

研究背景与问题重述

近年来,随着国际持枪犯罪案件数量的不断增加,准确有效地判断和分析枪弹痕迹已经成为涉枪案件侦破的关键因素。目前,尽管我国和部分发达国家开发了弹头识别系统,然而多数情况下对弹头痕迹的识别仍然采用直接观察的方法^[1]。这些经验性的传统方法存在效率低,弹头样本不易保存等诸多缺点。如何合理利用现代高精度数据采集设备以深入挖掘枪弹头痕迹体现出的相似性和相关性,成为摆在有关专家面前的重点研究问题。因此,建立数学模型对枪弹痕迹的自动比予以指导,根据已有数据判别一个弹头与样本弹头的相似程度,或判断两个未知弹头的相关程度,并充分提取枪弹痕迹自动比对过程中的关键点和特征因素,具有很大的现实意义和社会价值。同时在实现自动对比关键特征量的选取过程中,应该重点关注的是自动对比方法的准确性,智能性,自动性和可操作性。

77 制式手枪是一种常见的手枪,其弹头上留下的痕迹分为 4 片主棱线的痕迹和 4 片次棱线的痕迹。目前,通过光学设备我们得到该手枪的 4 条次棱弹头表面上若干点的空间坐标 (x,y,z) 数据。这 4 条次棱的数据属于 4 个独立的空间坐标系。测量时弹头圆柱中心线尽量保证平行于基准平面, y 轴尽量平行于擦痕的走向。目前我们得到了部分相同枪和未知枪发射发射的子弹 4 条次棱的坐标样本。通过以上数据,我们的研究重点内容将包含以下几个方面:

1. 由于在光学设备上采集弹头上痕迹的 3 维数据时,人工调节很难使两个弹头在同样位置和同样姿态情况下被测量。通常会造成一定的平移误差和转动误差。如何处理由位置和姿态造成的测量误差,使得两个弹头可以尽量在相同位置和相同姿态的情况下进行比对。
2. 弹头表面的损伤、锈迹、油斑、杂质物及痕迹生成的随机性,会造成数据误差和噪声。如何通过数值处理的方式消除这些误差和噪声。
3. 已知 6 支枪发射的 12 个弹头(每支枪对应两个弹头)4 个次棱部分按固定顺序的测量数据,依据这些数据,给出可以用于比对的特征;并给出提取这些特征的方法;依据这些数据,判断采用数据哪一部分用作比对,其效果比较好;给出完整的比对方案、算法,并在电子版附件中给出程序。
4. 已知另外 11 支枪发射的 22 个弹头(每支枪发射 2 个弹头)4 个次棱部分的测量数据,用解决第三个问题的方法给出这 22 个弹头痕迹两两之间的相似程度;根据以上结果,用列表的方式给出每个弹头按相似度由高到低给出与之相似的前 5 位的弹头文件号。

针对以上问题,我们处理以上问题的整体思路可用以下流程图进行整体说明。

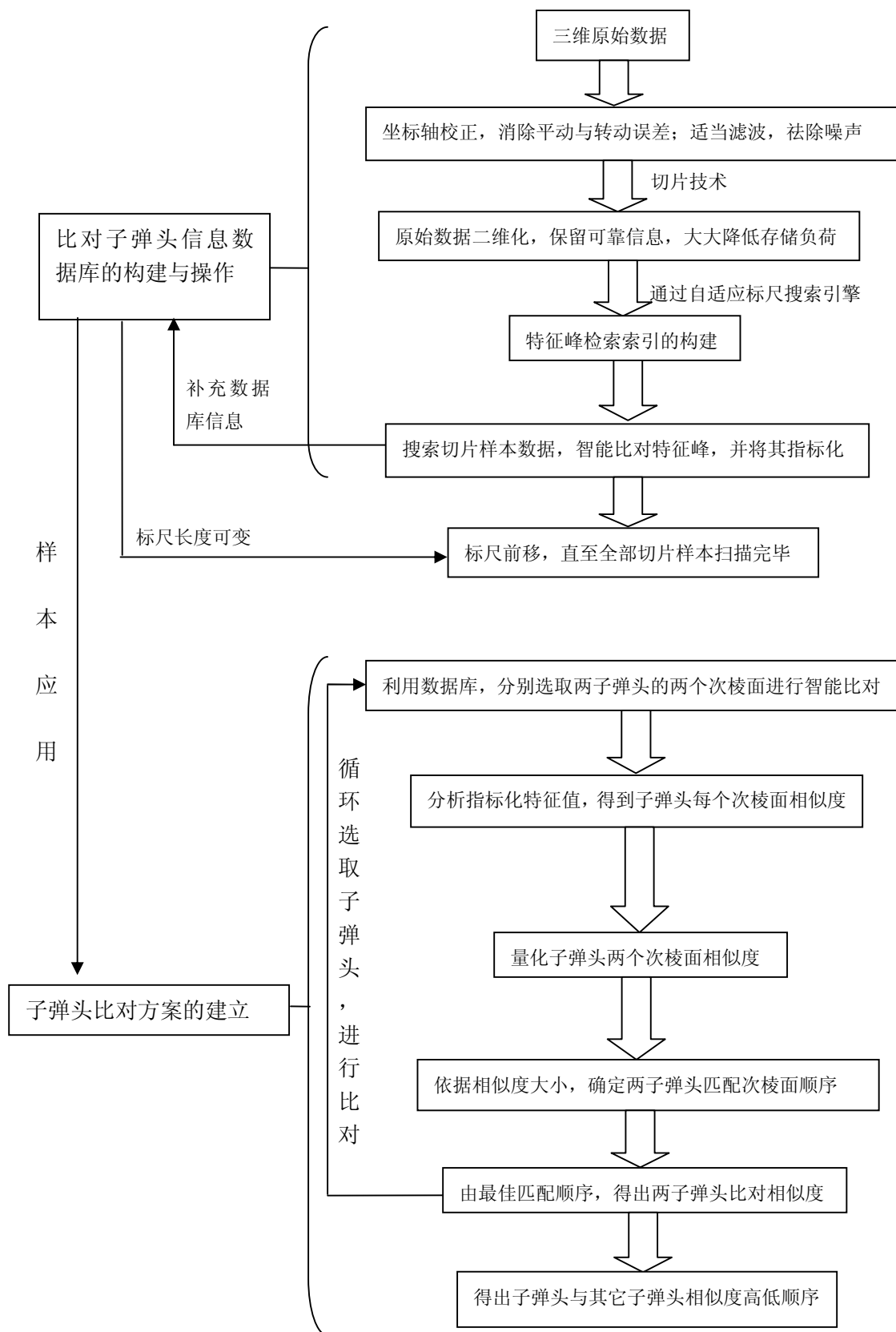


图 1: 本文研究的整体流程框架图

问题 1 的回答与求解：

问题分析：

由于测量条件带来的平动误差和转动误差，以及测量两个不同弹头时参考系的不统一性，导致了两个弹头无法在相同姿态和相同条件下进行对比。因此首先需要分析和运用已有数据，判别次棱部分划痕的位置，同时将不同弹头的 4 个次棱部分的空间坐标根据特定的空间变换方式转变到相同的坐标系或服从一定的统一标准条件下。从测量时的基本准则考虑（即弹头圆柱中心线尽量平行于基准平面；y 轴尽量平行于擦痕的走向），光学仪器实际测得的数据通常只会反应出擦痕走向和 y 轴的关系，而不会反映出弹头圆柱中心线与基准平面的关系，因此该测量准则很难成为次棱坐标进行空间变换的标准。从简单的理想情况出发，若子弹近似看做圆柱形，则划痕走向应将近似与基准平面平行（实际情况有偏差）。因此，我们将判别的基本准则定为：擦痕走向与基准平面平行；y 轴平行于擦痕的走向。

运用空间变换首先将 4 个次棱坐标通过旋转或平移变换为满足测量的基本准则可以实现两个弹头尽量在相同位置和相同姿态情况下进行比对的目标。

问题求解：

基于以上问题分析，将弹头坐标进行变换的过程示意图可简要由图 2 表示，图中 α , β , γ 分别代表擦痕偏离感光器正上方的角度；擦痕平面偏离基准 XOY 平面的角度和划痕偏离于 y 轴的角度。 α 主要由转动误差引起。 β , γ 角可能由转动误差，平移误差或人为误差共同导致。经过图 2A 过程，可使弹头划痕平行于 XOY 平面；图 2B 过程使弹头划痕平行于 y 轴。满足 2 个测量标准的一般步骤可归纳为：

1. 以一定步长用 XOZ 和 YOZ 平面分别沿 y 方向和 x 方向切割次棱坐标所构成的曲面空间。平面切割得到的两簇系列曲线可以分别判断出擦痕的起始和终止位置。同时，XOZ 平面切割得到的一簇曲线可以判断擦痕线与基准平面的偏离程度（若二者平行则跳过步骤 2）；YOZ 平面切割得到的一簇曲线可以判断擦痕与 y 轴的偏离程度（若二者平行则跳过步骤 3）。
2. 将次棱坐标在空间首先分别以一定步长绕 x 轴和 y 轴旋转一定角度（通常小于测量误差范围）后，以一定步长用 XOZ 平面分别沿 y 方向切割新生成的次棱坐标所构成的三维曲面空间。并通过得到的曲线判断擦痕线与基准平面是否平行。若不平行则继续重复以上步骤，直到判断二者达到平行。
3. 将步骤 2 得到的次棱坐标以一定步长绕 z 轴旋转一定角度（通常小于测量误差范围），以一定步长用 YOZ 平面沿 x 轴切割新生成的次棱坐标所构成的三维曲面空间，并通过得到的一簇曲线判断擦痕与 y 轴是否平行，若不平行则继续重复以上步骤，直到判断二者达到平行。

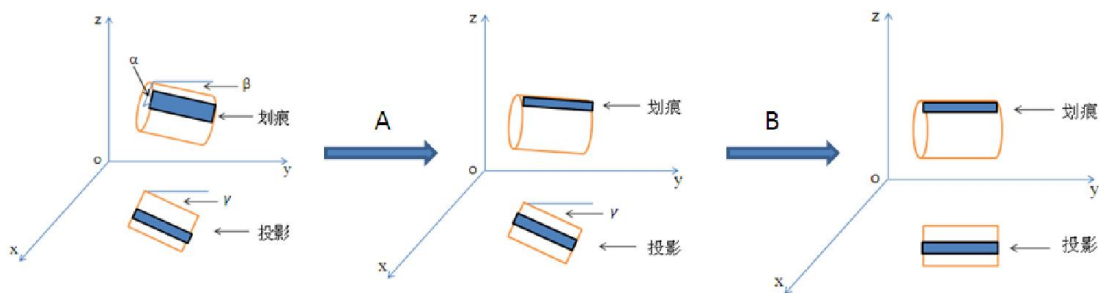


图 2：次棱坐标所构成的曲面变换示意图。图中 A 过程使弹头划痕平行于 XOY 平面（由步骤 B 实现）；B 过程使弹头划痕平行于 y 轴（由步骤 C 实现）。

考虑到实际情况中，由于擦痕处于弹头表面具有一定弧度，同时擦痕深浅可能不一。步骤 2，3 中的条件不易达到同时满足，然而通过以上方法可以尽量保证两个弹头在相同位置和相同姿态的情况下进行比对。

在上述步骤 2 和 3 中的坐标变换主要运用到齐次坐标变换方法。对于本问题，坐标绕 x, y, z 轴按一定角度旋转后生成的新坐标可由附录 1 中公式得到。

以下以问题 3 中（文件：6x2 次棱数据\1811345\77T1~1811345\c1）的数据为例说明上述方法的有效性和可行性。

对将该文件数据进行做图，得到如图 3 所示的次棱坐标所构成的曲面。从图中定性可以看出擦痕走向以及擦痕在次棱表面上的大概位置。

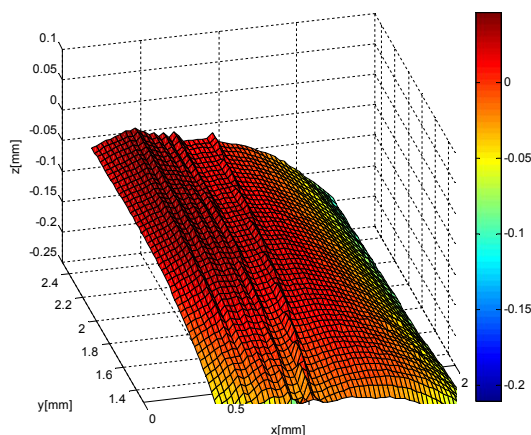


图 3：次棱坐标所构成的曲面图

首先用 XOZ 平面沿 y 方向以步长为 2.75 微米切割次棱坐标所构成的曲面空间，其中得到的部分切割曲线如图 4 所示。该图反映出空间变换前 XOZ 平面切割次棱线的平面变化图，系列走势有规律的曲线分布的变化可以反映出划痕在 y 轴和 x 轴方向的起始位置，同时划痕的宽度约为 0.5 毫米。划痕起始位置的斜率（如 y 为 2.0588 毫米对应的曲线斜率约为-0.0045）可知实际划痕平面略倾斜于基准 xoy 平面，即 α 角不为 0。

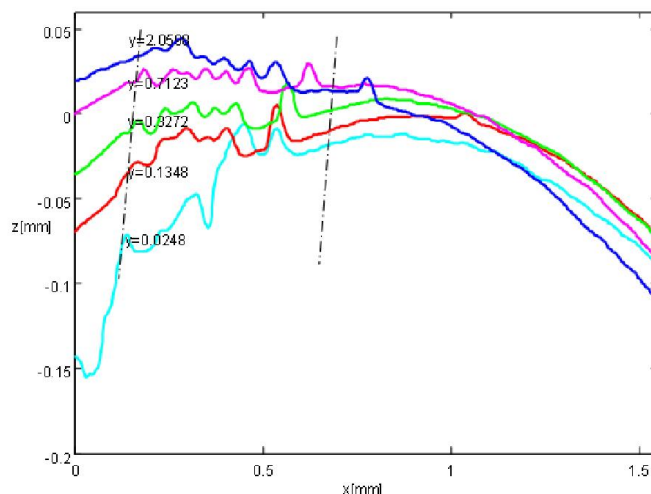


图 4: XOZ 平面沿 y 方向切割次棱坐标曲面空间得到的系列二维曲线图
(虚线标出擦痕范围)

图 5 为用 YOZ 平面沿 x 方向以步长为 2.75 微米切割次棱坐标曲面空间得到的部分一系列曲线变化图。该图反映出 x 较大时出现的系列有规律的 z 随着 y 的增大而递减的趋势，这表明测量的擦痕方向倾斜于 y 轴 (x 坐标为 1.0973 为斜率约为 -0.053)，即 β 和 γ 角不为 0，存在姿态误差。同时擦痕的凹凸峰在该图中也有清晰的反映 (如曲线中标注所示)。

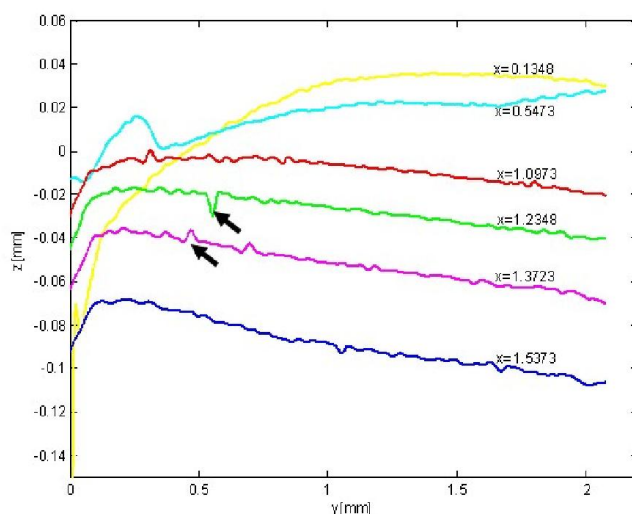


图 5: YOZ 平面沿 x 方向切割次棱坐标曲面空间得到的系列二维曲线图

以下分别将坐标分别绕 x 和 y 轴旋转进行坐标空间变换，使 α 和 β 角不断减小并趋向于 0，即将划痕平面平行于 XOY 基准平面。其中 α 角为 0 的严格判据为 XOZ 平面切割次棱线得到的划痕起始点对应坐标连线在 y 轴方向的导数为 0， β 角为 0 的严格判据为 YOZ 平面沿 x 方向切割次棱坐标得到的系列曲线中当 y 大于特定值时曲线的 z 值不随 y 值发生变化。通过程序搜索，得到空间坐标绕 x 轴旋转 0.0105 弧度 (0.602 度)，绕 y 轴旋转 0.0175 弧度 (1.003 度) 可使 α 和 β 角趋向于 0 (由于次棱坐标数值具有明显的波动性，不能使二者完全为 0，变化后图 6 的斜率为 0.000354 ($y=2.0588\text{mm}$) 几乎接近为 0；图 7 的斜率变化后约为 -0.0138

($x=1.0973\text{mm}$)。通过次棱曲面坐标变换后，用 XOZ 和 YOZ 平面切割其坐标空间曲线得到的两簇曲线图分别如图 6 和 7 所示。与坐标变换前的图 4 和图 5 相比，可以观察出擦痕附近曲线的变化情况。

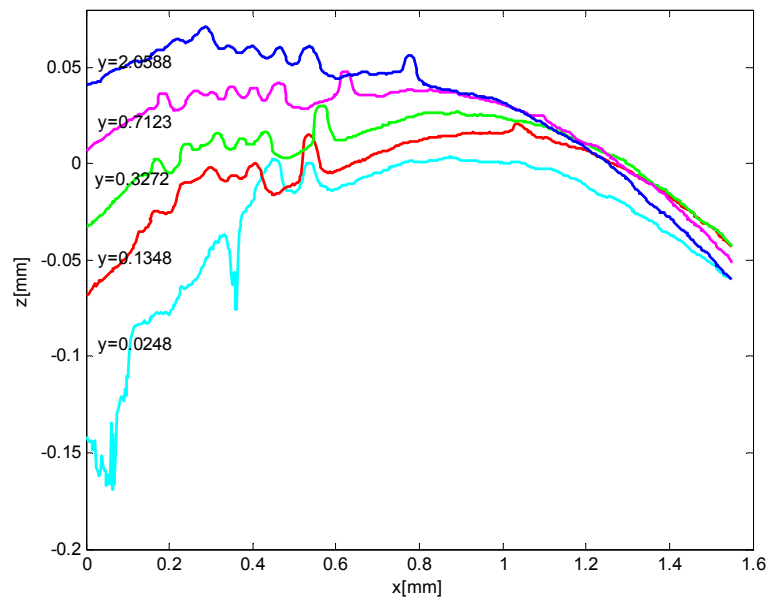


图 6: XOZ 平面切割坐标变换后次棱曲面空间得到的系列二维曲线图

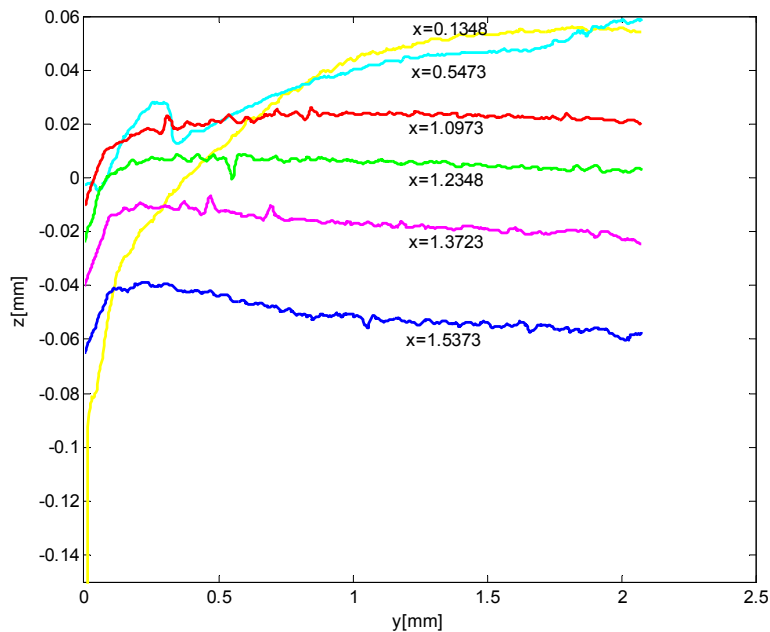


图 7: YOZ 平面切割坐标变换后次棱曲面空间得到的系列二维曲线图

以下通过将坐标空间绕 z 轴变化使得擦痕平面平行于 y 轴，使 γ 角趋向于 0。如图 8 所示，箭头为通过绕 z 轴旋转 2 度范围内的 y 轴切割空间坐标后得到曲线在 XOY 平面的方向导数变化情况。当擦痕偏离于 y 轴时， y 轴将跨过划痕而导致划痕峰处对应的曲线的导数存在较大梯度。当擦痕平行于 y 轴时， y 轴穿过曲面时得到的曲线的方向导数将变得十分平缓。

通过以上方式，可基本实现数据的校正。并消除转动和平动误差。

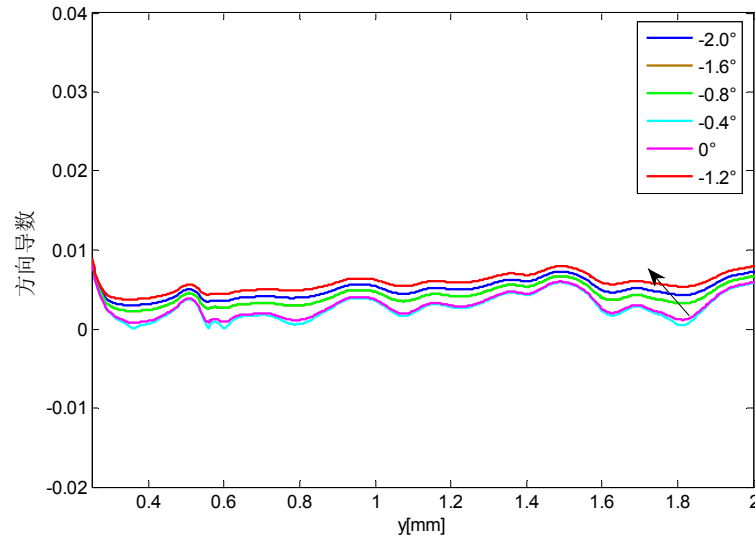


图 8: y 轴穿过旋转变换的次棱曲面空间得到的系列二维曲线图

问题 2 的回答与求解:

问题分析:

数据误差和噪声的来源和强弱具有不确定性和广泛性（如弹头表面的损伤、锈迹、油斑、杂质物及痕迹等）。文献中对于噪声的处理主要集中在弹头的图像处理等方面^[2,3]而通过数值方法进行消除弹头擦痕误差和噪声并不常见。从数学角度分析，消除误差和减小噪声的方法通常有平滑和滤波。其重要的特征是可以消除部分随机因素带来的噪声，同时强化特征信号。同时由于本题中由于擦痕空间坐标具有较强的实际性，我们考虑样条函数对此具有较大的灵活性和适应性，并利用其逼近真实的函数形式。

问题求解与实例:

将问题一中研究的子弹三维空间坐标数值进行三次样条函数滤波。得到的函数具有曲面光滑，二阶可导且连续的特点。该方法的主要思想是让样条函数和真实点之间的以下函数关系达到最小：

$$w \sum_{j=1}^n |y(j) - f(x(j))|^2 + (1-w) \int \left| \frac{d^2 f(x)}{dx^2} \right|^2 dx,$$

其中， $y(x)$ 和 $f(x)$ 为原函数和需要拟合的样条函数， $w = 1/(1 + h^3/m)$; $w \in [0,1]$ ， m 为一正数，通常小于 100。若 $w=0$ 则 $f(x)$ 还原为最小二乘直线拟合，若 $w=1$ 则 $f(x)$ 为三次样条差值函数。 h 为相邻数据点之间的距离。因此通过调节 w 的数值即可调节样条函数滤波程度，对于本问题，由于本题中 h 为 2.75 微米，导致 w 的数值趋近于 1。我们通过改变不同的 w 值来对比滤波效果的明显程度。在此，滤波要求达到的主要标准是要求滤掉用 XOZ 和 YOZ 平面切割次棱坐标曲面得到的曲线中较

大的噪音峰，同时擦痕峰本身体现的信息不能有较大丢失，该信息应该包含峰的位置，峰的形状和强度等。

以下举图 6 中 y 为 0.0248 毫米对应的曲线为例进行说明，该曲线和图 5 中其他系列曲线相比，其在 x 为 0 到 0.1 毫米之间，曲线具有较大的噪音峰。通过选取不同的 w 值 (w 分别取 0.9999, 0.99999 和 0.999999)，我们发现对于本问题涉及的曲线，滤波效果对 w 值具有较大的敏感性。通过与原始曲线图进行对比，选取 w 为 0.999999 作为样条函数滤波的基本参量。

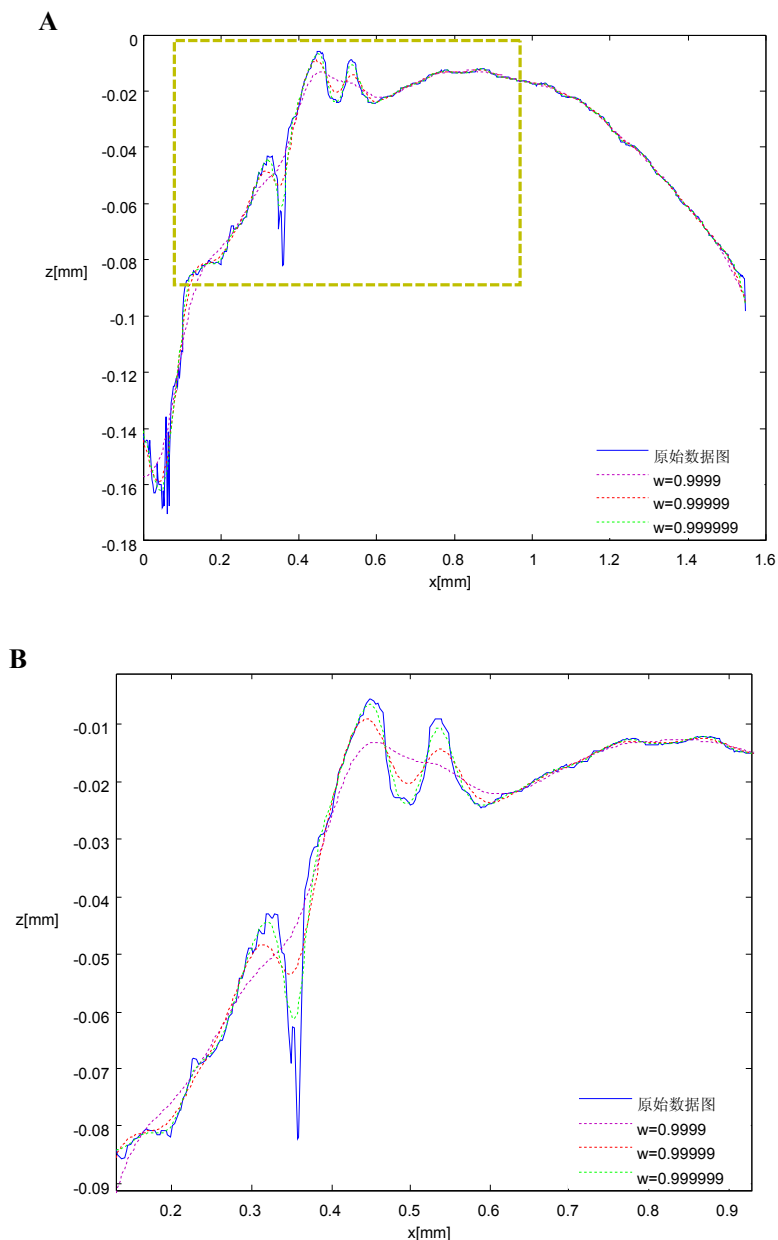


图 9: A: 不同 w 值进行滤波后的曲线与原始数据曲线的对比图; B: A 图中矩形区域放大后的对比图

问题 3 的回答与求解：

问题分析：

从自动比对方法的具体特征考虑，弹头擦痕的特征量需要具有可识别性强，不同枪发射的弹头之间差异尽量大的特点,同时还要兼顾操作过程的可行性和简易性。尽管传统的经验性判断和图像辨别方法总结出判断弹头擦痕的部分特征量，例如痕迹的起始位置，数量，面积，凹凸方向等^[2]，但是这些特征量在实际的提取过程中存在操作的复杂性和信息的不准确性（即容易收到噪音影响），同时受到人为因素的影响，使其难以实现智能识别，并造成较大的误差。

将擦痕识别问题联系互联网搜索技术中最常见的关键词搜索：通过输入几个关键词（类似于本题中的特征量）即可搜索与之相关或相匹配的信息。例如搜索信息度高度相关的文章或者比较不同论文之间的相似程度，我们最为直接的做法是比较二者之间关键词在文中出现的频率以及关键词之间的关系。

通过以上分析，我们在解决子弹头擦痕识别问题时，得到了如下启示：

1. 建立智能识别的方法或系统，提高操作的可行性或简易性有利于降低外界因素引起的不确定性。
2. 提取的特征量应易于定量表示，同时应尽量反映出经验中的特征要素体现的信息。特征量应该在原有信息基础上得到进一步强化，以消除误差和噪音。
3. 比较弹头划痕相关程度时应该重点考虑不同特征量出现的频率分布以及特征量之间的关系。

建立智能识别系统的前提是建立子弹擦痕样本信息的数据库。目前我们得到了 6 组不同枪发射的 12 个子弹的空间坐标作为原始信息样本，而每个弹头对应的次棱坐标数据量相对较大。通过问题 1 中的方法，运用 XOZ 和 YOZ 平面分别沿 y 方向和 x 方向切割次棱坐标的方法我们可以实现将三维数据信息转变为二维数据信息。而该数据信息可以充分反映擦痕的所有信息，用此方法的另外一个特点在于得到的擦痕附近的曲线具有较大的相似性（图 4 中可以直观分析出）。

基于强化擦痕信息以及统计擦痕特征出现频率的考虑，我们将不同步长得到的二维截面曲线（简称切片）向空间区域进行拓展相接，例如 XOZ 面截取的系列曲线在 x 轴方向进行依次相接，将此二维数据信息记为切片样本。例如图 10 的 A 和 B 分别为同一个弹头的两个不同的次棱面（c1 和 c3）所对应的切片样本。

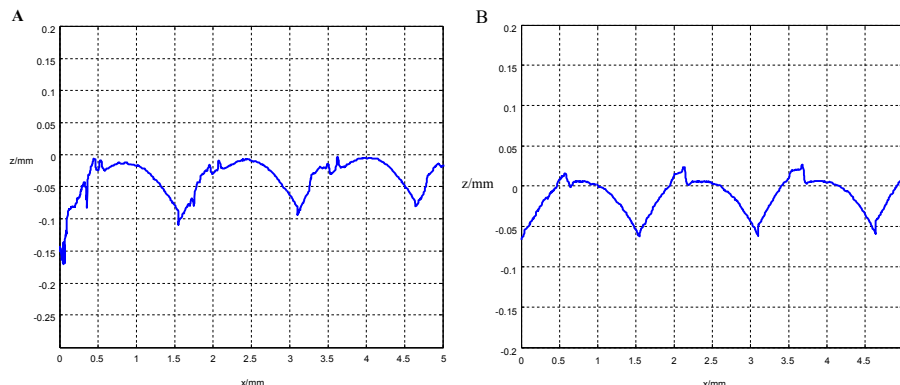


图 10：A：次棱面数据 c1 形成的切片样本；B：次棱面数据 c3 形成的切片样本（为直观只显示 3 个二维曲线连接的切片）

切片样本图较为直观的反映出划痕峰的变化情况，同时较好地保留了三维数据的信息。考虑到实际情况中，划痕处的所有峰（方向向下）的种类，峰形，峰面积分布等信息对于确定划痕具有重要作用，同时以上参数易于实现量化，符合前文中经验特征量的描述，将此类峰定义为划痕特征峰。以下通过 4 中参数描述划痕处的特征峰，同时通过确立特征峰类别，统计对比特征峰出现的种类，频率等信息建立比对方案的数学模型。

模型建立：

1. 特征峰的参数描述：

在切片样本中用以下 4 个参量来描述每个划痕的特征峰，如图 11 所示，这 4 个参数分别包括：

1. θ 角：线段 AC 与 BC 之间的夹角，用于反映划痕形貌特征的主要参数。
2. 线段 AC 长度 a 与 BC 长度 b 之间的比值 c (a, b 具有左右方向性)：描述特征峰形状的主要参数。
3. 三角形 ABC 的面积 S ：表征次棱擦痕深度，宽度的主要参数。与枪膛的特性具有重要相关性。
4. 曲边三角形 ABC 的面积 S' ：用于描述擦痕深度，凹凸性的主要参数。

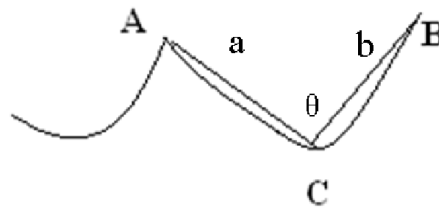


图 11：二维切片样本空间中定义的特征峰示意图

2. 特征峰的提取与分类：

在一个子弹次棱的切片样本空间中存在多个特征峰，当前，我们的目的是通过提取并统计特征峰的 4 个参数以有效实现特征峰的分类，并将不同类别的特征峰进行统计，通过不同类别特征峰出现的频率分布进一步判别两个不同次棱的相似程度。

步骤 1： 对一个次棱切片样本空间沿着水平方向以一个标尺长度（可伸缩）进行搜索。搜索的目标是得到所有沿着该方向的存在的特征峰。

搜索特征峰首先需要寻找一段搜索范围的极小值点 C 点（最小值为端点的不记为最低点），此时将切片样本按问题 2 中的方法进行滤波，以消除某些非极小值点的干扰，然后从极小点 C 出发向其左右进行搜索，搜索到其附近斜率变化为 0 或实际情况中变化范围很小的点（图 11 的 A 和 B 点）。至此完成一个特征峰的搜索，进一步求得该特征峰的 4 个参数，并计为第 1 类特征峰 A。

步骤 2： 继续沿着水平方向进行标尺搜索，并重复以上步骤，对特征峰进行搜索，并计算新搜索得到的特征峰的 4 个参数指标，判断这 4 个参数指标是否都在第一个特征峰 4 个指标 θ, c, S, S' 的 $\pm m\%$ （通常取 $m \leq 10$ ）的范围内，若是，则将

该特征峰记为第 1 类特征峰 A。反之，则将其新建一类特征峰 B，归为第二类特征峰。

步骤 3: 依次分别对该切片样本空间进行上述标尺搜索 n 次 (n 通常由标尺长度决定)，并按照步骤 2 的规则依次得到若干不同类别的特征峰，特征峰 4 参数以及各类特征峰对应的出现频率，并进行数据存储，建立该样本的数据库。

至此某次棱面的特征信息将全部体现在其提取的特征峰种类及其频率分布上。此思路可类比于互联网搜索引擎中用若干重要关键词来代替网络文章的搜索和匹配模式。

以下我们简要以 c1 特征峰搜索过程中为例进一步解释特征峰提取过程中由于峰形的差异可能出现的几种情形，以对上述步骤中 A,B,C 三点的确定进行补充说明，其中搜索尺度选取 50 个基本步长（基本步长为 2.75 微米），极小值搜索过程中可能出现以下情况：1.若最小点出现在端点处，则程序将不记录该最小点（图 12A 所示），此时进一步进行标尺搜索则可以搜索出 x 约为 0.35mm 处的特征峰极小值点 C 并进一步可以搜索出斜率为 0 的点 A 和 B(图 12B 所示)。图 12C 反映出搜索到的 B,C 两点处斜率为 0。图 12D 为反映出 B 点搜索过程中其斜率不为 0 但变化很小的情况。

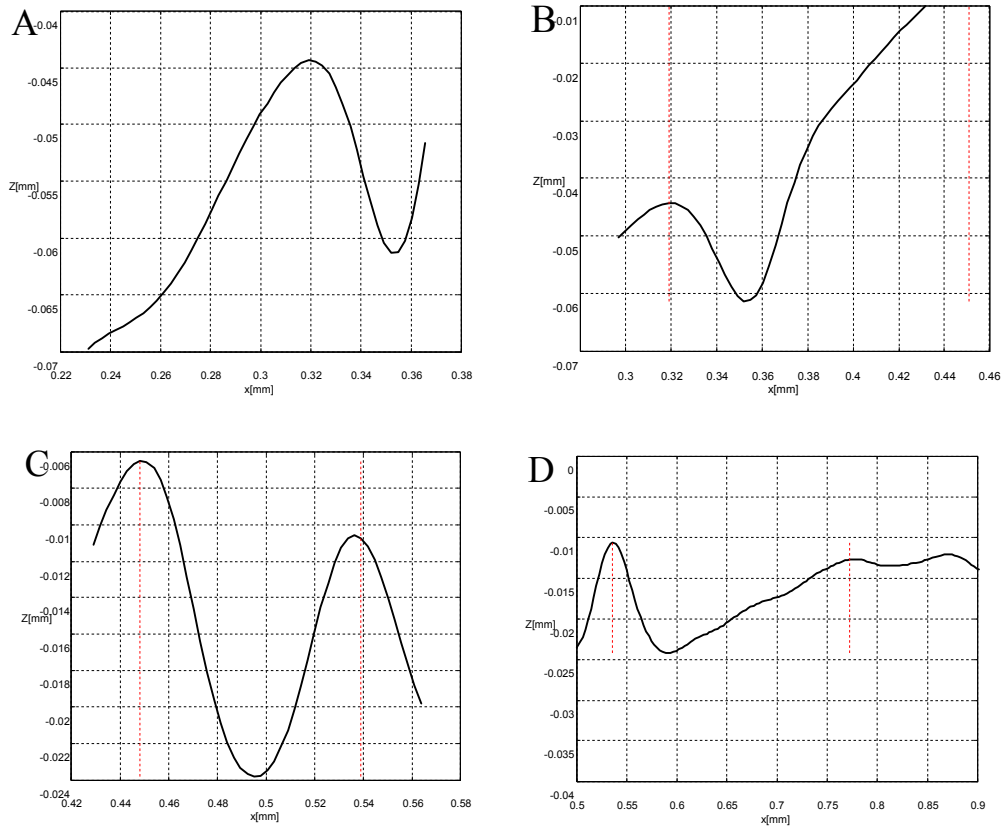


图 12：特征峰提取过程中可能出现的几种情形

3. 不同次棱面之间相似程度（或匹配关系）的确定

当比较任意两个次棱面之间的相关关系时，我们只需比较次棱面之间各特征峰的个数和频率分布即可。在两个次棱之间第 i 类特征峰相同的的总数记为 N_i ，

而互不相同的特征峰总数记为 N_2 , 则定义两个次棱面第 i 类特征峰的

$r_i = N_i / (N_1 + N_2); r_i \in [0, 1]$ 为两个次棱面之间的相似程度, 反映了二者对第 i 类特征峰的匹配关系。综合考虑所有 m 类特征峰的因素, 两个特征峰之间的相似程度为

$r = \frac{\sum_{i=1}^m r_i}{m}$ 。在理想情况下, 信息完全一致的两个次棱将出现 $r=1$ 的情形, 而在实际情

况中由于误差和噪音的干扰, r 值可能与 1 偏离较远。

对于本题而言, 我们选取问题 3 中 1811345 文件中 t1 枪的 c1 和 c3 两个次棱面为例说明实际操作中的具体处理。通过分析次棱面特征峰类别的数据我们发现除了较多特征峰数目分为一类以外, 还有许多单独分为一类的特征峰数目只有 1 个或很少, 因此考虑到可能是噪声或在滤波处理过程中产生的误差造成数目很少的特征峰归为一类。同时由于主要特征分对次棱面的影响最大, 同时不易受到外界噪声的影响。因此我们只考虑 c1 和 c3 次棱面影响最大的前 5 个特征峰, 即特征峰出现最多的种类。以 c1 按从大到小排序为标准计算得到的二者的不同类别的前 15 种特征峰的数目如表 1 所示。首先分别对 c1 和 c3 按照特征峰数目从大到小进行排序并分别以 c1 和 c3 为标准计算二者的 r 值。以 c1 按从大到小排序为标准计算得到的二者的相似程度 $r_{13}=0.7717$ 。以 c3 按从大到小排序为标准计算得到的二者的相似程度 $r_{31}=0.2946$ 。最后通过二者的代数平均求得这两个次棱面之间的相似程度为 0.53315。

表 1: 以 C1 次棱面由大到小排序的 C1 与 C3 次棱面前 15 个重要的特征峰数目表

特征峰类别	C1次棱面	C3次棱面	特征峰类别	C1次棱面	C3次棱面	特征峰类别	C1次棱面	C3次棱面
A	20	16	F	11	7	K	8	4
B	14	53	G	11	3	L	7	14
C	14	15	H	10	25	M	7	4
D	13	27	I	10	5	N	7	0
E	13	2	J	10	0	O	6	1

上述问题中, 计算两者之间的相似程度同时可以考虑其他方法, 如常见的模糊聚类分析(如马氏距离方法, 向量投影方法)和神经网络算法等^[5]。例如在本题中由于某个次棱面的不同某种特殊类别的特征向量个数已经确定, 因此可以构造两类次棱面的向量组, 同时选择两个向量之间的夹角余弦(即向量在空间中的投影)作为衡量二者之间相似程度。余弦夹角大表明向量之间相似程度越高, 反之则表明向量之间的相似程度越低。如在上述 c1 和 c3 的相似程度计算中, 按照 c1 按从大到小排序, 并计算向量夹角余弦可以得到二者的相似程度为 0.7734。这与代数运算计算方法十分接近, 从实际的可操作性和简单性原则考虑, 以下我们计算相似程度都统一采用代数运算方法。

4. 子弹头之间相关关系的确定

由于子弹头包含 4 个次棱面，同时 4 个次棱面之间的相对顺序固定。因此我们在比较两个不同的子弹头之间的相似程度时就有四种可能的排序。

设 P 和 Q 为任意两个待比较的两个子弹头。其中固定子弹头 P 的 4 个次棱面的逆时针排序为 $[p_1, p_2, p_3, p_4]$ (p_1 定义为与采集数据光源最近的次棱面)；同时改变子弹头 Q 的次棱面顺序，分别与 P 弹头的固定状态对比。4 种排序方式每一种将对应 4 个次棱面的相似程度，四中排序可构造如下矩阵 R 以记录 4 情况各棱面的匹配程度,其中 $r_{p_i q_j}$ 代表 P 弹头的 i 棱面与 Q 弹头的 j 次棱面的相关程度 ($i, j=1, 2, 3, 4$)。同时通过判别不同 4 种不同排序的所对应的次棱面的整体相关性矩阵

记为 R' (定义为 R 矩阵列的方向加和除以 4) 的大小还可以分析出不同子弹所对应的次棱面的排列顺序，显然某种排序对应的四个相似程度整体加和越大，则为最有可能的次棱排列顺序。利用 r 值比较匹配相关度大小时，若出现 r 值相等或者近似的情况下，则比较 r 的方差大小。方差小者为最佳匹配方式。

$$R = \begin{bmatrix} r_{p_1 q_1} & r_{p_1 q_2} & r_{p_1 q_3} & r_{p_1 q_4} \\ r_{p_2 q_1} & r_{p_2 q_2} & r_{p_2 q_3} & r_{p_2 q_4} \\ r_{p_3 q_1} & r_{p_3 q_2} & r_{p_3 q_3} & r_{p_3 q_4} \\ r_{p_4 q_1} & r_{p_4 q_2} & r_{p_4 q_3} & r_{p_4 q_4} \end{bmatrix}$$

我们以问题 3 中给出的 6 支枪分别发出的两枚子弹为例，检验上述方法的可行性， $R_{1811345}$ 为 1811345 号枪的发出的两枚子弹的相似程度矩阵。

$$R_{1811345} = \begin{bmatrix} 0.454038 & 0.280588 & 0.766448 & 0.877926 \\ 0.782578 & 0.997333 & 0.859922 & 0.553884 \\ 0.274723 & 0.480912 & 0.362577 & 0.926306 \\ 0.648766 & 0.728070 & 0.364852 & 0.753012 \end{bmatrix}$$

以下通过代数平均运算可以得到不同顺序的整体相似程度：

$$R'_{1811345} = [0.540026; 0.621726; 0.588450; 0.777782]$$

从整体相似程度分析，对于 1811345 号枪的发出的两枚子弹而言，次棱的排列顺序最为可能的是：77t1-1811345 号子弹的 $c1 \sim c2 \sim c3 \sim c4$ 顺序对应 77t2-1811345 号子弹的顺序为 $c4 \sim c1 \sim c2 \sim c3$ 。从图 13 也可以比较直观地看出：77t1-1811345 号子弹 $c1$ 图与 77t2-1811345 号子弹的 $c1 \sim c4$ 相比，其与 $c4$ 的相似程度较高（相似程度约为 0.878），而与 $c1$ 相似程度较低（相似程度约为 0.878）。77t1-1811345 号子弹的 $c2$ 次棱面与 77t2-1811345 的 $c2$ 次棱面从图像中分辨较差，这与较低的相似程度（约为 0.554）很低的结果是十分吻合的。以上可以从经验的角度侧面反映出模型的可行性和合理性。

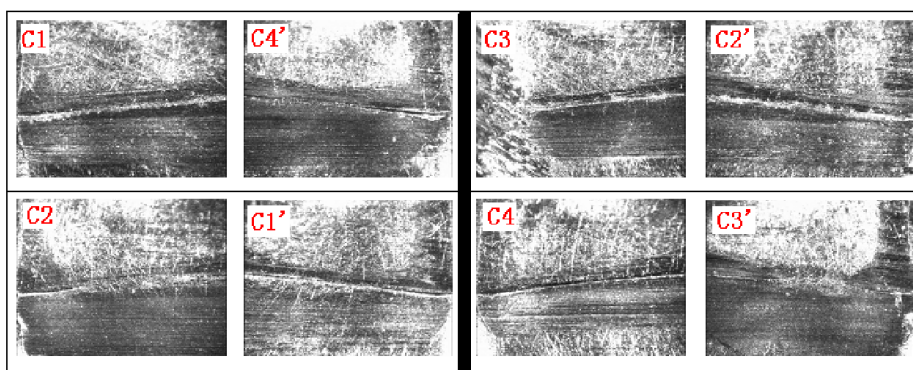


图 13：1811345 号枪的两枚子弹按计算得到的结果排序的光学显微镜照片对比图（照片具有镜像对应关系）

同时按照上述方法，问题 3 中其他 5 支枪分别发出的两枚子弹的相似程度矩阵 R ，整体相似程度矩阵 R' 和子弹次棱的一一对应关系也可以分别计算得到，如表 2 所示。

表 2：问题 3 中其他 5 支手枪发射子弹的相似程度矩阵 R ， R' 以及次棱对应顺序

序号	两枚子弹的相似程度矩阵 R	整体相似程度矩阵 R' ; 子弹 t_1 对应 t_2 的次棱对应顺序
2	$R_{1812492} = \begin{bmatrix} 0.421761 & 0.496201 & 0.698060 & 0.683006 \\ 0.823460 & 0.638709 & 0.731182 & 0.354342 \\ 0.372180 & 0.451887 & 0.632642 & 0.319890 \\ 0.286195 & 0.413194 & 0.270833 & 0.314285 \end{bmatrix}$	$R'_{1812492} = [0.475899; 0.499998; 0.583179; 0.417881]$ (t_1) $c_1 \sim c_2 \sim c_3 \sim c_4$: (t_2) $c_3 \sim c_4 \sim c_1 \sim c_2$
3	$R_{1504519} = \begin{bmatrix} 0.947871 & 0.544796 & 0.870383 & 0.323789 \\ 0.918110 & 0.917145 & 0.496942 & 0.660959 \\ 0.472425 & 0.583101 & 0.415661 & 0.808260 \\ 0.610363 & 0.875992 & 0.707349 & 0.972056 \end{bmatrix}$	$R'_{1504519} = [0.737192; 0.730259; 0.622584; 0.691267]$ (t_1) $c_1 \sim c_2 \sim c_3 \sim c_4$: (t_2) $c_1 \sim c_2 \sim c_3 \sim c_4$
4	$R_{1203959} = \begin{bmatrix} 0.622340 & 0.509361 & 0.342062 & 0.458769 \\ 0.529177 & 0.932748 & 0.778448 & 0.183270 \\ 0.912280 & 0.987179 & 0.181114 & 0.303177 \\ 0.743869 & 0.698950 & 0.771696 & 0.526188 \end{bmatrix}$	$R'_{1203959} = [0.701917; 0.782059; 0.518330; 0.367851]$ (t_1) $c_1 \sim c_2 \sim c_3 \sim c_4$: (t_2) $c_2 \sim c_3 \sim c_4 \sim c_1$
5	$R_{1923252} = \begin{bmatrix} 0.935526 & 0.741455 & 0.507233 & 0.381642 \\ 0.699530 & 0.900720 & 0.275881 & 0.714093 \\ 0.277448 & 0.323285 & 0.947181 & 0.632047 \\ 0.420731 & 0.785333 & 0.717029 & 0.977845 \end{bmatrix}$	$R'_{1923252} = [0.583309; 0.687698; 0.611831; 0.676407]$ (t_1) $c_1 \sim c_2 \sim c_3 \sim c_4$: (t_2) $c_2 \sim c_3 \sim c_4 \sim c_1$
6	$R_{1928033} = \begin{bmatrix} 0.839646 & 0.467914 & 0.379089 & 0.278409 \\ 0.478260 & 0.704686 & 0.354820 & 0.289855 \\ 0.448005 & 0.465183 & 0.319519 & 0.638198 \\ 0.278132 & 0.667844 & 0.510159 & 0.440567 \end{bmatrix}$	$R'_{1928033} = [0.511011; 0.576407; 0.390897; 0.411757]$ (t_1) $c_1 \sim c_2 \sim c_3 \sim c_4$: (t_2) $c_2 \sim c_3 \sim c_4 \sim c_1$

为了较好地证明方法的有效性，我们分别选取了 6 组不同枪发出的 2 个子弹形成 6 组组合并做相似程度矩阵 R 分析，得到的不同类别子弹的整体相似程度矩阵 R' 向量中的最大值如表 3 所示，其值一般分布在 0.44~0.56 之间。相比于同类枪发射的子弹相似程度较低。

表 3：不同枪发出子弹的最大相似程度值

不同子弹编号	77t1-1811345	77t1-1812492	77t1-1504519	77t2-1504519	77t2-1812492	77t2-1811345
	77t1-1928033	77t1-1923252	77t2-1203959	77t1-1203959	77t2-1923252	77t2-1928033
R' 矩阵最大值	0.518796	0.463473	0.511895	0.444141	0.478129	0.557642

5. 小结:

通过分析以上 6 支枪分别发出的两枚相同子弹的相似程度矩阵 R 以及整体相似程度数矩阵 R' 我们可以得到以下启示:

1. 对于两枚相同子弹而言, 最终的排序所分别对应的次棱面间的相关顺序有 2~3 个相似程度较高, 例如第 3, 4 支枪发出子弹的排序中有 2 个次棱面相似程度都达到 0.9 以上; 同时第 1~5 支枪发出子弹的排序中有 2 或 3 个次棱面相似程度都达到 0.7 或以上。这与同一支枪发射的两个弹头通常有 4 个次棱中通常可以有 2 至 3 个吻合比较好的经验规律十分吻合。
2. 从整体相似程度矩阵 R' 的结果看, 两枚子弹顺序匹配度最好的排列方式对应的相似程度达到 0.75 以上的有 2 对; 0.65 以上的有 4 对; 同时 6 组相似程度全部在 0.55 以上, 考虑到前文分析任意选取的两组次棱面的相似程度大约为 0.53。因此该方法在很大程度上将同一类子弹与不同类子弹进行有效分离。然而可以看出第 2 和第 6 组数据处于 0.55~0.60 之间相对其他同类弹头最大相似程度较低。此时通常任意两个截面的区分效果并不十分明显, 造成这样的原因可归为不同的次棱面的匹配相关度对弹头整体的贡献有一定的差别。但该结果仍然高于不同类弹头的最大相似程度值。
3. 对比同类和不同类的子弹头的相似程度值, 我们认为相似程度为 0.56 可以近似认为同类和不同类的一个分界的标准。

问题 4 的回答与求解:

对于本问题, 我们通过问题 3 中类似的方案对 22 个弹头 ($t_1 \sim t_{22}$) 进行了相关关系的计算, 表 4 列出了一部分 ($t_0 \sim t_7$) 之间的两两相似程度, 其他结果详见附录 2。

表 4: $t_0 \sim t_7$ 子弹两两相似程度表

相似程度	t_1	t_2	t_3	t_4	t_5	t_6	t_7
t_1	1.000	0.487	0.663	0.597	0.680	0.491	0.741
t_2	0.487	1.000	0.454	0.492	0.509	0.502	0.508
t_3	0.663	0.454	1.000	0.501	0.557	0.432	0.573
t_4	0.597	0.492	0.501	1.000	0.548	0.533	0.586
t_5	0.680	0.509	0.557	0.548	1.000	0.473	0.766
t_6	0.491	0.502	0.432	0.533	0.473	1.000	0.562
t_7	0.741	0.508	0.573	0.586	0.766	0.562	1.000

基于上述结果, 我们给出每个弹头按照相似度由高到低与之相似的前 5 个弹头, 同时我们给出与之相如表 5 所示:

表 5: 每个弹头按照相似程度由高到低前五排序表

子弹标号	按由高到低的相似程度排序前 5 个弹头
t_1	t_{18} ; t_7 ; t_{13} ; t_{22} ; t_{11}
t_2	t_{12} ; t_{10} ; t_{18} ; t_{19} ; t_{13}

t3	t22; t1; t17; t20; t14
t4	t21; t15; t9; t1; t17
t5	t7; t13; t22; t18; t1
t6	t12; t16; t10; t21; t18
t7	t13; t5; t11; t1; t18
t8	t1; t20; t22; t18; t13
t9	t20; t10; t13; t22; t1
t10	t9; t16; t7; t13; t20
t11	t7; t13; t22; t1; t20
t12	t6; t19; t2; t7; t1
t13	t22; t7; t20; t5; t1
t14	t13; t7; t16; t21; t22
t15	t18; t1; t22; t4; t5
t16	t14; t22; t10; t21; t5
t17	t7; t22; t1; t13; t5
t18	t22; t1; t7; t15; t5
t19	t12; t21; t3; t6; t14
t20	t13; t9; t22; t7; t11
t21	t4; t14; t16; t13; t7
t22	t13; t18; t5; t11; t20

通过以上表格我们可以发现并得到以下一些规律：

1. 由题意可知，上表中的弹头是分别由 11 支不同枪发出，即最为理想的匹配情况是分别有 11 组子弹两两互相优先识别并完成配对，但是实际计算结果中这种情况只发生在 t4 和 t21；t6 和 t12；t13 和 t22 这 3 组。出现这种情况的原因可能是由于实际情况存在的噪声和误差，或不同枪发出的子弹本身存在某种相似性。
2. 有某些子弹头与其他几个子弹头的匹配程度都比较高。例如 t13 和 t22，这是一组互相匹配识别的弹头，它们与其他弹头（例如 t20, t13）之间较高的识别可能是由于其特征峰的分布具有一定普适性。

若我们假设互相识别的两类弹头是同一支枪发射出的，则排除了上述 6 个弹头与其他弹头互配的可能性，若这 6 个弹头与其他弹头相似程度开始排序最高，则此时对应排序第二位的可以自动成为相似程度最高的子弹，其他排位则相应升高一位。此时，我们可以找出另外 3 对互相识别的子弹为 t1 和 t18；t5 和 t7；t9 和 t20；依此，继续排除这 6 枚子弹与其他互配的可能，同时相应的排位继续上升。此时可找出 t14 和 t16；t8 和 t15 两组匹配。依照上面的思路我们可以继续得到剩余的子弹匹配分别为 t3 和 t17；t2 和 t19；t10 和 t11。因此以上 22 枚子弹均可以在基本假设条件下实现两两配对。然而，在实际操作中，会出现相似程度系数差别较大的在最终筛选中实现匹配（例如具体排序中 t8 和 t15 相差较远），我们认为这样的匹配可能会与实际情况有些偏差，但是相似程度较高的子弹两两配对判别出于同一支枪的可信度是比较高的。

我们分别对比了以上配对的 11 组匹配结果的相似程度，发现匹配的相似程度仅 3 组在 0.6 以下（其中包括最后通过筛选进行组合的 2 组）。这一点与我们在问题三中发现的相同类别子弹相似程度大于 0.56 的判据是基本一致的。

模型优缺点分析:

模型优点:

1. 将 3 维数据原始样本通过“切片”展开方法转变为二维的切片样本,较好得保持了原有数据的信息,同时降低了数据存储的复杂性和繁琐性。运用样条函数滤波的方法,通过参数调变,保留了原始数据峰的特征,同时去除了噪音和误差的干扰。
2. 通过类比互联网技术中的关键词搜索引擎,对弹头的擦痕对比问题以 4 个指标量化特征峰,并以特征峰出现的次数和频率分布作为任意两个次棱面比较的依据,提取了两棱面相关的有利信息,抓住了问题的关键。
3. 通过特征峰识别以及研究 6 组相同弹头与不同弹头之间的相似程度之间的差异,我们找出了相似程度为 0.56 可以近似看做同类和不同类弹头的一个评价的标准,并予以验证。

模型缺点:

为简化模型形式,在计算任意两个次棱面的相似程度时,没有过多采用其他现有方法(如模糊聚类等)进行与我们现有的方法对比,这在一定程度上可能会导致了两棱面间相关关系存在一定误差。

对于以上不足,我们将在以后进行更加深入的研究和讨论。

模型延伸与展望:

本论文中的模型从庞大的三维数据样本中通过切片技术提取了关键特征峰,并通过特征峰的个数,频率,分布作为判断次棱面,并进一步识别子弹头的依据。我们认为在此过程中值得深入挖掘和研究的内容包括:

1. 不断建立和扩充弹头识别次棱的特征峰种类及其分布的数据库,有利于弹头痕迹辨别方法由经验性向自动性和智能识别系统的跨越。在本文中我们引入了 4 个参数以量化特征峰种类。由于特征峰的判断标准不同,可能其种类分布也会出现差别。因此我们认为可以通过进一步细化特征峰的量化参数可能会使基于特征峰识别的弹头痕迹比对方法更加精确。
2. 研究和总结不同弹头的特征峰种类的概率分布,以及不同的搜索特征峰的方式(例如进一步开发文中搜索特征峰标尺的可伸缩性和步长的可调性)可以进一步充分了解弹头擦痕的形貌特征。可能对特征峰的分类提供更好的依据。

参考文献:

- [1]杨敏,牟丽,弹头痕迹多尺度扩展分形检验分析,6,677-681,2009.
- [2]潘自勤,对枪弹弹头痕迹特征自动检索技术的探讨,6,29-31,2003.
- [3]鲁静,李德华,吴士泓,基于图像纹理分布的算法在子弹自动比对中的应用,武汉

大学学报(理学版), 53, 543-547, 2007.

[4]李德仲, 李安国.枪弹痕迹检验技术.北京: 警察教育出版社, 1996.

[5]姜启源, 谢金星, 叶俊编著. 数学模型 (第三版), 北京: 高等教育出版社, 2003.

附录:

附录 1: 在右手坐标系下相对坐标系原点绕坐标轴旋转 θ 角的变换公式:

将三维 $[x \ y \ z]$ 拓展为四维 $[x \ y \ z \ 1]$

(1)绕 x 轴旋转

$$\begin{bmatrix} x^* & y^* & z^* & 1 \end{bmatrix} = \begin{bmatrix} x & y & z & 1 \end{bmatrix} \bullet \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & \sin \theta & 0 \\ 0 & -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

(1)绕 y 轴旋转

$$\begin{bmatrix} x^* & y^* & z^* & 1 \end{bmatrix} = \begin{bmatrix} x & y & z & 1 \end{bmatrix} \bullet \begin{bmatrix} \cos \theta & 0 & \sin \theta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \theta & 0 & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

(1)绕 z 轴旋转

$$\begin{bmatrix} x^* & y^* & z^* & 1 \end{bmatrix} = \begin{bmatrix} x & y & z & 1 \end{bmatrix} \bullet \begin{bmatrix} \cos \theta & \sin \theta & 0 & 0 \\ -\sin \theta & \cos \theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

附录 2: 问题 4 中 t1~t22 的两两之间相关程度表

相似程度	t1	t2	t3	t4	t5	t6	t7	t8
t1	1.000	0.487	0.663	0.597	0.680	0.491	0.741	0.677
t2	0.487	1.000	0.454	0.492	0.509	0.502	0.508	0.444
t3	0.663	0.454	1.000	0.501	0.557	0.432	0.573	0.581
t4	0.597	0.492	0.501	1.000	0.548	0.533	0.586	0.552
t5	0.680	0.509	0.557	0.548	1.000	0.473	0.766	0.620
y6	0.491	0.502	0.432	0.533	0.473	1.000	0.562	0.472
t7	0.741	0.508	0.573	0.586	0.766	0.562	1.000	0.590
t8	0.677	0.444	0.581	0.552	0.620	0.472	0.590	1.000
t9	0.657	0.482	0.621	0.597	0.655	0.467	0.592	0.614
t10	0.567	0.582	0.616	0.528	0.631	0.604	0.658	0.549
t11	0.690	0.498	0.615	0.468	0.651	0.453	0.745	0.588
t12	0.571	0.610	0.410	0.534	0.438	0.675	0.572	0.526
t13	0.721	0.518	0.590	0.589	0.742	0.472	0.790	0.636
t14	0.577	0.492	0.639	0.559	0.637	0.542	0.689	0.541
t15	0.679	0.428	0.583	0.655	0.649	0.452	0.630	0.595

t16	0.619	0.504	0.570	0.549	0.656	0.635	0.566	0.559
t17	0.680	0.498	0.652	0.597	0.654	0.512	0.690	0.541
t18	0.743	0.536	0.584	0.540	0.685	0.573	0.738	0.642
t19	0.438	0.525	0.571	0.520	0.555	0.568	0.529	0.456
t20	0.659	0.512	0.645	0.557	0.660	0.423	0.689	0.673
t21	0.609	0.473	0.573	0.667	0.604	0.601	0.627	0.537
t22	0.696	0.449	0.689	0.525	0.729	0.492	0.682	0.667
相似程度	t9	t10	t11	t12	t13	t14	t15	t16
t1	0.657	0.567	0.690	0.571	0.721	0.577	0.679	0.619
t2	0.482	0.582	0.498	0.610	0.518	0.492	0.428	0.504
t3	0.621	0.616	0.615	0.410	0.590	0.639	0.583	0.570
t4	0.597	0.528	0.468	0.534	0.589	0.559	0.655	0.549
t5	0.655	0.631	0.651	0.438	0.742	0.637	0.649	0.656
t6	0.467	0.604	0.453	0.675	0.472	0.542	0.452	0.635
t7	0.592	0.658	0.745	0.572	0.790	0.689	0.630	0.566
t8	0.614	0.549	0.588	0.526	0.636	0.541	0.595	0.559
t9	1.000	0.715	0.599	0.389	0.702	0.609	0.567	0.596
t10	0.715	1.000	0.556	0.471	0.635	0.586	0.590	0.673
t11	0.599	0.556	1.000	0.439	0.720	0.595	0.579	0.623
t12	0.389	0.471	0.439	1.000	0.453	0.539	0.441	0.550
t13	0.702	0.635	0.720	0.453	1.000	0.701	0.617	0.557
t14	0.609	0.586	0.595	0.539	0.701	1.000	0.596	0.684
t15	0.567	0.590	0.579	0.441	0.617	0.596	1.000	0.585
t16	0.596	0.673	0.623	0.550	0.557	0.684	0.585	1.000
t17	0.544	0.544	0.649	0.506	0.679	0.589	0.594	0.566
t18	0.641	0.633	0.628	0.549	0.669	0.638	0.719	0.651
t19	0.477	0.537	0.498	0.613	0.440	0.556	0.491	0.545
t20	0.718	0.633	0.689	0.483	0.749	0.606	0.595	0.582
t21	0.539	0.561	0.576	0.557	0.645	0.666	0.597	0.660
t22	0.681	0.613	0.715	0.522	0.801	0.643	0.670	0.683
相关程度	t17	t18	t19	t20	t21	t22		
t1	0.680	0.743	0.438	0.659	0.609	0.696		
t2	0.498	0.536	0.525	0.512	0.473	0.449		
t3	0.652	0.584	0.571	0.645	0.573	0.689		
t4	0.597	0.540	0.520	0.557	0.667	0.525		
t5	0.654	0.685	0.555	0.660	0.604	0.729		
t6	0.512	0.573	0.568	0.423	0.601	0.492		
t7	0.690	0.738	0.529	0.689	0.627	0.682		
t8	0.541	0.642	0.456	0.673	0.537	0.667		
t9	0.544	0.641	0.477	0.718	0.539	0.681		
t10	0.544	0.633	0.537	0.633	0.561	0.613		
t11	0.649	0.628	0.498	0.689	0.576	0.715		
t12	0.506	0.549	0.613	0.483	0.557	0.522		

t13	0.679	0.669	0.440	0.749	0.645	0.801
t14	0.589	0.638	0.556	0.606	0.666	0.643
t15	0.594	0.719	0.491	0.595	0.597	0.670
t16	0.566	0.651	0.545	0.582	0.660	0.683
t17	1.000	0.618	0.528	0.638	0.622	0.690
t18	0.618	1.000	0.437	0.673	0.574	0.774
t19	0.528	0.437	1.000	0.548	0.587	0.498
t20	0.638	0.673	0.548	1.000	0.597	0.712
t21	0.622	0.574	0.587	0.597	1.000	0.556
t22	0.690	0.774	0.498	0.712	0.556	1.000

附录 3：相关程序的电子文档说明：

1. Main.m

本程序为主程序，从原始数据出发，生成切片二维数据，建立核心搜索算法，逐一扫描整个切片图形，提取特征峰。存储检索索引信息（即全部的特征峰信息）。

2. Mpipei.m

子弹匹配程序. 对第 4 题，每两颗子弹对比，每个子弹的任意两两面匹配 输出 resD4。

3. Mshuchu.m

对两两子弹之间的特征峰匹配数据分析，得出最终的匹配度. 对信息进行处理，最后生成一个 $N \times N$ 的矩阵，第 (m, n) 个元素代表第 m 个子弹和第 n 个子弹的匹配程度。最后输出每个子弹前 5 个最匹配的子弹及匹配概率。

4. Mjihebianhuan.m

这是对原始数据做几何变换的程序，后面发现输入面的空间位置偏差对我们设计的搜索算法不影响。所以并不需要几何变化，即可处理各子弹数据。

5. resdata.dat

这是存储数据库，由于第四题的所有子弹信息为 36m，文件过大。这里只保留了第一条信息。注意，运行程序需要完全信息的 resdata。

6. resD4.dat

这是存储第 4 题两两子弹匹配信息的数据。