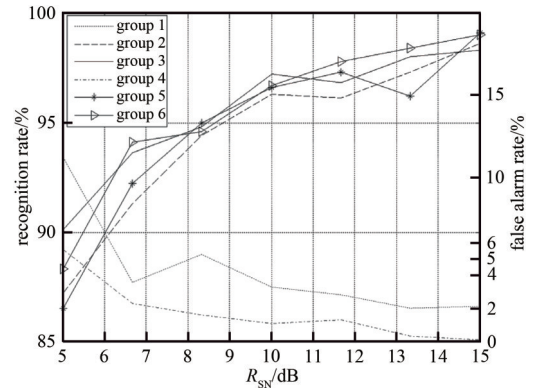


Fig.7 Identification results under different SNR
图 7 不同信噪比下识别结果

表 2 对比算法的识别结果

Table 2 Results of contrast algorithm

time/s	algorithm 1 (accuracy/%)	algorithm 2 (accuracy/%)
8	96.23	97.87
7	96.54	97.26
6	96.40	95.64
5	96.16	93.98
4	95.97	90.12
3	96.21	86.52
2	92.65	80.10