

声呐数据图像化处理的研究与实现

韦波¹, 姚宇青¹, 邹国华², 陈帅¹, 张胜茂¹, 戴阳^{1*}

¹中国水产科学研究院东海水产研究所, 上海

²上海渔联网科技有限公司, 上海

Email: *13162787852@163.com

收稿日期: 2021年3月28日; 录用日期: 2021年4月21日; 发布日期: 2021年4月28日

摘要

针对部分定制普通声呐探测到数据却无法分析数据的问题, 本文运用Python设计一种程序, 将普通声呐所探测到的数据进行图像化处理, 使普通声呐所探测的数据可视化。首先对数据的格式进行解析, 然后运用Python进行异常数据删除和数据格式化, 最后将数据转化为十进制数并画出信号强度图。结果表明, 通过本程序得出的信号强度图能够达到声呐信号强度图的标准。

关键词

可视化, Python, 数据格式化, 信号强度图

Research and Realization of Image Processing of Sonar Data

Bo Wei¹, Yuqing Yao¹, Guohua Zou², Shuai Chen¹, Shengmao Zhang¹, Yang Dai^{1*}

¹East China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Shanghai

²Shanghai Fishing Network Technology Co., Ltd., Shanghai

Email: *13162787852@163.com

Received: Mar. 28th, 2021; accepted: Apr. 21st, 2021; published: Apr. 28th, 2021

Abstract

Aiming at the problem that some custom-made ordinary sonar detects data but cannot analyze the data, this paper uses Python to design a program to image the data detected by ordinary sonar to visualize the data detected by ordinary sonar. First, analyze the format of the data, then use Py-

*通讯作者。

thon to delete abnormal data and format the data, and finally convert the data to decimal numbers and draw the signal strength graph. The result shows that the signal intensity map obtained by this program can meet the standards of the sonar signal intensity map.

Keywords

Visualization, Python, Data Formatting, Signal Strength Graph

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

科学声呐因其价格昂贵，多应用于大型科学研究中，却无法适用于一般的民用作业。因此一般的渔业探测都会选择普通的简易声呐，普通声呐因其价格亲民，且精度能够满足基本的渔业需求，成为很多民用作业的首选。

但是，普通声呐往往无法做到科学声呐的全套功能，如普通声呐中的部分定制声呐，对收集到的数据没有成像的功能，缺少一个专门的信号强度数据成像软件[1]。这一问题也是一直困扰部分定制声呐使用者的一个问题。本文针对这一问题，以定制声呐所采集的数据为样本，对数据进行详细解析并生成图像[2]。

对声呐数据的图像化处理，能够对定制声呐所探测到的数据得到一个可视化的界面，更加方便使用者的探测研究，特别是在鱼群探测、海底地貌勘测、水文测量、海洋资源勘测、军事领域等，声呐数据的图像化显示已经成为了一个不可或缺的部分。

声呐数据可视化就是运用计算机图形学和图像处理技术，将科学计算过程产生的数据和计算结果转换为图形或图像在屏幕上显示。当前的国内外声呐数据图像化处理从数据的可视化上来分，有二维、三维，以至多维，但主要还是以二维图像和三维图像为主。目前国内的声呐数据图像化处理已经相对成熟，在很多民用作业上，均有涉及图像化处理的技术。且多为自主研发，受益于计算机性能的提高，声呐数据图像化处理所需要的实时性、精确性皆能达到基本需求。而国外的声呐数据图像化处理技术，不管是技术水平还是研究资金来说，都具有雄厚的实力，以英国 2076 声呐系统为例，其不管是在数据成像的实时性和精确性上，都是世界领先水平。

2. 系统结构与数据解析

本文中数据的来源均通过声呐采集所得。其中声呐所探测的其中一 ping 数据的形式为 8 bit 一个字节，一个字节为我们常规熟知的两位。本节分为两个部分，第一部分为介绍定制声呐的系统结构，第二部分是对数据的解析。

2.1. 系统结构

本文所分析的数据为声呐为定制声呐所接收的数据，其中定制声呐的系统结构图如下图 1 所示。

声呐的整体系统主要分为两个部分，分别是浮标和接收模块两部分。浮标部分主要包括包含声呐、控制电路、传输模块和天线[3]。声呐位于最底层，负责探测数据，在声呐探测到数据后，由控制电路控

制传输模块通过天线进行数据传输。当浮标部分发出数据时[4]，接收模块开始工作，同样由传输模块和天线协同工作接收浮标部分发出的数据，在接收部分接收到数据后，同时开始数据处理，最终将输出形成可视化[5]。

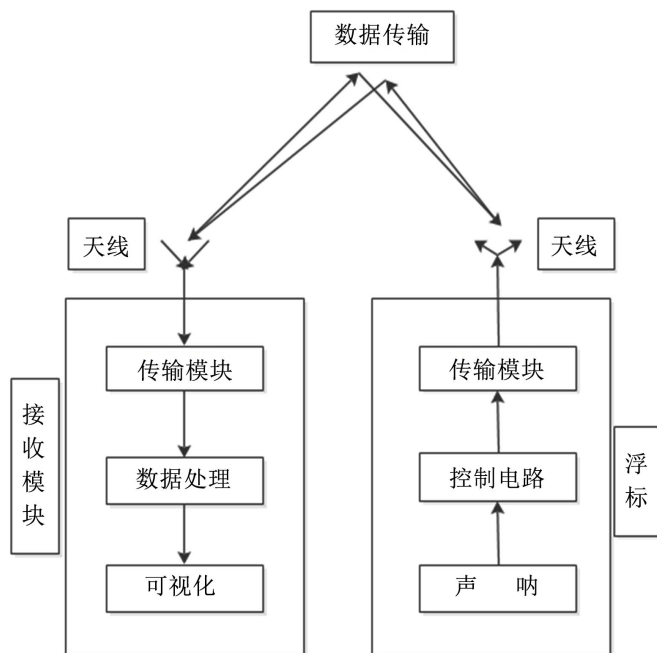


Figure 1. Sonar system structure
图 1. 声呐系统结构

本文的主要目的就是解决数据处理与可视化这两个部分。对接收到的数据进行处理，并同时生成可视化的界面。

2.2. 数据解析

在本文的研究中，数据的来源为通过定制的双频率声呐所探测到的数据，其中声呐的系统结构与工作方式 1.1 中已分析。由于声呐所探测的数据，在是正确数据格式的前提下，其数据格式的组成方式是一样的，因此本文以一 ping 的数据为例，对一次探测所得到的声呐数据进行解析[6]。

在一 ping 的数据中，每两个为一个字节，如开头的“23”就是为一个字节长度。前三个字节为数据头，其中以十六进制数 23 开始，后面两个字节比如为“FE 03”，表示数据长度，该数据长度为小端存储的整型数，因此此处就应该为十六进制数“03FE”，也就是十进制数 1022。代表这一 ping 的数据体有 1022 个字节，即代表“FE 03”后面的 1022 个字节为探测数据。

在数据体中，第一个字节表示硬件增益，下一个字节为工作频率，指出是哪一个光束在工作，第三个字节表示接收通道。第四到七字节为流水号，区分是否丢帧，第八到十一字节为毫秒同步时钟。第十二到十五字节为采样上表面，十六到十九字节为采样最大深度，二十到二十一为回波图的图像像素。接下来的第二十二到二十九字节表示的是采样点的纬度，三十到三十七字节为采样点的经度，第三十八到四十一字节为简化的 GPS 时间，四十二到四十三字节为前行方向，四十四到四十五字节为速度，四十六到四十九字节为天线离海平面高度，五十到五十三字节为大地椭球面相对海平面高度第五十四位表示状态，五十五位到五十八位表示船的实时俯仰角，五十九位到六十二位表示船的实时侧倾角。接下来的两

位表示电压，再接下来的两位表示温度，最后四位表示为当前水深。在全部的数据体中，除去前面分析的七十一个字节，剩余的全部为声呐所探测的信号强度值，为十六进制表示。

在数据体之后的数据的第一位数据为校验位，是数据体各字节的异或校验。接下来的十九个字节不再是十六进制数据，而是用十六进制表示的 ASCII 码，例如“24 4F 52 46 46 49 2C 30 2C 30 2C 30 2C 30 2A 54 0D 0A 00”其表示的意思为“\$ORFFI, 0, 0, 0, 0 * T”，含义为“\$ORFFI, <depth0: 0~99999>, <size0>, <depth1>, <size1> * XOR”，Depth0/1 代表鱼群深度，单位为 cm，Size0/1 代表鱼群信号强度，和回波幅值及持续时间相关，一般有效范围 10~100 之间。

最后的十七位字节为数据传输完毕的结束语句，此处也为 ASCII 码，比如“24 4F 52 43 4D 44 2C 4C 4E 4B 5F 45 4F 54 2A 0D 0A”意思为“\$ORCMD, LNK_EOT*”，含义是表示数据接收完毕的意思。

3. 数据处理

数据处理主要分为两个部分，分别是对无用数据删除和正确但是非格式化数据的格式化两部分[7]。

3.1. 异常数据删除

在对所接收的数据进行观察后不难发现，对于声呐所接收的数据，无用数据主要是“7777772E 75 73 72 2E 63 6E 20 4F 4B”这样的连续数据。实际这些数据并非异常数据，而是心跳包，但是对本文的数据画图并无用处，因此这里将其作为异常数据删除。

对于接收到的声呐数据，将其存放在 txt 文档中，并用 PyCharm 打开，本次对数据处理的 PyCharm 版本为 2019.3，Python 版本为 3.7.0。在 PyCharm 中打开 txt 文档后，通过模式匹配方法，将文档中与“7777772E 75 73 72 2E 63 6E 20 4F 4B”相匹配的字节直接删除。

同时还有一种情况是在接收数据时，会存在一行数据前若干字符是符合正确格式数据样式的形式，但是突然会断开，直接开始下一 ping 的数据录入。这就导致录入的数据不完整，则对于未完整录入的数据，也需要进行删除处理，此处删除的方法是，寻找完整的数据作为起始，接着寻找下一个完整的数据，若下一个数据并非 1.1 数据解析满足的格式，则可断定为异常数据，可通过代码将此异常数据删除。

3.2. 数据格式化

通过在 PyCharm 中打开数据可以发现，对于声呐所接收的数据，正确的数据都是每一 ping 占一行。要进行数据格式化是因为在数据接收过程中，有些数据并没有进行换行，而是接在了上一行数据的后面。而有些数据则是在没有完全接收完毕的情况下就自动换行。

针对上述问题，同样运用 Python 语言，对 txt 文本中的数据进行遍历，由于在 2.1 中已经对数据进行了处理。因此此处的数据均为正确数据，只需要对数据的格式进行统一化处理即可，也就是设置为每一 ping 的数据刚好占一行的形式，方便下一步的画图处理。在对文本的遍历过程中，只需要按照 1.1 的格式要求，将每一 ping 的数据独立占一行即可。

4. 数据图像化

4.1. 数据成像

在经过以上 1 和 2 的处理后，本节就是完成对数据的画图部分。在实际接收数据的过程中，由于每个数据体的字节数是不一样的，因此这里为了方便画图，利用 Python 在每行数据的开头添加一个接收时间，格式可为【xxxx】，其中括号的内部可以为时间，或者其他一些信息，由于此处只是为了方便后面的画图，因此对于括号的内容不作过多要求，最终可得到如下图 2 所示的在 Python 中打开的数据头。

```

3      [2021-01-16 15:13:31:330] 23 2E 04 18 01 00 80
4      [2021-01-16 15:13:33:649] 23 2E 04 18 01 00 81
5      [2021-01-16 15:13:35:336] 23 2E 04 18 01 00 82
6      [2021-01-16 15:13:37:249] 23 2E 04 18 01 00 83
7      [2021-01-16 15:13:39:487] 23 2E 04 18 01 00 84
8      [2021-01-16 15:13:41:490] 23 2E 04 18 01 00 85
9      [2021-01-16 15:13:43:190] 23 2E 04 18 01 00 86
10     [2021-01-16 15:13:45:284] 23 2E 04 18 01 00 87
11     [2021-01-16 15:13:47:289] 23 2E 04 18 01 00 88
12     [2021-01-16 15:13:51:335] 23 2E 04 18 01 00 8A
13     [2021-01-16 15:13:53:330] 23 2E 04 18 01 00 8B
14     [2021-01-16 15:13:57:372] 23 2E 04 18 01 00 8D
    
```

Figure 2. Data header after formatting
图 2. 格式化处理后的数据头部分

在完成以上步骤后，接下来就是对数据的画图。这里仍然选择用 Python 作为画图工具。

由于前面部分我们已经对数据进行解析并且处理完毕，因此这里对于画图部分，需要做的就是将数据转化为常规所认知十进制数。完成数据转化后，就是对数据的坐标定义，通过对十进制信号强度数据统计可以发现，信号强度均在 0~130 之间。且信号强度越小，数据量越大。因此，为了成像效果的显著性，这里选择将 81~130 设置为一个颜色，同理分别将 41~80、31~40、21~30、16~20、11~15、6~10、4~5、3、2、0~1 设置为不同颜色。同时也可以针对鱼群的实际分布数量及信号强度大小设置不同的颜色区间。

在横纵坐标的确定上，横坐标设置为第一行的横坐标为 1，第二行数据的横坐标为 2，依此类推。而纵坐标的确定上，需要结合信号强度值的数量以及深度来确定，用深度除以信号强度值的数量，即可得出纵坐标的间隔，从而确定纵坐标。

在完成以上步骤后，运用 Python 的 matplotlib 函数[8] [9]，通过对样例数据的测试，可得出如下图所示的信号强度图，其中横坐标单位为个，纵坐标单位为米。

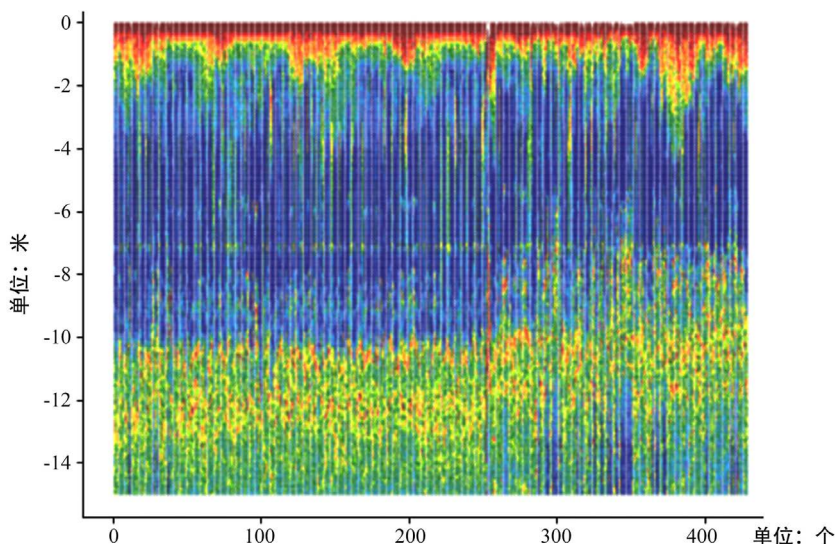


Figure 3. Sonar signal intensity graph
图 3. 声呐信号强度图

结合图 3 的信号强度图可以发现, 在深度 0 到 0.7 左右的深度信号强度较大, 这与实际探测得到的结果一致, 原因是水面的噪声比较大, 导致出现水面的信号强度明显大于水底的现象。中部水深 6 左右的深度间断性出现非一致强度的信号, 即是有物体的信号, 也就是可能有鱼。在深度 7 左右的地方, 出现相对连续性的信号强度, 这可能是出现连续性物体, 比如渔网之类的物体。深度 10 以下的信号强度不同于其上面, 但信号强度又相对上面较大且均匀, 这可能是因为探测水域为淤泥质底部的原因。

4.2. 方法评价

以上通过 Python 编程对定制声呐所探测到的数据进行图像化处理的方法, 能够很好的将定制声呐所探测到的数据进行图像化处理, 得到可视化的结果, 且成像的代码较为简单, 方法思路清晰, 具有一定的参考价值。但是, 此方法也存在一个适用范围的问题, 想要适用于其他种类声呐时, 需要对代码进行修改。

5. 总结

声呐数据作为水声数据的一种, 对于科研和水上作业有着重要的作用。普通声呐虽然没有科学声呐的精度, 但是综合对比科学声呐, 普通声呐更加适合于一般的民用作业或精度要求不是特别高的场景。声呐数据因其数据量大的特点, 一般方法处理起来相对困难, 且使用不同的方法, 声呐的显示效果图也不同。本文使用 Python 对数据进行处理和成图, 经过对比测试, 基本满足了需求, 能够作为参考样例。且本文的执行较为方便, 具有一定的学习价值。在未来的声呐发展中, 声呐数据成像是一个不可避免的趋势, 数据成像是整体数据的分析, 对研究者具有极其重要的作用。同时, 声呐数据图像化在未来对于日常作业、科学研究等都将更加普及。

6. 项目基金

中国水产科学研究院东海水产研究所中央公益性科研院所基本业务费专项(2019T03)。

上海市科技计划项目资助, 项目编号: 202H1047900, 上海市杨浦区 2020 年博士后创新实践基地科研项目资助。

参考文献

- [1] 闫志国, 宛楠, 严迪, 许超, 秦逸飞, 齐前. 基于 Scrapy 爬取电商药品数据及数据可视化分析处理[J]. 轻工科技, 2021, 37(3): 98-100.
- [2] 刘忠杰. 声呐系统的高分辨率多波束成像技术[J]. 舰船科学技术, 2020, 42(12): 154-156.
- [3] 李季, 孙同晶, 刘桐. 融入频域先验信息压缩感知方法的主动声呐回波信号处理[J]. 声学技术, 2020, 39(5): 532-539.
- [4] 程浩. 声呐浮标的关键结构设计[J]. 声学与电子工程, 2020(3): 23-25+46.
- [5] 蒯小燕. 水声网络中 DSSS/OFDM 通信技术研究[D]: [博士学位论文]. 厦门: 厦门大学, 2017.
- [6] 赵昆, 王方旗, 丁继胜, 隋海琛, 许方正. 侧扫声呐原始数据解析与可视化显示[J]. 海岸工程, 2020, 39(2): 103-110.
- [7] 陈雷, 王建功, 张旭. 通用雷达数据格式化分析与预处理系统[J]. 中国科技信息, 2011(8): 103+109.
- [8] 齐虎春. 基于网络信息采集的多线程与多进程性能研究[J]. 内蒙古科技与经济, 2020(24): 97-98.
- [9] 薛先贵, 黎路. 数据可视化教学中 Matplotlib 教学应用探讨[J]. 福建电脑, 2020, 36(12): 215-216.