

# 基于 NSCT 的 SMP 立体匹配算法研究

丁 辉,付梦印,王美玲

(北京理工大学信息科学技术学院,北京 100081)

**摘 要:** 针对区域立体匹配算法中匹配窗口自适应选择的难点,提出一种基于多尺度分析 Nonsubsampled Contourlet 变换(NSCT)的快速区域立体匹配新算法.首先介绍 NSCT 并对其时移不变性进行讨论.采用由粗到精的匹配策略和快速的 SMP 匹配代价函数降低运算量.实验结果表明,该算法具有很高准确率.

**关键词:** 区域匹配; Nonsubsampled Contourlet 变换; 时移不变性; 图像多尺度几何分析

**中图分类号:** TP391 **文献标识码:** A **文章编号:** 0372-2112 (2008) 04-0772-05

## SMP Stereo Matching Based on Shift-Invariant NSCT

DING Hui, FU Meng-yin, WANG Mei-ling

(School of Information Science and Technology, Beijing Institute of Technology, Beijing 100081, China)

**Abstract:** One of the central problems in stereo matching is the selection of optimal window sizes for comparing image regions. This paper presents a fast and reliable multi-scale stereo matching algorithm based on the Nonsubsampled Contourlet Transform (NSCT), which is a fully shift-invariant, multi-scale, and multi-direction expansion that has better directional frequency localization. This method consists of multiple passes which compute stereo matches with a coarse-to-fine and sparse-to-dense paradigm. Then, we using a fast local algorithm, referred to as single matching phase (SMP), which reduce computation and is suitable to real time applications. The proposed approach is also compared with bidirectional matching (BM), since the latter is the basic method for detecting unreliable matches in most area-based stereo algorithms. Experimental results with real images are presented.

**Key words:** area-based matching; nonsubsampled contourlet transform; shift-invariance; image multi-scale geometric analysis

## 1 引言

自主车(简称 ALV)是指基于多传感器的移动机器人系统,由于其重大的军事及民用价值已经成为当前研究的一个热点.立体视觉是自主车视觉导航的重要组成部分,基本方法是从两个或多个视点去观察同一场景,获得不同视角下的一组图像,然后通过三角测量原理获得不同图像中对应像素的位置差别,即视差,进而推断场景中目标的空间位置.

立体视觉中最关键和最困难的步骤是立体匹配.它与普通的图像配准不同,立体图像对之间的差异是由拍摄时观察点的不同引起的,而不是由其他景物本身的变化、运动所引起的.目前,已提出的获得致密视差图立体匹配算法大致可以分为基于全局的立体匹配算法和基于区域的立体匹配算法两类<sup>[1]</sup>.针对自主车视觉导航的实际情况,即立体视觉是用于检测障碍物,而不是试图去精细的描述整个道路场景,其中最常用的是基于区域

的立体匹配方法.这种算法使用对应像素具有相似亮度的假设,意味着对于每一个匹配窗口计算一个且仅一个视差,因此对匹配窗口大小的选择要求严格,既要求匹配窗口不能过大以保证匹配窗口内视差一致,又要求匹配窗口必须足够大以提高图像区域的信噪比.为了解决立体视觉匹配问题,提出了很多方法如固定窗口匹配<sup>[2]</sup>、自适应窗口匹配<sup>[3]</sup>,基于高斯滤波的多尺度匹配和利用 Laplacian 金字塔结构的匹配方法以及采用小波变换等多分辨率分析<sup>[4]</sup>实现图像的立体匹配等.另外,基于分割的立体匹配<sup>[5,6]</sup>,近年来也受到越来越多的关注,这类方法的匹配精度依赖于采用的分割算法,而好的分割算法复杂度往往很高,不适用于立体视觉领域的实时应用场合.

本文提出了一种新的基于多尺度分析的快速区域立体匹配算法.对于自主车上采用两个同样的 CCD 摄像头组成的双目视觉系统,采集的立体图像对可以看成是其中一幅图像由另一幅图像经过空间平移得到的,因

此匹配过程要求变换具有时移不变性(shift-invariance)。多尺度分析是一种金字塔式的结构,这对于进行由粗到精的快速图像匹配是非常有利的。然而利用 Mallat 算法的多分辨率分析方法都缺乏时移不变性,对图像平移是很敏感的,处于不同位置的同一图像在各尺度下变换系数的能量会产生较大的差异,这种平移的不稳定性,给基于多尺度表示的图像匹配带来了一定的困难。本文的匹配算法采用 Nonsubsampled Contourlet 变换<sup>[7]</sup>进行分解,克服了多分辨率分析缺乏时移不变性的缺陷,而且在多分辨率之外又加上了多方向分析能够更好的保存纹理细节信息,并在实现过程中考虑匹配代价函数的快速运算。实验结果表明,该算法能够获得较高匹配率的视差效果图,同时可以降低匹配过程的计算量。

## 2 多尺度几何分析算法

### 2.1 Contourlet 变换

小波变换在分析点瞬态特征的奇异性时是最优的,但在表示图像结构的直线/曲线奇异性时却不是最优的。为解决这一局限性,多尺度几何分析的思想应运而生。Contourlet 变换<sup>[8,9]</sup>是由 M. N. Do 等人提出的一种离散图像的方向多尺度计算框架,它的定义如同微积分的定义一样,认为在足够小的尺度下,曲线可以被看作为直线,曲线奇异性就可以由直线奇异性来表示。Contourlet 变换是一种“真正”的二维图像稀疏表达方法,它不仅继承了小波变换的多分辨率时频分析特征,而且拥有良好的各向异性特征,它能用比小波变换更少的系数来表达光滑的曲线。

### 2.2 NSCT 变换及其平移不变性

虽然 Contourlet 变换能够很好描述图像特征,但是由于在构造中使用了 Laplacian 金字塔滤波器,因此在计算中不可避免地要进行下采样运算,所以不具有时移不变性。

所谓时移不变性是指如果将输入信号沿时间轴平移,则变换后的输出信号除平移同样长度以外其他不变。时间平移不变性可以推广到二维的图像空间,即若输入图像相对于其原点有一平移,则输出图像除了相同的平移外其他不变。

在图像对的立体匹配过程中,可以将立体图像对看成是其中一幅图像是对另一幅图像在不同观察点获得的,所以要求变换过程必须满足时移不变性。NSCT 具有与 Contourlet 变换相同的构造思想和过程,它们的不同之处在于多尺度分解过程中 NSCT 采用非下采样的金字塔结构,多方向分析中采用非下采样的滤波器组实现。因此 NSCT 在变换过程中满足时移不变性,它可以看成是 Contourlet 变换的时移不变性算法。图 1 所示为分解尺度取 3 时 NSCT 的分解过程。

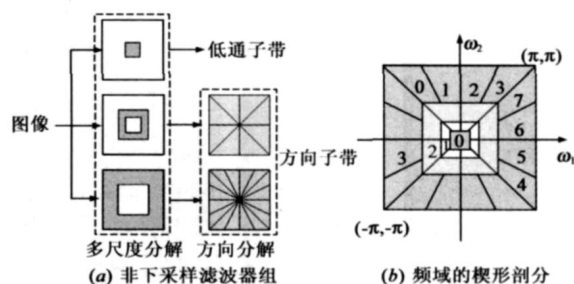


图 1 Nonsubsampled Contourlet 变换

## 3 算法原理及实现

### 3.1 视差

因为视角的不同,左右摄像头所观察到的景象是有区别的。设空间中一物点  $P$  的深度距离为  $Z$ ,在图像坐标系下,它在左右摄像机上投影像点的坐标分别为  $(x_L, y_L)$  和  $(x_R, y_R)$ ,根据三角测量法得到  $Z = Bf \frac{1}{x_L - x_R}$ ,其中  $B$  为基线距离,  $f$  为相机的焦距。记  $P$  点的视差为  $d = x_L - x_R$ ,立体匹配的目的就是求出左右图像之间的视差,然后进行三维重构获得目标的深度信息。本文算法的设计和实现遵循如下视觉理论的约束条件:外极线约束;匹配唯一性约束;视差连续性约束。

### 3.2 代价函数

基于区域的立体匹配算法是在实际应用中使用的最广泛的算法,因为它有一个最大的优点——计算简单,一次匹配速度可以很快,同时可以获得致密的视差图不需在后期处理中进行插值运算。对于自主车导航来说,如果计算能够以很快的速度完成,则系统就可以提早探测到障碍物而及时采取有利的行动。这里,首先采用像素灰度差的绝对值平方和(SSD)作为亮度相似性的度量因子,匹配的代价函数定义为:

$$C_{SSD}(x, y, d) = \sum_{(x, y) \in W_p} (I_L(x, y) - I_R(x, y))^2 \quad (1)$$

其中  $W_p$  为匹配区域,  $W_p$  的大小一般选取为  $(2m+1) \times (2n+1)$ 。文献[10]根据单匹配相位(SMP)原理,分析了在计算图像中每个像素的相关亮度时,模板在水平方向和竖直方向上具有平移特性。根据此原理可以对代价函数  $C_{SSD}$  进行计算优化,将其改写成如下形式:

$$C_{SMP}(x, y+1, d) = C_{SSD}(x, y, d) + C(x, y+1, d) \quad (2)$$

$C(x, y+1, d)$  表示匹配窗口中最上一行与最下一行之间  $C_{SSD}$  的差,如图 2 中深蓝色行:

$$C(x, y+1, d) = \sum_{j=-n}^n |L(x+j, y+n+1) - R(x+d+j, y+n+1)|^2 - \sum_{j=-n}^n |L(x+j, y-n) - R(x+d+j, y-n)|^2 \quad (3)$$

考虑图 2 中紫色像素点,  $C(x, y+1, d)$  可由  $C(x-1, y+1, d)$  计算得到:

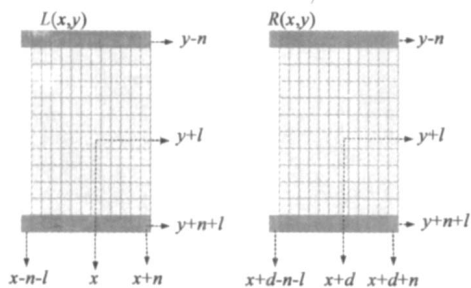


图2 增量的计算区域

$$C(x, y+1, d) = C(x-1, y+1, d) + |L(x+n, y+n+1) - R(x+d+n, y+n+1)|^2 - |L(x+n, y-n) - R(x+d+n, y-n)|^2 - |L(x-n-1, y+n+1) - R(x+d-n-1, y+n+1)|^2 + |L(x-n-1, y-n) - R(x+d-n-1, y-n)|^2 \quad (4)$$

比较式(1)和式(2), 可以看到当采用传统 SSD 计算方法的时候, 每一对待匹配点需要进行  $(2m+1)(2n+1)$  次平方运算, 采用 SMP 优化算法后, 如果  $C(x-1, y+1, d)$  未知, 则需要进行  $2 \times (2m+1)$  次平方运算, 如果  $C(x-1, y+1, d)$  已知, 则只需进行四次平方运算. 在匹配过程中选择采用(2)式作为代价函数可以大大降低匹配过程的计算量.

### 3.3 带通信号的获取

NSCT 分解过程主要由两个步骤实现: 子带分解和方向变换. 首先利用两通道非下采样塔式滤波器将左右图像分解成对应于不同空间频率的不同图层, 如图 3(a) 所示.

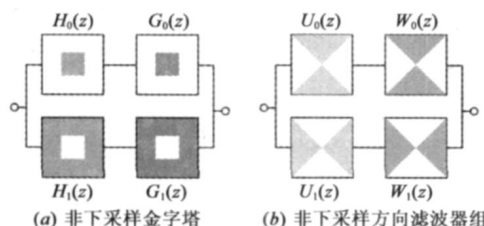


图3 子带获取

其中  $H_0(z)$  具有低通性质,  $H_1(z)$  具有高通性质, 图像经过一次分解后可以得到图像的概貌系数(低频部分)和细节系数(高频部分). 如果将 NSCT 分解输出的概貌部分继续进行分解, 就可得到原始图像在不同尺度上的细节和概貌, 里面包含了原始图像在不同尺度上的特征. 多分辨率分析只是对低频部分进行进一步的分解, 使频率的分辨率越来越高. NSCT 算法的精确重构条件为:

$$H_0(z)G_0(z) + H_1(z)G_1(z) = 1 \quad (5)$$

利用 a trous 分解算法, 当分解尺度为  $k$  时, 非下采样塔式滤波器将图像分解成如下频率子带:

$$H_n(z) = \begin{cases} H_1(z^{2^{n-1}})_{j=0}^{n-2} H_0(z^{2^j}), & 1 \leq n < 2^k \\ H_0(z^{2^j}), & n = 2^k \end{cases} \quad (6)$$

然后由扇形方向滤波器组(Directional Filter Bank)实现如图 3(b) 所示的频带划分, 它利用  $N$  级树形结构的扇形滤波器组把整个频域划分成  $2^N$  个楔形区域. Bamrberger 及 Smith 等首先提出了无冗余且能完全重构的方向滤波器组<sup>[11]</sup>. NSCT 分解中去除了 DFB 中的采样步骤, 是一种具有时移不变性的多尺度几何分析方法.

### 3.4 相似性匹配

对要匹配的图像对作 NSCT 变换, 分别得到左右图像的多分辨率、多方向分解金字塔. 图像匹配开始于图像的最粗一层, 在该层上, 左图和右图经过变换分别得到不同方向上的带通子图, 频率范围的划分如图 1(b) 所示.

如果用  $i$  表示子带序号,  $j$  表示尺度分解的级数, 在第  $j$  级中, 定义带通子图的匹配准则函数为:

$$CN_{SMP}(z_1, z_2, d) = \frac{1}{W_p} |NSCTL_{i,j}(z_1, z_2) - NSCTR_{i,j}(z_1, z_2)|^2 \quad (7)$$

对于左图像中的某一区域, 在右图像的外极线上搜索, 使准则函数  $CN_{SMP}(z_1, z_2, d)$  达最小值的点为最佳匹配. 这种方法在低分辨率层上选用大窗口进行匹配, 在两图像中的每个像素点计算大尺度的区域亮度值, 通过  $CN_{SMP}(z_1, z_2, d)$  代价函数的计算, 使左右两个图像相关联. 然后通过迭代的方式将上一级的匹配结果送到下一级图像中作为给定位置区域的正确视差的一个参考, 如此重复直到期望的分辨率层. 匹配窗口的选择可根据分辨率的增高而适当减小, 对于像自主车躲避障碍物这样的任务可能只需在粗分辨率下的信息, 这样就提高了效率.

## 4 实验结果及分析

实验用的图像是经过校正的灰度图立体图像对, 采用由左到右的匹配策略, 并以左图像为基准. 为验证本文算法的实际效果, 在 Inter(R) Pentium(R) 4 CPU, 2.8GHz, 内存为 256MDDR 的 PC 机上进行试验, 软件环境为 Matlab7.0. 首先对传统 SSD 算法、SMP 优化算法和文献[12] Progressive 区域匹配算法的计算时间进行比较, 实验中取图像大小为  $288 \times 384$ , 匹配窗口为  $15 \times 15$ , 仿真结果如表 1 所示. SSD 总的运算时间为 2.609s, SMP 总的运算时间为 0.735s.

表1 运算时间比较

Algorithm	Image Size	Scale Range	Runtime
SSD	288 × 384	15	2.609s
Progressive	288 × 384	15	1.209s
SMP	288 × 384	15	0.735s

使用 Tsukuba 和 Venus 标准立体图像对对本文算法进行测试, 采用 3 层 NSCT 分解, 每一层上做  $2^j$  ( $j$  为尺度

分解子带)的方向分解. Tsukuba 图像大小为  $288 \times 384$  像素,视差搜索范围为 0 到 15,Venus 图像大小为  $383 \times 484$  像素,视差搜索范围为 0 到 8. 利用本文算法分别与基本

SSD 的 BM 匹配方法、多尺度 Laplacian 金字塔 (LP) 匹配方法、Contourlet 分解匹配以及 Progressive 匹配方法比较,匹配如表 2 所示. 图 4 和图 5 为匹配的视差效果图.

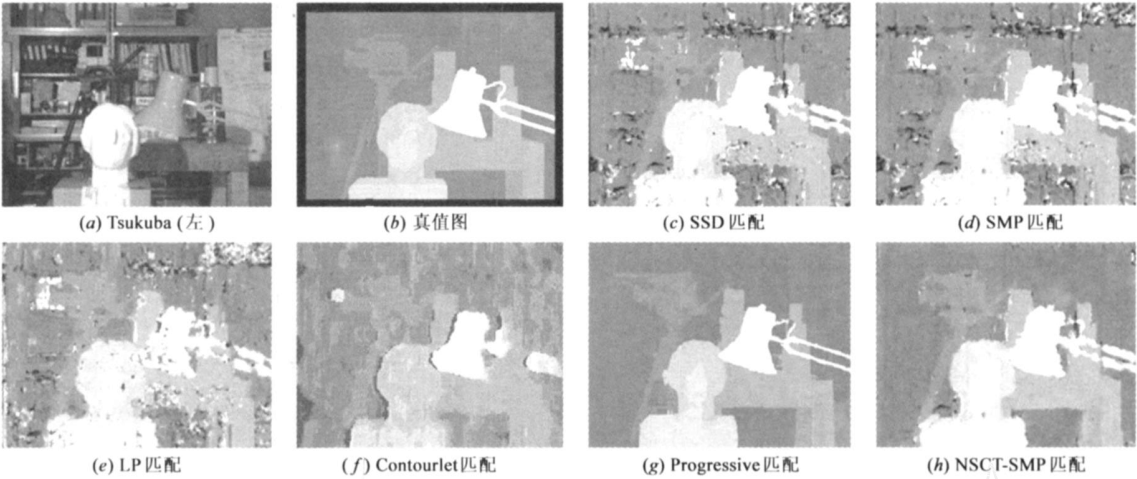


图 4 Tsukuba 视差匹配结果比较

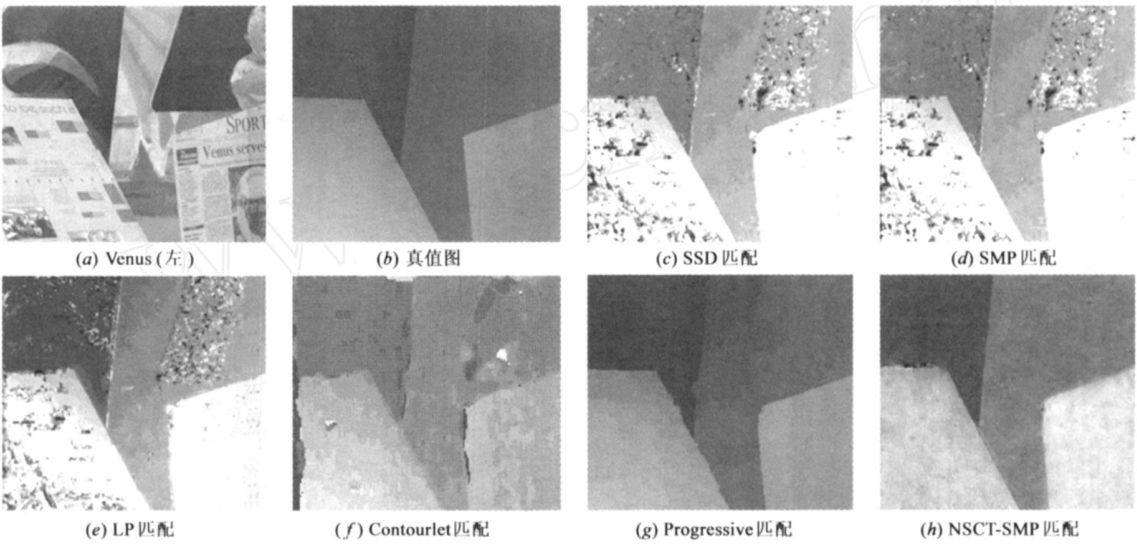


图 5 Venus 视差匹配结果比较

表 2 算法匹配结果比较

Algorithm	BM	SMP	LP	Contourlet	Progressive	NSCT-SMP
Image						
Tsukuba	86.2 %	88.4 %	88.2 %	68.0 %	97.1 %	96.7 %
Venus	80.2 %	85.2 %	74.4 %	69.7 %	85.5 %	93.3 %

从实验结果可以看出利用本文算法进行匹配具有较高的正确率,观察匹配结果发现,NSCT 分解与 Contourlet 变换虽然思想和方法都相同,但是因为 Contourlet 变换缺乏时移不变性导致匹配结果明显不好.虽然 Progressive 匹配方法在有些情况下,匹配效果优于本文的算法,但是本文的算法在计算的快速性上明显优于 Progressive 匹配算法.

5 结束语

本文所提出的基于 NSCT 的区域匹配算法,是一种

多尺度几何分析方法.匹配首先在最低分辨率层进行,然后逐层推进直至图像的最高分辨率.这种由粗到精的匹配控制策略,能够很好的解决区域匹配算法中匹配窗口的选择问题,在低分辨率层图像灰度比较均匀匹配窗口可以选取较大,在分辨率较高的子带可以将匹配窗口适当减小.同时在运算过程中对代价函数的计算方法进行了改进,显著地降低了计算量.实验表明,该算法有很高的匹配准确率.

为了使该算法具有更强的实用性,还需要进一步完善.如增加相邻像素点的对应匹配点也相邻等强约束条件以处理遮挡区域.对近距离小物体目标情况下,本文的算法采用外极线的搜索可能会造成误匹配,可以考虑利用图像分割方法进行预处理.本文算法由于使用滤波器组和图像的重复卷积运算,而滤波器的长

度设计的很短,且实部虚部全为整数,从而保证了算法的运算速度,并且可以通过硬件进行加速适用于自主车等对实时性要求较高的系统.另外,根据应用场合具体要求的不同,可以对图像匹配的稠密度和运算速度进行折衷考虑.

#### 参考文献:

- [1] Anandan P. A computational framework and an algorithm for the measurement of visual motion[J]. In International Journal of Computer Vision, 1989, 2(3): 283 - 310.
- [2] Kanade T, A Yoshida, et al. A stereo machine for video-rate dense depth mapping and its new applications [A]. In IEEE Conference on Computer Vision and Pattern Recognition [C]. San Francisco, 1996. 196 - 202.
- [3] Patricio M P, F Cabestaing, et al. A similarity-based adaptive neighborhood method for correlation-based stereo matching. in Image Processing [A]. In International Conference on Image Progress [C]. Singapore, 2004(2), 1341 - 1344.
- [4] Anderson R, N Kingsbury, et al. Coarse-level object recognition using interlevel products of complex wavelets [A]. In IEEE International Conference on Image Progress [C]. Genoa, 2005, (1): 745 - 748.
- [5] Bleyer M, Gelautz M. A layered stereo matching algorithm using image segmentation and global visibility constraints [A]. In International Conference on Image Processing [C]. Singapore, 2005, 128 - 129.
- [6] Hong L, Chen G. Segment based Stereo Matching Using Graph Cuts [A]. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition [C]. Washington, 2004, (1): 74 - 81.
- [7] Da Cunha A L, J Zhou, et al. The nonsubsampled contourlet transform: theory, design, and applications [J]. In IEEE Transactions on Image Proceeding, Genova, 2006, 15 (10): 3089 - 3101.
- [8] M N Do, M Vetterli. The contourlet transform: an efficient directional multiresolution image representation [J]. In IEEE Transactions on Image Proceeding, Genova, 2005, 14 (12): 2091 - 2106.
- [9] Duncan D, M N Do. Directional multiscale modeling of images using the contourlet transform [J]. In IEEE Transactions on Image Proceeding, 2006, 15 (6): 1610 - 1620.
- [10] Luigi di Stefano, Massimiliano Marchionni, Stefano Mattocchia. A fast area-based stereo matching algorithm [J]. Image Vision Comput, 2004, 22 (12): 983 - 1005.
- [11] R H Bamberg, M J T Smith. A filter bank for the directional decomposition of images: Theory and design [J]. IEEE Transactions on Signal Proceeding, 1992, 40 (4): 882 - 893.
- [12] Yichen W, Q Long. Region-based progressive stereo matching [A]. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition [C]. Washington, 2004 (1), 106 - 113.

#### 作者简介:



丁辉女, 1977年4月出生于黑龙江省双鸭山市, 2006年博士毕业于北京理工大学信息科学技术学院自动控制系统, 其后进入清华大学电子工程系, 现为博士后, 从事视频处理、智能监控及智能导航与组合导航方面的有关研究.

E-mail: ariel\_dh@gmail.com



付梦印男, 1964年11月出生于内蒙古, 教授、博士生导师, 中国惯性技术学会副理事长、科普部主任, 兵器工业光电技术专业组成员, 北京市科协理事, 1987年、1992年、2000年分别于辽宁大学、北京理工大学和中科院测地所获得工学学士、工学硕士、工学博士学位, 现为北京理工大学计算机科学技术学院、软件学院院长, 主要从事导航制导与控制方面的研究工作.

王美玲女, 1970年1月出生于山东省乳山市, 副教授、硕士生导师, 中国惯性技术学会会员, 1992年、1995年和2007年分别在北京理工大学获得工学学士、工学硕士和工学博士学位. 长期从事组合导航与智能导航技术研究工作.