

# 基于随机抽样一致性算法的运动平面上点的视觉测量及其应用

RANSAC based visual measurement of point on ground plane  
and application in sport

王岚<sup>1</sup>, 沈乐君<sup>1</sup>, 柯遵渝<sup>1</sup>

Wang Lan<sup>1</sup>, Shen Le-jun<sup>1</sup>, Ke Zun-yu<sup>1</sup>

**摘要:** 使用计算机视觉技术, 对球场中运动员和关键点的位置进行测量, 是基于视频的技战术分析系统的核心技术, 也是运动解析的重要原理, 更是体育科学研究的重要工具。本文介绍了其理论基础与实现方法, 探讨了测量误差和 RANSAC 容错算法, 介绍了使用三维模型进行标定和结果演示。大量实验证明, 本文提出的方法和传统方法相比, 提高了纠错能力和适应性, 在体育科研中有广泛的应用价值。

**关键词:** 视觉测量; RANSAC; 单应性矩阵; 图像坐标; 球场坐标; 球落点; 网球; 运动员跟踪;

**Abstract:** computer vision based visual measurement of points on ground plane via video or digital image are introduced by paper. What we presented is fundamental for sport software and crucial for the tactic analysis. The paper introduces the basic theories; discuss the error of measurement and use RANSAC algorithm framework to improve algorithm accuracy. The 3D model can visible and very useful when we do calibration and presentation of analysis. The results of experiments showed that the method in the paper is more robust, flexible than traditional method and can be widely used in many fields of sport science.

**Keywords:** visual measurement; RANSAC; homography matrix; image coordinates; stadium coordinates; serve placement; tennis; athlete tracking;

作者简介:

1、王岚(1961-),女,成都人,工程师,本科,研究方向为运动训练学, Tel(028)85095382, E-mail: wwanglan2002@yahoo.com.cn;

2、沈乐君(1976-),男,重庆人,讲师,博士研究生,研究方向为计算机视觉在运动训练中的应用, Tel(028)86412928

3、柯遵渝(1953-),男,四川成都人,教授,硕士,研究方向为体育测量与评价; Tel(028)85093764, E-mail ke5380@163.com ;

作者单位:

1、成都体育学院,成都,610041

作者单位(英文)

1、Chengdu Sport University, Sichuan Chengdu,610041

## 1 引言

分析视频图像，获得运动员的技战术数据，不但具有客观、无干扰的特点，而且能够在比赛条件下进行的，是目前最为有效的科研手段之一<sup>[1][2]</sup>。通过图像分析获得目标的定量信息，如运动员在球场上的位置<sup>[3]</sup>，一直都是计算机视觉研究的主要任务，也是体育科学研究的迫切要求。数字化设备采集的图像，都以像素为基本单位，只有将之变换为公制单位，基于图像的定量分析才具有客观性和实用性。视觉测量，就是应用图像处理方法，计算图像中目标的大小、位置、速度等客观信息。

本文涉及的几个概念。如果将运动场地抽象成一个二维平面，且场地上运动员的尺寸忽略不计，也抽象成平面上一个点，本文称这个平面为“运动平面”。具体而言，由摄像机拍摄比赛成为数字图像，叫做“图像平面”（单位：pixel）。待测量的点在图像平面上的位置称为“图像坐标”，它在运动平面上的对应位置叫做“球场坐标”（单位：m）。

如图 1 所示，A、B 和 C 三个红色圆点，本是球场上同一个位置 D，但是，由于摄像机拍摄角度不同，远近不同，在三幅图像中有不同的图像坐标值。本文首先介绍了运动平面和图像平面的点的投影几何关系，然后探讨了数字模型和图像像素进行拟合的原理，然后介绍了图像坐标转换到球场坐标的具体公式。

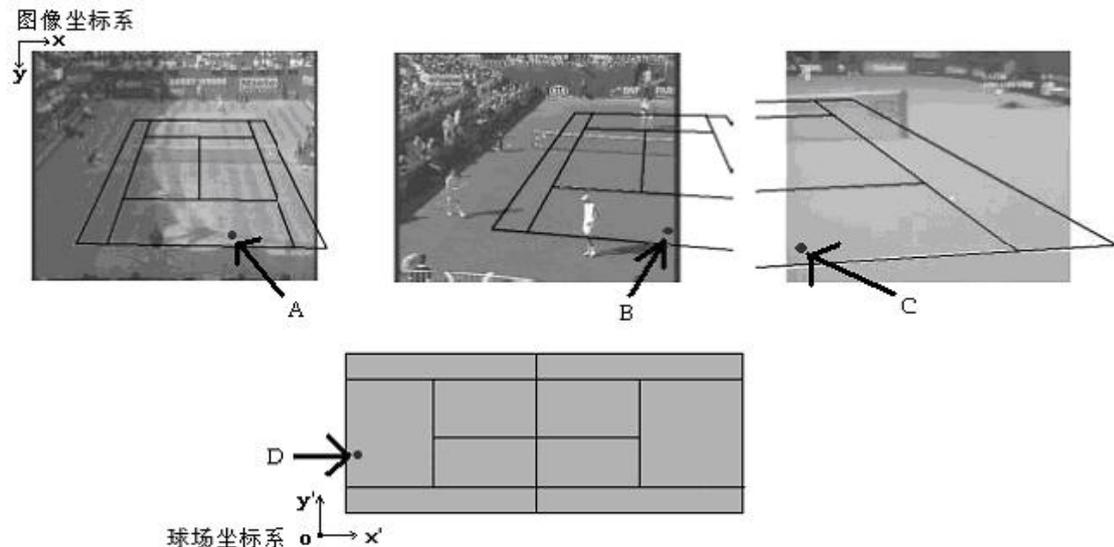


图 1 球场坐标和图像坐标

如何评价和避免测量误差<sup>[4]</sup>，直接决定本研究是否有应用价值。本文探讨了由摄像机导致的光学误差，人为选点误差和测量方法本身的误差。实验证实，通过增加标定点数量，可以提高可靠性和测量精度。同时，应用RANSAC算法框架，从多个点筛选出质量“好”的点，可避免数据异常点导致的测量偏差，提高精度。

随机抽样一致性算法RANSAC (Random Sample Consensus) 是应用广泛的模型拟合算法，对于错误率超过 50% 的数据仍然能够处理<sup>[5]</sup>。这些由于人为误差、系统误差、光电干扰造成的数据外点，若直接进行模型拟合将导致灾难性结果。使用RANSAC，可以有效的筛除外点，提高计算的鲁棒性。它在很多领域得到应用，如双目视觉中基础矩阵估计<sup>[6]</sup>、特征点匹配<sup>[7]</sup>、机器人运动估计、飞行器飞行时移动目标检测等。

典型应用之一：球的落点分析。在隔网对抗类项群中，双方运动员的相互牵制是通过球的速度、旋转、力量、落点和弧线等要素来实现的。人们对球的速度、

旋转等较为重视，但球的“落点”在比赛中，常常能够起到四两拨千斤的作用<sup>[8]</sup>，体现出运动员的技术水平、战术习惯、临场变化能力。本文提出的技术，能够对球在运动平面上的落点进行测量，其结果便于科研人员进行分析统计，后半部分有具体示例。

典型应用之二：运动员的跑动研究。在足球等项目中，由于场地比较宽阔，可以将球场上的运动员抽象成“运动平面”上的“点”。有研究使用SMI Scout™，获得足球运动员比赛的运动距离、运动速度等参数，从而对运动员比赛活动能力进行定量分析<sup>[9]</sup>。若进一步结合运动员实时跟踪技术<sup>[3]</sup>，可以在比赛条件下，自动、实时的获得运动员的位置信息。本文给出了实验结果。

其他的应用。由于水平和时间有限，本文只给出了初步的实验结果。显然，只要满足“运动平面”的假设和点在平面上的假设，本研究的应用领域十分广泛。

## 2 视觉测量原理

建立图像平面上的点和运动平面上的点之间的变换关系。根据投影几何理论<sup>[10]</sup>，将图像平面  $\pi$  上的点用齐次坐标表示  $p=(x, y, 1)T$ ，将运动平面  $\Pi$  上的点用齐次坐标表示为  $q=(x', y', 1)T$ 。这两点是同一条光线，分别在两个平面上的投影点，如下图。

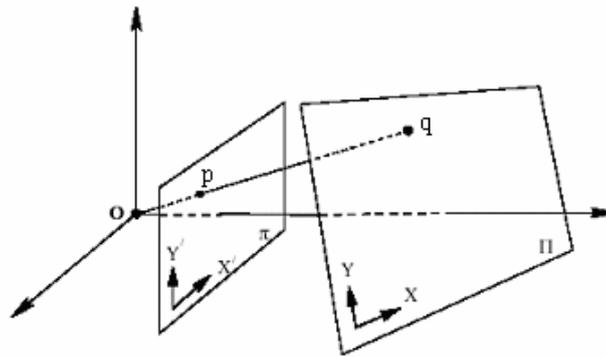


图 2，投影几何空间中的点与两个平面

投影变换是最一般的图像变换，它可以描述针孔摄像机成像平面上点  $p$  和世界坐标中点  $q$  之间的几何关系<sup>[10]</sup>。 $p$  和  $q$  之间的变换用  $3 \times 3$  的矩阵  $H$  表示，即  $q=Hp$ 。 $H$  是投影变换矩阵，又叫单应性矩阵 (Homography)，它可以描述一系列放大缩小，旋转、平移操作和透视效果，是计算机图形学中生成三维画面的关键。变换用公式 (1) 表示。

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (1)$$

视觉测量分为 2 个阶段：标定和变换。建立图像平面  $\pi$  和运动平面  $\Pi$  之间点与点的对应关系： $p_i \leftrightarrow q_i$  ( $i=1, 2, 3 \dots n$ )，求得变换矩阵  $H$  中各个未知参数，叫做标定。就是由标定得到的  $H$  和公式 (1)，计算图像平面上待测量的像素点  $p$  对应的球场坐标  $q$ ，即变换阶段。

以下是标定过程求  $H$  的公式推导。若已知对应点  $p$  和  $q$ ，由公式 (1) 可得：

$$x' = \frac{h_{11}x + h_{12}y + h_{13}}{h_{31}x + h_{32}y + h_{33}}, y' = \frac{h_{21}x + h_{22}y + h_{23}}{h_{31}x + h_{32}y + h_{33}} \quad (2)$$

若已知 n 对点，整理 (2) 得方程组：

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & x'_1 x_1 & x'_1 y_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & x'_2 x_2 & x'_2 y_2 \\ \dots & & & & & & & \\ x_n & y_n & 1 & 0 & 0 & 0 & x'_n x_n & x'_n y_n \\ 0 & 0 & 0 & x_1 & y_1 & 1 & y'_1 x_1 & y'_1 y_1 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & y'_2 x_2 & y'_2 y_2 \\ \dots & & & & & & & \\ 0 & 0 & 0 & x_n & y_n & 1 & y'_n x_n & y'_n y_n \end{bmatrix} \begin{bmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \\ x'_1 \\ x'_2 \\ \dots \\ x'_n \end{bmatrix} \quad (3)$$

已知  $h_{33}=1$ ，故 H 的自由度为  $8=(9-1)$ 。只要 n 大于等于 4，就可以解线性方程 (3) 求出变换矩阵 H。显然，标定在视觉测量中起着关键作用，下一节主要讨论标定过程出现的误差。

### 3 测量误差

首先确立评价测量误差的指标。如图 3，以图像中的 A 为基准点 ( $x=176, y=181$ )，使用 H 将之变换到运动平面，离理想点 ( $x=5.485, y=5.485$ ) 的距离 D 为评价指标。距离 D 越小，模型拟合越好，反之误差越大。

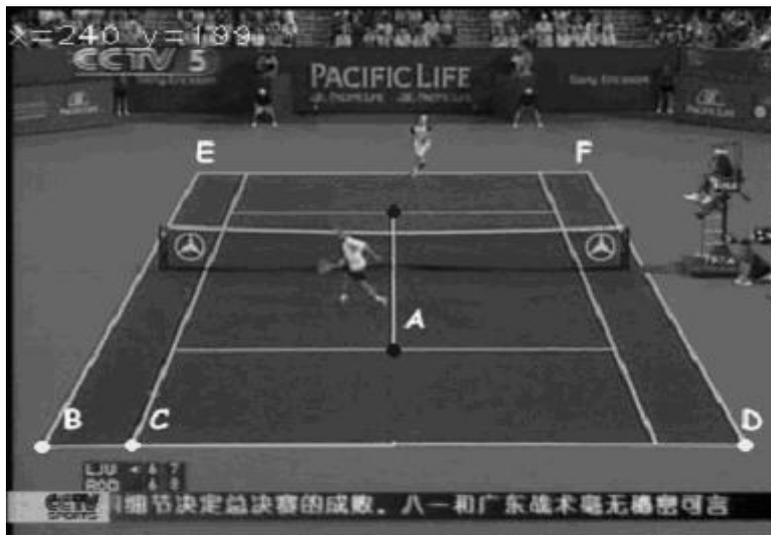


图 3 测量误差

光学器件的量化误差。本文实验所用的所有图像，均采集于 CCTV-5 的 AV 信号，其分辨率较低，由此导致误差也较大。如图 3 中，接近摄像机的球场边线 BD 的误差约为 0.034 m/pixel，远端边线 EF 误差为 0.062 m/pixel。这种由于数字化器件量化导致的测量误差无法克服，采用更高解析度的 DV 设备有利于提高精度。

选点误差。有两个途径对图像平面进行标定。其一，使用者用鼠标选择图像中的坐标点，如图 3 中的黄色点。其二，计算机自动选取特征点，然后进行特征点匹配。无论是手工标定还是自动标定，都容易出现选点误差。尤其是离摄像机较远的标定点，出现偏离后，对变换矩阵的计算影响更大，故标定时宜多选接近摄像机的点。

系统误差。将运动员抽象成运动平面上的点，球员的身高被忽略不计，会导致测量出现误差。可以通过对运动员的位置进行矫正来克服。

算法误差。采用4点法计算H过程中，若出现3点共线，将导致H计算结果退化（病态），结果不可靠。如图3，近端边线有3个黄色标定点B、C、D共线，公式（3）求出的H矩阵退化。若使用病态的H进行测量，将导致误差指标畸高 $D=5.74532\text{ m}$ 。所以，标定过程中，要尽量避免3点共线的情况出现，或者增加标定点数。

#### 4 提高精度和避免误差

通过提高标定点数量，提高测量精度。虽然标定过程只需要4个点即可，可统计学告诉我们，样本数越多<sup>[4]</sup>，变换矩阵的计算结果就应该越精确。当点数n超过4时，使用最小二乘法求解公式(3)，可得均方意义下的最优解。以下简称直接法。

如图4，左图采用4点标定，右图采用5点标定。5点标定计算H的误差明显小于4点，只有0.004 m。分析原因，一是新增点在近端，引入的光学误差本身就小，二是多点标定后，方程个数大于未知量个数，解的平方误差最小。



4点标定，误差  $D=0.0302395\text{m}$

5点标定，误差  $D=0.0038644\text{m}$

图4 多个标定点的误差

刚才讨论标定点超过4个的理想情况，但是选点误差的最差情况也可能发生，如用户错选模型点。一个有严重误差的点，会使计算结果远远偏离理想值。以直线拟合为例，如图5所示，用7个点来拟合一条直线，6个点有效，第7点有严重误差（称为外点 outliers），若使用最小二乘法直接求解，这1个点足以导致结果严重出错（虚线表示），远远偏离理想值（实线表示）。所以，为了避免最差情况，不能直接法计算变换矩阵H。

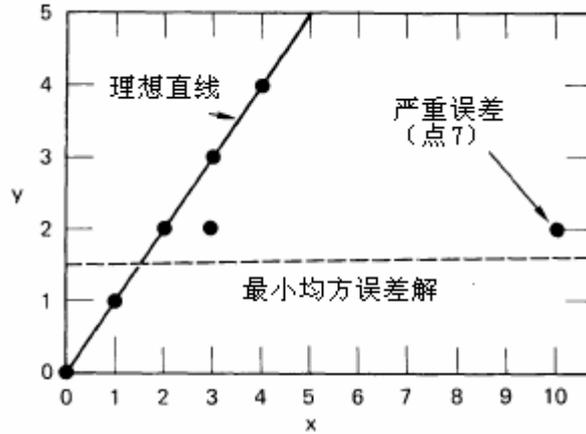


图 5 RANSAC的鲁棒性<sup>[5]</sup>

为筛除模型拟合中的外点，本文引入RANSAC算法<sup>[6][11]</sup>，其算法框架如下：

第 1 步：由置信概率  $P$  计算最小抽样数  $k$

第 2 步：从数据中均匀随机的采样  $n$  个点，对这  $n$  个点进行拟合

第 3 步：对  $n$  个点以外的点，比较到直线的距离小于  $t$ ，表示这个点很靠近，是好点。

第 4 步：若由  $d$  个或更多的点靠近直线，那么是一个好的拟合，计算拟合误差  $E$ 。

第 5 步：回到第 2 步，迭代  $k$  次。

第 6 步：使用拟合误差  $E$  作为准则，挑选出最好的拟合作为解。

RANSAC 法和直接法对比实验。通过同一场景的不同标定点，分析 RANSAC 算法和直接法的误差和出错率。由于算法本身带有随机性，为了真实反映算法的特性，都是 1000 次计算的误差平均值  $D_{avg}$ 。1000 次实验中，超过误差阈值的结果，被认为计算  $H$  出错，求得出错率  $E$ 。实验参数： $n=4, d=5, P=0.99, t=0.03$ ，误差阈值= $0.1m$ 。

表 1 RANSAC 法求解

序号	外点数	总点数	出错率 $E$	平均误差 (单位 $m$ )	备注
(1)	0	4	0%	0.023396	图 6
(2)	1	5	0%	0.023396	图 6
(3)	3	7	0%	0.049429	
(4)	4	9	5%	0.104627	图 6
(5)	8	16	70%	0.242713	图 7

表 2 直接法求解

序号	外点数	总点数	出错率 $E$	平均误差 (单位 $m$ )	备注
(6)	0	4	0%	0.02339	图 6
(7)	1	5	100%	2.66056	图 6
(8)	3	7	100%	3.37114	
(9)	4	9	100%	4.66382	图 6
(10)	8	16	100%	1.88813	

RANSAC 实验结果分析。如图 6，浅色的点表示 RANSAC 判断的“好”点，深色的点表示“坏”点。实验 1、2、3 中，能够计算出正确的  $H$ ，误差指标  $D$  的平均值小于光学误差。实验 4 中，虽然数据外点占 45%，出错率却很低，原因是 RANSAC 选择时，球场边线四角的 4 点是好点，它们对模型拟合的影响较大。实

验 5 的外点占 50%，但由于外点呈随机分布，且设定的出错阈值 0.1 较高，导致出错率很高。

直接法实验结果分析。如表 2，实验 6 和实验 1 点数等于 4，RANSAC 法和直接法完全等效。实验 7 至实验 10 中，直接法无法区分好点和坏点，都用于求方程，所以，错误率达 100%。误差也很高，即使有 1 个错误点的误差也超过 2m。显然，直接法求 H 根本没有任何实用价值，必须引入 RANSAC。



实验 1、6

实验 2、7

实验 4、9

图 6 对比实验

相反，如图 7 中左图所示，实验 5 是模拟实验中最差的情况，外点和好点随机出现。即使有 50% 的标定错误，RANSAC 算法也能找到其中和球场模型匹配很好的 6 个黄色的“好点”。1000 次实验中，其大部分误差都小于 0.3m，表明算法即使找不到最优解，也有能力找到近似的“好点”，远远小于采用直接法计算的误差。

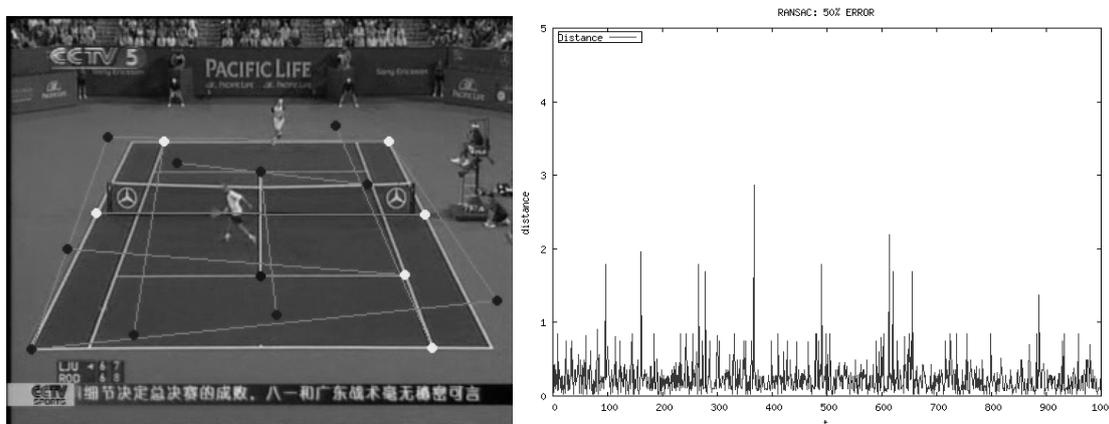


图 7 实验 5 的好点及其测量误差

总之，参与标定点数越多，可靠性越高，测量精度也越高。引入 RANSAC 算法后，极大的提高了计算 H 的鲁棒性，减小了标定过程选点误差的影响。

## 5 视觉测量的应用

### 5.1 自定义三维模型提高应用的普适性

本研究可自己定义任意的三维场景模型参数，并以此进行标定，所以适用于任何尺寸的场地，包括非标准的场地。其次，自定义场景模型中定义了需要标定的点的精确位置，可以避免标定过程中可能的选点误差。再次，由于三维模型可以在任意角度进行观察，不拘泥于整个场景的，故依据部分球场图像就可以进行标定。如图 8，根据足球场的部分图像，即禁区线上 5 个点就可以计算出 H。最后，在三维空间中观察运动平面，可以任意旋转缩放，有利于分析战术，将结果

可视化地呈现给教练员。如图 8 左图，可以明显看到被测量点和禁区线平行。

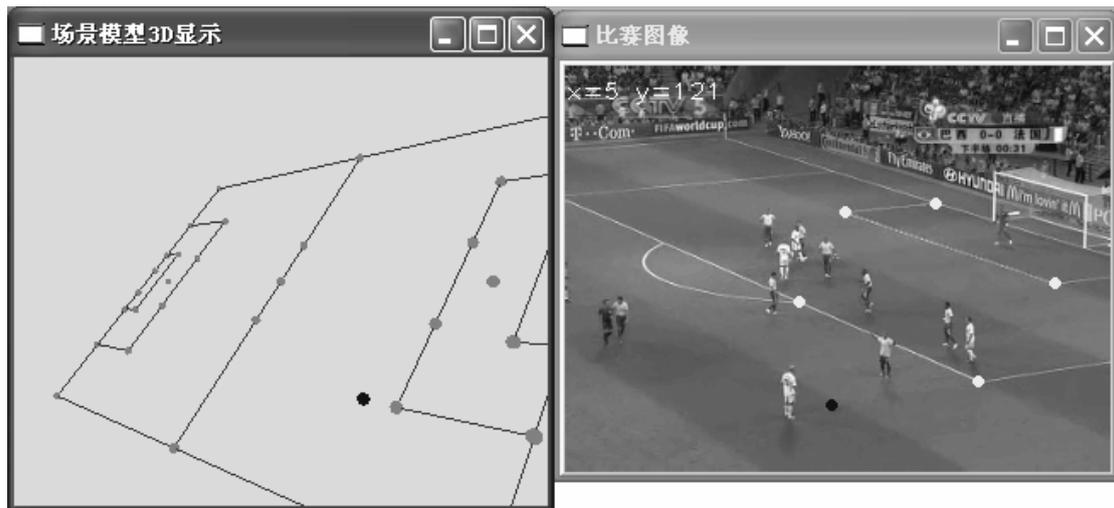


图 8 三维场景模型、被标定图像和测量点

### 5.2 球落点统计

在隔网对抗类项群中，接发球技术对比赛胜负起着关键性的作用<sup>[12]</sup>。尤其是在比赛过程中，且不能实地拍摄情况下，要想获取技战术信息，只能通过计算机视觉技术分析电视画面。本次实验的视频来源为ATP大师赛印第安威尔斯站的CCTV现场直播，世界排名 2、3 的男子网球运动员Nadal与Roddick的第一局较量，时间 42 分钟。如图 9，运动员交换场地、摄像机镜头频繁切换、镜头焦距不断变换和平移较多、只拍摄部分网球场等因素，都给标定带来困难。但是，凭借自定义三维模型和RANSAC算法，我们都能够寻找到对应的变换矩阵H，将测量点图像坐标转换到球场坐标，然后统计分析。



图 9 电视转播的多角度画面

示例 1，接发球站位统计。在 Roddick 的发球局中，对 Nadal 的接球站位进行统计。如图 10，原点在左上角，左为发球方，右为接球方，其接球时站位明显聚集于 2 个区域。

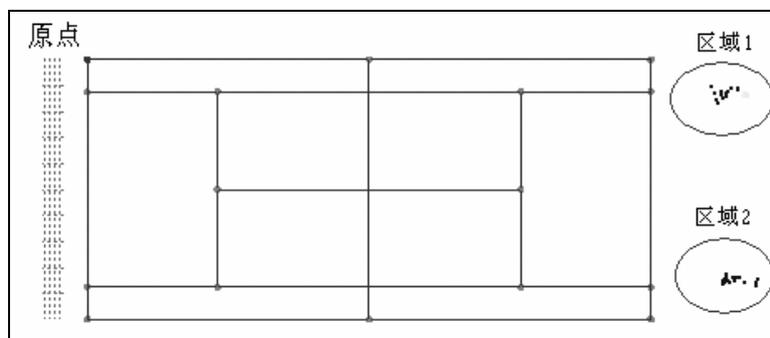


图 10 Nadal 接发球位置

对两个区域的点分别求平均值得到下表 3，这对我国运动员的训练有借鉴意义。

表 3 Nadal 接球区域平均坐标

名字	X (单位: m)	Y (单位: m)
区域 1	1.4361074	26.940498
区域 2	9.249542778	27.28566261

示例 2, 双方发球落点统计分析。

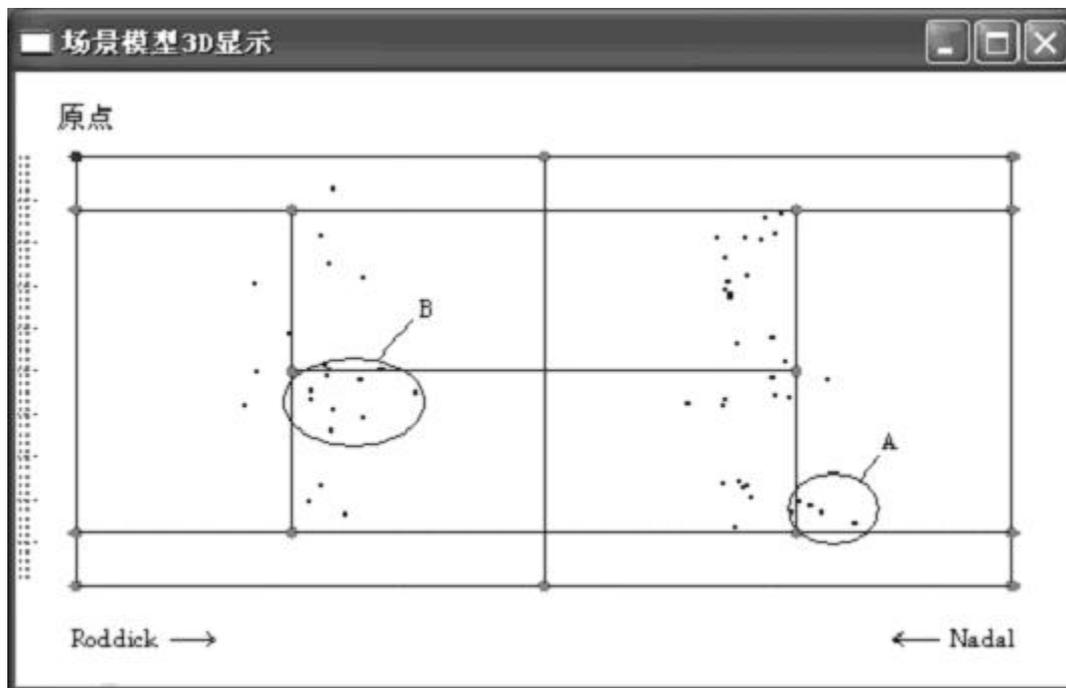


图 11 发球落点统计

如图 11, Roddick (从左向右) 在发球局中, 将球发到对方半场共 33 次 (含发球失误落点), 同样, Nadal (从右向左) 在发球局中, 将球发到对方半场共 23 次。从落点分布中, 可看出某些规律: Roddick 在的发球失误落点, 多在 A 区域, 其外角球成功率偏低; Nadal 的发球落点, 集中在 B 区域; Roddick 发球偏重于内外角和对方反手 (落点分布的方差为 2.587m), ACE 球比对手多; Nadal 发球更为随机 (方差为 1.886m), 落点不易预测。

若具有多的电视录像资料, 结合发球成功率、击球线路等信息, 视觉测量和落点分析可能会更有价值。

### 5.3 足球运动员实时跟踪

只有将屏幕坐标转换为球场坐标, 足球运动员的视觉跟踪系统<sup>[3]</sup>才有实用价值。因此, 将粒子滤波器输出的坐标值乘以变换矩阵H, 就可以获得该目标的球场坐标, 然后计算运动员的跑动距离、跑动速度等。



图 12 目标跟踪的球场坐标和图像坐标

示例 1，运动员跟踪。试验素材为成都体育学院某比赛的视频片断，场地尺寸  $90 \times 68\text{m}$ ，时间长度 1 分 3 秒。选择的目标是红队某后卫，自动跟踪他从慢跑 to 加速跑，再到慢跑的全过程。如图 12，右图中蓝色点连成了运动轨迹（图像坐标），左图曲线为坐标变换后的真实轨迹（球场坐标）。该后卫跑动总距离 125.8m，平均速度 2.0449m/s，最大速度 8.08m/s，发生在第 949 帧处。如图 13，速度曲线不够平滑，是因为自动跟踪算法的输出坐标没有平滑处理，和运动员步伐导致的坐标值小幅摇摆波动。

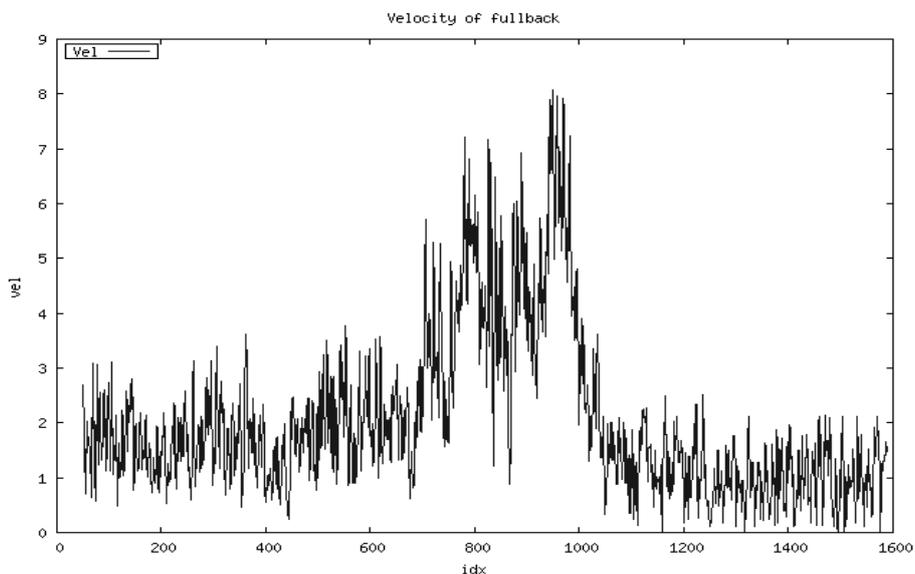


图 13 某队员的跑动速度-时间曲线

示例 2，发球点和落点。有研究使用观察法结合分区，对足球落点进行统计分析，得出定位球和角球攻防的重点区域<sup>[13]</sup>。本文提出的方法，可以精确的测量其落点，如图 8 是 2006 足球世界杯决赛中，齐达内的第一个任意球，测量其发球点，球场坐标位置  $x=82.9765\text{m}$ ,  $y=14.7029\text{m}$ ，几乎和禁区线平行。

## 6 进一步工作：

目前场景标定工作尚倚赖手工，虽然在摄像机镜头固定时标定一次即可轻松完成点测量，但是，当镜头变化较多时（如电视体育转播画面），需要不断进行修正，消耗时间较长。下一步工作，将研究球场自动识别和自标定技术。

可进一步对测量技术进行扩展。如对球场固定区域的落点检测与反馈,可以帮助教练员和运动员在训练中进行落点监控和准确性专项练习。再如,对田径训练中运动员自动跟踪测量和反馈,可以实时监控训练过程中的体能指标。

由于时间和精力有限,应用研究面不够广。需要本文程序的研究者可向作者索取,以便完善提高。

#### 参考文献:

- [1] 李爱东. 视频图像技术的发展及其在田径项目训练中的应用[J]. 体育科学, 2005, 25(6): 42-46.
- [2] 刘建和. 运动竞赛学:现状与重点关注问题[J]. 北京体育大学学报, 2004, 27(9):1262-1263.
- [3] 沈乐君,柯遵渝. 球类比赛中运动员的实时跟踪[J]. 体育科学, 2007, 27(1):64-67.
- [4] 魏登云. 体育测量客观性、可靠性和有效性的对比分析[J]. 北京体育大学学报, 2002, 25(4):499-501.
- [5] M. FISCHLER, R. BOLLES. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography[J]. CACM, 1981, 24(6):381 - 395.
- [6] DAVID A FORSYTH, JEAN PONCE. Computer Vision:A Modern Approach[M]. 北京:电子工业出版社, 2004.
- [7] D. LOWE. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2):91 - 110.
- [8] 张秋芬. 47 届乒乓球锦标赛男子团体决赛击球落点分析[J]. 中国体育科技, 2004, 40(6):64-66.
- [9] 刘丹. 对国家男子足球队运动员比赛活动能力的研究[J]. 中国体育科技, 2006, 42(4):10-15.
- [10] R. HARTLEY AND A. ZISSERMAN. Multiple View Geometry in Computer Vision[M]. Cambridge: Cambridge University Press, 2000.
- [11] CHUM O, MATAS J. Matching with PROSAC -- Progressive Sample Consensus[C]. Conference on Computer Vision and Pattern Recognition (CVPR), 2005, 220-226.
- [12] 李庆有. 网球接发球研究. 中国体育科技[J], 2006, 42(5):62-65.
- [13] 杨次榆. 第 18 届世界杯足球赛前场定位球进攻战术运用情况分析[J]. 成都体育学院学报, 2007, 33(1):81-84.