

# 一种基于多模式单演特征检测与匹配的 三维视觉测量方法

李进军<sup>1,2</sup> 赵 宏<sup>2</sup>

(<sup>1</sup> 第二炮兵工程学院, 陕西 西安 710025  
<sup>2</sup> 西安交通大学机械制造系统工程国家重点实验室, 陕西 西安 710049)

**摘要** 三维视觉测量技术从一幅或多幅二维图像或者视频剪辑自动获取现实世界中目标和场景的几何模型、运动和形变等信息, 从而识别和理解三维现实环境。针对传统三维视觉测量易受亮度变化、尺度变化、旋转变化及噪声影响的缺点, 提出了一种基于多尺度单演特征检测与匹配的三维视觉测量新方法。该方法采用泊松(Poisson)滤波算子和里斯(Riesz)变换提取图像的多模式单演特征, 包括局部幅值、局部方向、局部相位和彩色相位; 模拟人眼视觉功能构建多模态的局部特征描述子及特征空间的相似性距离测量函数; 将该相似性测量应用于立体匹配和三维视觉重建中, 并进行了定性和定量的实验评估。

**关键词** 三维视觉测量; 单演信号分析; 立体匹配

**中图分类号** TH741 **文献标识码** A **doi**: 10.3788/AOS201131.0712007

## Three-Dimensional Vision Measuring Technique Based on Multi-Modal Monogenic Features Detecting and Matching

Li Jinjun<sup>1,2</sup> Zhao Hong<sup>2</sup>

(<sup>1</sup> The Second Artillery Engineering Institute, Xi'an, Shaanxi 710025, China  
<sup>2</sup> State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China)

**Abstract** Three-dimensional vision measuring technique aims at automatically reconstructing the properties of object and scene such as shape, motion, and deformation from one or more images or video clips in the real world around us, so that three-dimensional true environment can be recognized and understood. To solve the problems such as brightness change, scale change, rotation change and noise in conventional approaches, a new three-dimensional vision measuring technique is proposed. Firstly, the multi-modal isotropic monogenic features which include geometric feature (local orientation), structure feature (local amplitude and local phase), color phase and color texture are extracted by Poisson kernel and Riesz transforms. Secondly, a novel multi-modal feature descriptor and a similarity measurement function are constructed by simulating human vision function. Finally, the proposed similarity measurement function is applied to stereo matching and three-dimensional vision reconstruction. At the same time, the qualitative and quantitative evaluations are finished on the synthetic and natural stereo images.

**Key words** three-dimensional vision measuring; monogenic signal analysis; stereo matching

**OCIS codes** 120.4630; 100.2960; 330.1400; 330.7325

**收稿日期**: 2010-12-20; **收到修改稿日期**: 2011-03-14

**基金项目**: 国家 863 计划(2008AA04Z121)、国家自然科学基金(50575169)和陕西省西安市科技创新支撑计划(2010CXY1007)资助课题。

**作者简介**: 李进军(1975—), 男, 博士研究生, 主要从事机器视觉、图像处理和嵌入式系统开发等方面的研究。

E-mail: jinjun\_lee@stu.xjtu.edu.cn

**导师简介**: 赵 宏(1964—), 男, 教授, 主要从事智能光电检测、实时在线测试以及机器人视觉技术、生物组织成像等方面的研究。E-mail: zhaohong@mail.xjtu.edu.cn

## 1 引 言

三维视觉测量从一幅或多幅二维图像或者视频剪辑自动获取现实世界中目标和场景的几何模型、运动和形变等信息,从而识别和理解三维现实环境,为科学和工程数据分析提供定性和定量的几何信息。三维视觉技术经过近 40 年的研究,已经取得了巨大的成功和进展,并广泛应用于机器人导航、机器检测、目标识别、模型重建、运动跟踪、汽车安全、手术、三维电视、三维动漫、三维游戏、安全监控以及日常生活等各个领域<sup>[1~6]</sup>。然而,与人类自身的视觉功能相比,视觉理论与技术还很不健全,还处在一个十分不成熟的阶段,其技术上的难点主要是图像特征的检测与匹配<sup>[7]</sup>。一方面是因为,视觉重建过程可认为是成像过程的逆过程,由于成像过程中存在射影变换、成像混合、畸变与噪声等诸多因素影响,使得特征检测与立体匹配过程是不适定的和十分困难的。在存在灰度失真、几何失真、透视、旋转、缩放等失真以及噪声的情况下,立体匹配的过程是一项颇为困难的工作,目前还没有一个完全可靠的、通用的匹配方法,尤其是在失真和噪声比较严重的情况下,匹配过程显得更为困难。另一方面,现实世界中物体间的相互遮挡以及物体本身各部分间的遮挡,局部区域的缺失,也增加了图像特征检测与匹配的难度。目前,图像特征检测的主要方法有灰度自相关矩阵法、角点检测法、自适应非最大抑制法、特征重现性估计法、尺度不变法、旋转不变法和仿射不变法;而立体匹配主要采用基于区域、基于特征和基于相位的方法<sup>[8]</sup>。这些方法由于直接利用了图像的局部灰度和彩色信息,其结果易受几何失真、透视、旋转和缩放等失真以及噪声等因素的影响<sup>[9]</sup>。

为克服传统特征检测与立体匹配方法中存在的不足,本文采用二维图像信号在泊松尺度空间的建模分析技术和二维希尔伯特变换,代替传统的梯度法和 Gauss 微分法,从原始图像提取多尺度或多分辨的各向同性单演特征,包括几何特征(局部方向)、结构特征(局部振幅、局部相位)和彩色纹理特征(局部彩色向量、局部彩色单演相位)。以此为基础,模拟人眼视觉功能,建立了多模态的局部特征描述子,提出了特征空间的相似性距离测量函数及立体匹配价值函数,并在三维视觉重建中进行了定性和定量评估。

## 2 单演信号分析与局部特征检测

图像信号分析是计算机视觉任务的基础和关

键,图像处理和分析方法直接影响后续高级视觉过程的处理效果和质量。图像分析最理想的情况是无损失地融合图像中局部邻域内的所有原始信息,也就是要从二维图像中尽可能地提取能够真实、全面地反映目标本质属性的亮度、颜色、反射率以及几何、结构、纹理等特征的信息<sup>[10]</sup>。二维单演信号采用希尔伯特变换代替传统的梯度或微分运算,能够从原始图像提取多模式的单演局部特征,如几何特征(局部方向)、结构特征(局部振幅、局部相位)和彩色纹理特征(局部彩色信息、局部彩色单演相位)。单演信号具有稳定的局部几何特征和局部相位定位能力,基于单演相位的边缘识别精度可达亚像素级<sup>[11]</sup>。

### 2.1 尺度空间二维信号建模

根据傅里叶理论和泛函分析结果,每一个二维信号都可表示为任意方向一维余弦波的叠加<sup>[9]</sup>。利用泊松核滤波原始信号  $f$ , 形成一个经过平滑处理的局部信号模型

$$(p_s * f)(\mathbf{z}) = \sum_{v=1}^n a_v(\mathbf{z}, s) \cos[\langle \mathbf{z}, \bar{\mathbf{o}}_v(\mathbf{z}, s) \rangle + \phi_v(\mathbf{z}, s)], \quad (1)$$

式中  $\mathbf{z} = (x, y)$ ,  $*$  是卷积运算,  $\bar{\mathbf{o}}_v(\mathbf{z}, s) = [\cos \theta_v(\mathbf{z}, s), \sin \theta_v(\mathbf{z}, s)]^T$  表示法向方向,  $\langle \cdot, \cdot \rangle$  表示内积。泊松卷积核可表示为

$$p_s(\mathbf{z}) = \frac{s}{2\pi(s^2 + \|\mathbf{z}\|^2)^{3/2}}. \quad (2)$$

对某一确定的尺度参数  $s \in R_+$ , 泊松核在原始信号上的作用类似于低通滤波。为了得到感兴趣的频率范围,实际上通常使用泊松核的差(DOP)

$$p_{s_f, s_c}(\mathbf{z}) = p_{s_f}(\mathbf{z}) - p_{s_c}(\mathbf{z}), \quad (3)$$

式中  $s_c > s_f > 0$ ,  $s_c$  是粗尺度,  $s_f$  是精尺度。使用泊松核的差分对原始信号进行滤波后,可以得到原始信号的较小带宽的频谱。该局部信号模型允许在尺度空间内对任何纹理或结构,如直线、边缘、角点和连接,进行建模。特定的信号模型建立后,数学任务就是在每一个位置  $\mathbf{z} \in \Omega$  和每一个尺度空间参数  $s > 0$  恢复信号参数  $[\theta_v(\mathbf{z}, s), \phi_v(\mathbf{z}, s), a_v(\mathbf{z}, s)]$ 。不失一般性,信号模型在局部坐标系的原点  $|\mathbf{z}| = 0$  处变为

$$f_p(\mathbf{z}, s) = (p_s * f)(\mathbf{z}) = \sum_{v=1}^n a_v(\mathbf{z}, s) \cos \phi_v(\mathbf{z}, s). \quad (4)$$

图像分析时,可采用该方法表示图像中的直线、边缘、连接和角点。此时,图像信号处理的任务就是确定某一尺度参数  $s$  和某一位置  $\mathbf{z}$  的局部幅值

$a(\mathbf{z}, s)$ 、局部方向  $\theta_v(\mathbf{z}, s)$  和局部相位  $\phi(\mathbf{z}, s)$ 。对一维信号, 这些局部参数可采用希尔伯特变换通过传统的分析信号求解, 而对于内在结构维数为一维的二维图像信号, 这些局部参数则可采用里斯(Riesz)变换通过二维单演信号求解。Riesz 变换是一维希尔伯特变换在二维图像信号分析中的推广, 包括沿水平和垂直两个方向的希尔伯特变换。

## 2.2 二维单演信号及其局部特征

假设  $s: \mathbb{R} \rightarrow \mathbb{R}$  是一维实信号,  $f: \mathbb{R} \rightarrow \mathbb{R}_{2,0}$  是向量信号, 且有  $f(x) = s(x)e_2$ 。信号分析的目的是构建一个满足 Dirac 等式的函数, 其实部是实信号, 而虚部是实信号的希尔伯特变换。这实际上是一个黎曼问题, 即第二类边界求解问题。设  $D = e_1 \frac{\partial}{\partial x} + e_2 \frac{\partial}{\partial y}$  和  $\Delta = D^2$ , 有<sup>[12,13]</sup>

$$\begin{cases} \Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 & y > 0 \\ e_2 \frac{\partial u}{\partial y} = f(x) & y = 0 \end{cases} \quad (5)$$

方程组(5)的第一个方程是二维拉普拉斯方程, 其值严格限制在  $y > 0$  的半空间内; 第二个方程称为边界条件和基向量,  $e_2$  是复数的虚部。利用二维拉普拉斯方程的基本解, 可得

$$f_A(x, y) = p_1 * f(x, y) + p_1 * h_1 * f(x, y), \quad (6)$$

式中  $p_1 = \frac{y}{\pi(x^2 + y^2)}$  是一维泊松滤波核,  $h_1 = \frac{1}{\pi x} e_{12}$  是希尔伯特变换核, 变量  $y$  是尺度参数, 当  $y = 0$  时, 就可得到分析信号。

2001年, Felsberg 等<sup>[11]</sup>采用 Riesz 变换将一维分析信号推广到二维分析信号, 提出了二维图像的单演分析理论与方法。假设  $s: \mathbb{R}^2 \rightarrow \mathbb{R}$  是二维实信号,  $f: \mathbb{R}^2 \rightarrow \mathbb{R}_{3,0}$  同样是一个向量信号,  $\{e_i\} (i = 1, 2, 3)$  是  $\mathbb{R}^3$  的正交基, 因此有  $f(x, y) = f_3(x, y)e_3$ 。根据  $z > 0$  半空间内三维拉普拉斯方程和第二类边界条件, 可定义单演分析信号为

$$f_M(x, y, s) = p_2 * f(x, y) + p_2 * h_x * f(x, y) + p_2 * h_y * f(x, y), \quad (7)$$

式中  $p_2 = \frac{s}{2\pi(x^2 + y^2 + s^2)^{3/2}}$  是二维泊松滤波核,

$h_R = (h_x, h_y) = \left[ \frac{x e_1}{2\pi(x^2 + y^2)^{3/2}}, \frac{y e_2}{2\pi(x^2 + y^2)^{3/2}} \right]$  是 Riesz 变换核,  $s$  是尺度参数。现在假设  $f_M(x, y, s) = f_p(x, y, s) + f_x(x, y, s) + f_y(x, y, s)$  和  $f_q(x, y, s) = \begin{bmatrix} f_x \\ f_y \end{bmatrix} (x, y, s)$ , 对内在结构维数为 1 的图像信号, 可

得单演信号的三个分量为

$$(f_p, f_x, f_y) = a(\cos \phi, \sin \phi \cos \theta, \sin \phi \sin \theta), \quad (8)$$

式中  $a(x, y, s), \theta(x, y, s), \phi(x, y, s)$  分别代表局部幅值、局部方向和局部相位

$$a(x, y, s) = \sqrt{f_p^2(x, y, s) + f_x^2(x, y, s) + f_y^2(x, y, s)}, \quad (9)$$

$$\theta(x, y, s) = \arctan[f_y(x, y, s)/f_x(x, y, s)], \quad (10)$$

$$\phi(x, y, s) = \arctan\left[\frac{\sqrt{f_x^2(x, y, s) + f_y^2(x, y, s)}}{f_p(x, y, s)}\right]. \quad (11)$$

由于相位的正负与信号的方向有关, 因此相位的真实值为

$$\begin{aligned} \phi(x, y, s) &= \frac{f_q(x, y, s)}{|f_q(x, y, s)|} \times \\ &\arg[f_p(x, y, s) + i|f_q(x, y, s)|]. \end{aligned} \quad (12)$$

## 2.3 彩色单演信号及其局部彩色相位

实际上自然图像绝大部分是彩色的, 同时, 彩色图像也携带了更多的视觉信息, 在泊松尺度空间内提取彩色图像的彩色纹理、彩色相位等信息, 对目标和场景的高保真三维重建具有更加重要的意义。因此, 可将灰度级的单演信号分析扩展到彩色级单演图像处理, 并根据 Clifford 代数的几何积<sup>[7]</sup>提取局部彩色相位特征<sup>[13]</sup>。对灰度图像, 二维单演信号是一个向量信号  $f_M(\mathbf{z}, s)$ 。对彩色图像, 有三个通道  $(R, G, B)$ , 而每一个通道都可得到一个单演信号

$$\begin{aligned} f_M^c(\mathbf{z}, s) &= [f_p^c(\mathbf{z}, s), f_x^c(\mathbf{z}, s), f_y^c(\mathbf{z}, s)]^T, \\ c &= R, G, B \end{aligned} \quad (13)$$

设  $f_C = f_M^R + f_M^G + f_M^B$ , 可得

$$\begin{aligned} f_C &= h_p^R * R + h_p^G * G + h_p^B * B + h_R^R * h_R * R + \\ &h_p^G * h_R * G + h_p^B * h_R * B, \end{aligned} \quad (14)$$

式中  $h_p^i (i = R, G, B)$  是泊松核,  $h_R(z) = h^{(1)}(z)$  是一次广义希尔伯特变换核, 即 Riesz 变换核。

根据 Clifford 代数的几何空间关系<sup>[12]</sup>, 欧几里得向量空间  $\mathbb{R}^n$  (通常用  $\mathbb{R}_{n,0}$  表示) 中的两个向量  $\mathbf{a}$  和  $\mathbf{b}$  的积为

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \wedge \mathbf{b}, \quad (15)$$

式中  $\mathbf{a} \cdot \mathbf{b}$  是  $\mathbf{a}$  和  $\mathbf{b}$  的内积,  $\mathbf{a} \wedge \mathbf{b}$  是  $\mathbf{a}$  和  $\mathbf{b}$  的楔积(或外积), 并且是一个双向量。这个积通常称为几何积, 其中内积是对称的, 而楔积是斜对称的。彩色单演信号  $f_C$  实际上属于  $\mathbb{R}^5$  的向量空间。如果用  $\{e_i, i = 1, \dots, 5\}$  表示  $\mathbb{R}_{5,0}$  的基向量, 则  $\{e_1, e_2\}$  表示图像的两个坐标,  $\{e_3, e_4, e_5\}$  代表彩色分量  $\{R, G, B\}$  (当

然,也可是其他彩色空间,如 CIE  $L^* a^* b^*$  色彩空间)的基向量。为更直观地表示,令  $(f_3, f_4, f_5) = (R, G, B)$  和  $(h_p^3, h_p^4, h_p^5) = (h_p^R, h_p^G, h_p^B)$ , 重写彩色单演信号  $f_c$  的表达式为

$$f_c = A_1 e_1 + A_2 e_2 + A_3 e_3 + A_4 e_4 + A_5 e_5, \quad (16)$$

式中

$$A_1 = h_p^3 * h_{R_x} * f_3 + h_p^4 * h_{R_x} * f_4 + h_p^5 * h_{R_x} * f_5, \quad (17)$$

$$A_2 = h_p^3 * h_{R_y} * f_3 + h_p^4 * h_{R_y} * f_4 + h_p^5 * h_{R_y} * f_5, \quad (18)$$

$$A_i = h_p^i * f_i, i = 3, 4, 5. \quad (19)$$

因此,彩色单演信号包含了两种信息:1)  $A_1$  和  $A_2$  对应于 Riesz 变换在水平和垂直方向平滑滤波后的结构;2)  $A_3, A_4$  和  $A_5$  对应于由  $\{e_3, e_4, e_5\}$  基向量张成的空间中的彩色值  $(f_3, f_4, f_5)$  的平滑滤波值。

如果令  $\mathbf{V} = ue_1 + ve_2 + ae_3 + be_4 + ce_5 \in \mathbb{R}_{5.0}$  是一个包含图像结构信息  $(u, v)$  和彩色信息  $(a, b, c)$  的参考向量,那么,在  $\mathbb{R}_{5.0}$  中的几何积  $f_c \mathbf{V}$  可表示为

$$f_c \mathbf{V} = \langle f_c \mathbf{V} \rangle_0 + \langle f_c \mathbf{V} \rangle_2, \quad (20)$$

式中 0 级部分  $\langle f_c \mathbf{V} \rangle_0$  是标量部分,2 级部分  $\langle f_c \mathbf{V} \rangle_2$  是双向量部分。根据 Clifford 代数,可表示为  $f_c \mathbf{V} = \langle f_c \mathbf{V} \rangle_0 + \frac{\langle f_c \mathbf{V} \rangle_2}{|\langle f_c \mathbf{V} \rangle_2|} |\langle f_c \mathbf{V} \rangle_2|$ , 因此可把  $f_c \mathbf{V}$  看作是一个复数

$$f_c \mathbf{V} = \langle f_c \mathbf{V} \rangle_0 + i |\langle f_c \mathbf{V} \rangle_2|. \quad (21)$$

这意味着  $f_c \mathbf{V}$  是一个旋量,它相当于在由双向量  $\langle f_c \mathbf{V} \rangle_2$  张成的平面内的旋转。这样,局部彩色相位就是旋转角  $\varphi_c$ 。

$$\varphi_c = \arg(f_c \mathbf{V}) = \arctan\left(\frac{|\langle f_c \mathbf{V} \rangle_2|}{\langle f_c \mathbf{V} \rangle_0}\right), \quad (22)$$

该彩色相位描述了  $f_c$  和给定参考向量  $\mathbf{V}$  在  $\mathbb{R}_{5.0}$  中的角距离。它给出了图像像素(包含彩色和结构信息)与参考向量(包含选择的彩色和结构)的相关性测量。

### 3 多模式局部特征描述及其相似性测量

由单演信号分析方法,在泊松尺度空间内提取了原始图像的几何特征(局部方向)、结构特征(局部振幅、局部相位)和彩色纹理特征(局部彩色信息、局部彩色单演相位)。为建立这些特征之间的内在联系,模拟人眼视觉对自然特征的处理和描述方法来建立特征间的相互联系<sup>[14]</sup>,实现特征处理中的相互融合。为直观而紧凑地表达图像的局部信号特征,首先定义一个多模式单演特征描述子

$$\varepsilon = (\mathbf{X}, \hat{\phi}, \varphi_c, \mathbf{C}). \quad (23)$$

由于局部特征检测时利用的是图像局部区域信息,因而该特征描述子实际上表示了该局部区域内的特征信息,区域的大小可定义为一个  $n \times n$  ( $n \in \mathbb{R}^+, n > 1$ ) 大小的局部面片。 $\mathbf{X} = (x, y)$  代表局部面片的中心坐标,  $\hat{\phi}$  是彩色图像转换为灰度图像时的瞬时相位,  $\varphi_c$  是局部彩色相位,  $\mathbf{C} = (R, G, B)$  是 RGB 彩色空间中的彩色值。

基于上述多模式特征描述子,可以建立立体图像在特征空间的相似性测量函数,包括相同大小局部面片内立体图像的瞬时相位、彩色相位和彩色值的绝对差

$$e(\varepsilon_l, \varepsilon_r) = \alpha d_\phi(\hat{\phi}_l, \hat{\phi}_r) + \beta d_\varphi(\varphi_l, \varphi_r) + \gamma d_C(\mathbf{C}_l, \mathbf{C}_r) = \alpha |\hat{\phi}_l - \hat{\phi}_r| + \beta |\varphi_{cl} - \varphi_{cr}| + \gamma \sqrt{\sum_{i \in \{R, G, B\}} (c_l^i - c_r^i)^2}, \quad (24)$$

式中  $l, r$  分别代表立体图像对中的左图像和右图像,  $d_\phi \in [0, \pi]$  是瞬时相位  $[\phi \in [-\pi, \pi]]$  特征的距离,  $d_\varphi \in [0, \pi]$  是局部彩色相位  $[\varphi_c \in [0, \pi]]$  特征的距离,  $d_C \in [0, \sqrt{3}]$  是图像在 RGB 彩色空间中局部彩色值的距离,且有  $\mathbf{C} \in [0, 1] \times [0, 1] \times [0, 1]$  和  $C_R + C_G + C_B = 1$ 。 $\alpha, \beta, \gamma$  是特征间距离测量的权重系数,且有  $\alpha, \beta \in [0, 1]$  和  $\gamma \in [0, 0.5]$ 。为了在自然图像中得到更准确和连续一致的特征测量,采用自适应权重方法计算局部面片内的相似性测量误差和。根据完形原则,权重系数和特征空间相似性测量误差和分别为<sup>[15]</sup>

$$w(x, y, m, n) = \exp\left[-\left(\frac{\Delta c_{x, y, m, n}}{\gamma_c} + \frac{\Delta q_{x, y, m, n}}{\gamma_g}\right)\right], \quad (25)$$

$$E(x_l, y_l, x_r, y_r) = \frac{\sum_{m, n \in [-p, p]} w_l(x_l, y_l, m, n) w_r(x_r, y_r, m, n) e(\varepsilon_l, \varepsilon_r)}{\sum_{m, n \in [-p, p]} w(x, y, m, n) w_r(x_r, y_r, m, n)}, \quad (26)$$

式中  $(x_l, y_l)$  和  $(x_r, y_r)$  分别是左图像和右图像的对比区域,  $(m, n)$  是局部面片内像素的相对位置,  $w(x, y, m, n)$  代表邻域像素  $(x + m, y + n)$  的位置,  $\Delta c_{x, y, m, n}$  是像素  $(x, y)$  和  $(x + m, y + n)$  间的彩色值之差,  $\Delta q_{x, y, m, n}$  是像素  $(x, y)$  和  $(x + m, y + n)$  间的欧氏距离,  $\gamma_c$  和  $\gamma_g$  是用户确定的经验参数,  $E(x_l, y_l, x_r, y_r)$  是左图像中坐标为  $(x_l, y_l)$  的局部面片与右图像中坐标为  $(x_r, y_r)$  的局部面片的特征相似性测量误差。立体图像间的特征相似性测量函数确定后,接下来就是最小化价值函数。为了减少噪声影

响并提高匹配效率,立体匹配中采用多尺度空间和局部最优搜索方法来优化立体视差图。

## 4 实验及结果

为定性和定量评估多模式局部特征描述子及其相似性测量函数,完成了三个方面的实验:1)合成立体图像上视觉估计实验,定量评估和对比研究算法的匹配精度;2)在噪声和亮度变化条件下,立体匹配的稳健性和稳定性测试实验;3)自然目标(场景)的形貌和运动重建实验,定性评估算法的可行性。

### 4.1 视差估计和对比实验

实验中,选择文献[16]中的 Cloth1 立体图像对进行立体匹配实验,左图像、右图像和真实视差图分别如图 1(a)~(c)所示。为清楚地描述视差估计全

过程,下面给出算法的具体实现过程(立体图像的对应极线已经水平):

第一步:立体图像对的单演信号分析和局部特征提取。左、右图像分别进行最大尺度  $S_{\max} = 3$  的单演信号分析,由(9)~(11)式在泊松尺度空间中提取每个尺度的局部振幅、局部方向和瞬时相位特征,由(12)式确定瞬时相位的真实值。图 2(a)~(c)分别显示了左图像在三个不同尺度时的瞬时相位大小。同时,根据 2.3 节内容,分别计算左、右图像在三个不同尺度的彩色相位,其中参考向量选择  $\mathbf{V} = (e_3 + e_4 + e_5) / \sqrt{3}$ ,其局部几何结构  $(u, v) = (0, 0)$ ,单位彩色向量值为  $(a, b, c) = (1, 1, 1) / \sqrt{3}$ 。图 3(a)~(c)分别显示了左图像在三个不同尺度时的彩色相位图。

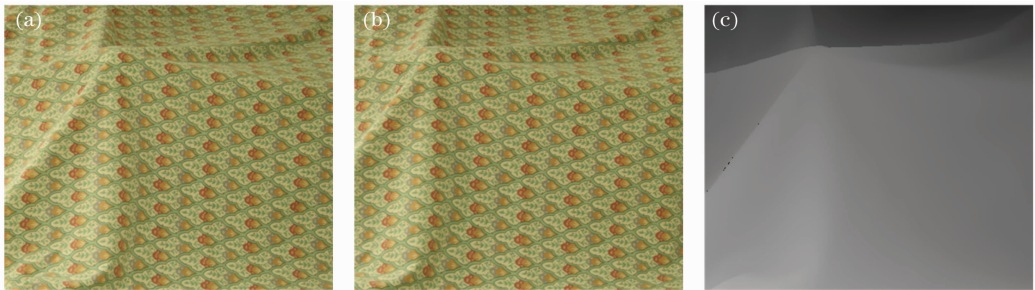


图 1 立体图像对。(a)左图像;(b)右图像;(c)真实视差图

Fig. 1 Stereo images. (a) Left image; (b) right image; (c) ground disparity map

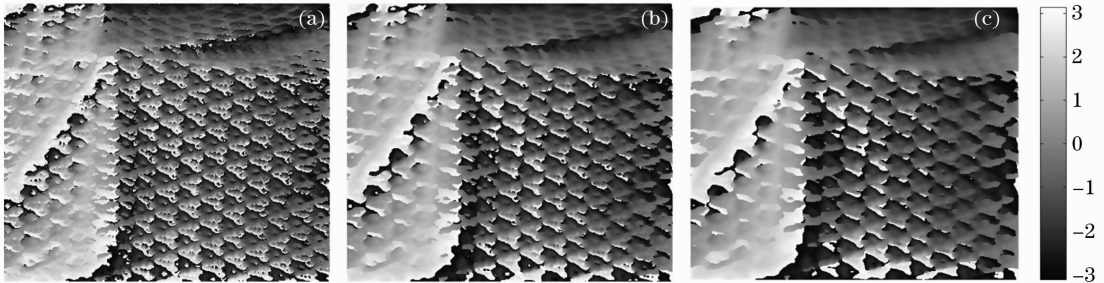


图 2 局部瞬时相位图。(a)第 1 尺度;(b)第 2 尺度;(c)第 3 尺度

Fig. 2 Local instantaneous phase. (a) First scale; (b) second scale; (c) third scale

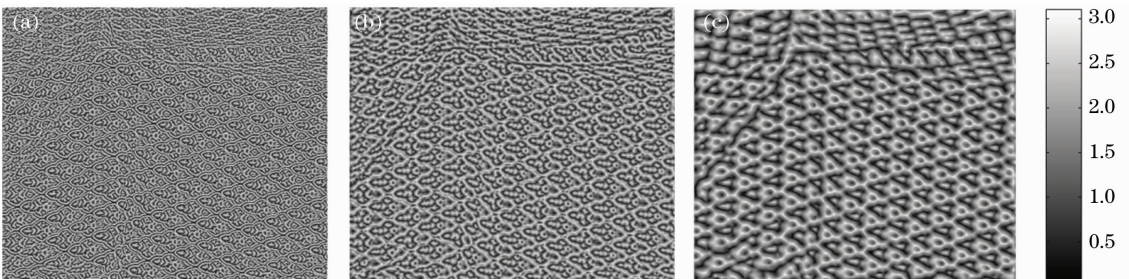


图 3 局部彩色相位图。(a)第 1 尺度;(b)第 2 尺度;(c)第 3 尺度

Fig. 3 Local color phase. (a) First scale; (b) second scale; (c) third scale

第二步:估计尺度空间最精层( $s = S_{\max}$ )的视差。将视差搜索范围设置为该层最大值  $d_{\max}$ ,忽略匹配价值函数的光滑误差项和连续误差项,采用最优局部搜索方法最小化匹配价值函数,估计视差的初始值。同时,按照文献的方法,考虑匹配光滑误差项和连续误差项,采用全局优化方法最小化视差图。

第三步:由粗到精的视差估计( $1 \leq s < S_{\max}$ )。由尺度空间的  $s+1$  层视差图初始化该  $s$  层的视差图,设置视差搜索范围为邻域视差的最大值,采用全局优化方法最小化视差图,重复该步直到  $s=1$ 。

第四步:视差图后续处理。按照匹配唯一性和顺序性原则对视差图进行滤波、遮挡、缺失等处理,验证匹配正确性。

第五步:子像素视差估计。采用双线性插值法,由最精层( $s=1$ )的局部相位计算子像素的视差图。图 4(a)~(d) 分别显示了( $\alpha=0, \beta=0, \gamma=0.5$ ), ( $\alpha=1, \beta=0, \gamma=0$ ), ( $\alpha=0, \beta=1, \gamma=0$ ) 和( $\alpha=$

$1, \beta=1, \gamma=0.5$ ) 时的视觉图。

从图 4 中,可以算出当( $\alpha=1, \beta=1, \gamma=0.5$ ) 时视差图的质量明显优于其他情况。该视差估计算法及过程通过小波金字塔来实现,实现了由粗到精,再到子像素的视差估计,不仅能够消除噪声、遮挡、缺失等因素对匹配的影响,而且计算时间短,匹配效率高。为对算法进行定量的评估,还进行了立体匹配对比实验。从文献中选择了一些有竞争力的匹配算法,包括绝对和差法(SAD)<sup>[17]</sup>、平方和差(SSD)<sup>[17]</sup>、图分割法(GC)<sup>[17]</sup>、离散小波变换(DWT)<sup>[18]</sup>、复小波变换(CWT)<sup>[18]</sup>和四元数小波(QWT)<sup>[19]</sup>。对比实验中,统计了所有立体视差图的均方根误差(RMSE)和误匹配百分比(PBD),统计结果如表 1 所示。从表中可以看出,当  $\alpha=1, \beta=1, \gamma=0.5$  时,取得了最好立体匹配结果:RMSE 为 1.82, PBD 为 0.09。

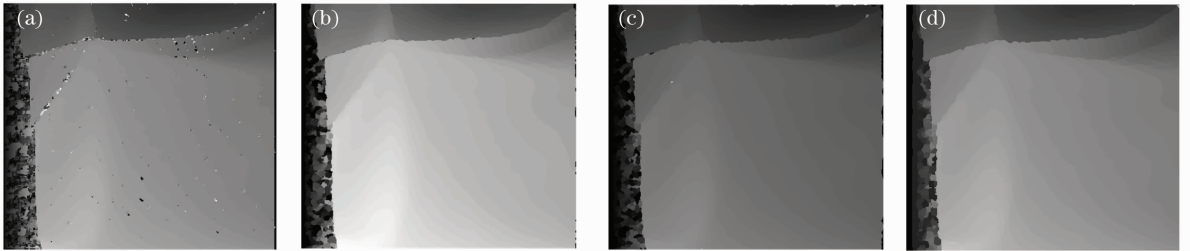


图 4 估计的视差图。(a)  $\alpha=0, \beta=0, \gamma=0.5$ ; (b)  $\alpha=1, \beta=0, \gamma=0$ ; (c)  $\alpha=0, \beta=1, \gamma=0$ ; (d)  $\alpha=1, \beta=1, \gamma=0.5$   
Fig. 4 Estimated disparity map. (a)  $\alpha=0, \beta=0, \gamma=0.5$ ; (b)  $\alpha=1, \beta=0, \gamma=0$ ; (c)  $\alpha=0, \beta=1, \gamma=0$ ; (d)  $\alpha=1, \beta=1, \gamma=0.5$

表 1 不同立体匹配算法的实验结果对比  
Table 1 Experimental results for different matching algorithms

| Algorithm                       | RMSE | PBD  |
|---------------------------------|------|------|
| SAD                             | 4.23 | 0.37 |
| SSD                             | 3.78 | 0.31 |
| GC                              | 2.25 | 0.15 |
| CWT                             | 2.69 | 0.19 |
| DWT                             | 3.17 | 0.29 |
| QWT                             | 2.53 | 0.18 |
| $\alpha=0, \beta=0, \gamma=0.5$ | 3.24 | 0.30 |
| $\alpha=1, \beta=0, \gamma=0$   | 2.61 | 0.20 |
| $\alpha=0, \beta=1, \gamma=0$   | 1.98 | 0.13 |
| $\alpha=1, \beta=1, \gamma=0.5$ | 1.82 | 0.09 |

#### 4.2 噪声和亮度变化的稳健性实验

通过实验研究算法对噪声和光照变化的稳健性。第一步,分别将  $\sigma=0.05, 0.10, 0.15, 0.20, 0.25, 0.30$  的高斯噪声增加到彩色立体图像对,使

用 4 种不同的( $\alpha, \beta, \gamma$ )值估计视差图,并统计每个视差图的 RMSE 和 PBD。第二步,分别将  $I=-30, -20, -10, 10, 20, 30$  的亮度值增加到右图像,分别统计 4 种不同的( $\alpha, \beta, \gamma$ )值估计的视差图的 RMSE 和 PBD。全部实验结果如表 2 所示。分析统计结果可知,当  $\alpha=1, \beta=1, \gamma=0.5$  时不易受噪声和亮度变化的影响,而当  $\alpha=0, \beta=0, \gamma=0.5$  时视差图的结果随噪声和亮度的波动较大,稳健性较差。

#### 4.3 自然目标的三维重建实验

实验中,在多视图立体三维重建中利用多模式特征相似性测量函数来重建目标表面的密集点云。首先,使用两台已标定的立体相机(AVT Guppy F146C, 1024 pixel  $\times$  768 pixel)采取自然目标(一个静止彩色纸杯和一张变形打印纸)的立体图像或立体图像序列。其次,采用 SURF 特征检测方法<sup>[20]</sup>提取所有图像的特征点,并使用极线约束和  $\alpha=1,$

表 2 噪声和亮度变化的稳健性实验结果

Table 2 Experimental results for noise/brightness change

| Noise/brightness change | RMSE                            | PBD  | RMSE                          | PBD  | RMSE                          | PBD  | RMSE                            | PBD  |
|-------------------------|---------------------------------|------|-------------------------------|------|-------------------------------|------|---------------------------------|------|
|                         | $\alpha=0, \beta=0, \gamma=0.5$ |      | $\alpha=1, \beta=0, \gamma=0$ |      | $\alpha=0, \beta=1, \gamma=0$ |      | $\alpha=1, \beta=1, \gamma=0.5$ |      |
| None                    | 3.24                            | 0.30 | 2.61                          | 0.20 | 1.98                          | 0.13 | 1.82                            | 0.09 |
| $\sigma=0.05$           | 3.46                            | 0.34 | 2.73                          | 0.22 | 1.99                          | 0.14 | 1.89                            | 0.10 |
| $\sigma=0.10$           | 3.73                            | 0.37 | 2.89                          | 0.24 | 2.13                          | 0.15 | 1.96                            | 0.11 |
| $\sigma=0.15$           | 3.98                            | 0.41 | 2.99                          | 0.26 | 2.18                          | 0.16 | 2.04                            | 0.12 |
| $\sigma=0.20$           | 4.41                            | 0.46 | 3.17                          | 0.28 | 2.29                          | 0.18 | 2.12                            | 0.13 |
| $\sigma=0.25$           | 5.09                            | 0.51 | 3.29                          | 0.31 | 2.41                          | 0.20 | 2.19                            | 0.15 |
| $\sigma=0.30$           | 6.12                            | 0.59 | 3.50                          | 0.35 | 2.53                          | 0.22 | 2.26                            | 0.18 |
| $I=-30$                 | 6.47                            | 0.62 | 2.86                          | 0.28 | 2.30                          | 0.19 | 2.07                            | 0.14 |
| $I=-20$                 | 5.35                            | 0.54 | 2.67                          | 0.25 | 2.12                          | 0.16 | 1.96                            | 0.12 |
| $I=-10$                 | 4.05                            | 0.41 | 2.43                          | 0.23 | 1.97                          | 0.14 | 1.89                            | 0.10 |
| $I=10$                  | 4.02                            | 0.39 | 2.39                          | 0.22 | 1.96                          | 0.13 | 1.86                            | 0.10 |
| $I=20$                  | 5.10                            | 0.49 | 2.46                          | 0.24 | 2.14                          | 0.15 | 1.94                            | 0.11 |
| $I=30$                  | 6.59                            | 0.61 | 2.72                          | 0.27 | 2.27                          | 0.18 | 2.09                            | 0.14 |

$\beta=1, \gamma=0.5$  的特征相似性测量函数匹配提取的特征。根据这些初始匹配,可以按立体视觉原理重建目标表面的稀疏三维点云。然后,以每个稀疏点为中心,在其邻域内进一步膨胀并重建更多的表面三维点,邻域内的左右图像匹配仍使用本文提出的特征相似性测量函数,并通过目标表面连续性约束和彩色一致性约束对重建的稀疏点云进行滤波。最后,将该膨胀、重建和滤波过程进行多次重复,直到

获得稠密的表面点云。三维重建过程中,以左图像作为参考图像,并试图在左图像的每个  $2 \times 2$  局部图片上重建一个三维表面点。对静止的彩色纸杯,用两台摄像机从不同方位共采集了 16 张图像,每张图像的畸变和位姿已标定。使用已标定的 16 张图像,重建了纸杯表面的稠密点云及其三维网络模型,并在网格模型上融合了表面纹理和彩色信息。图 5 分别显示了第 1,7,13 个视图的原始左图像(a)~(c)、

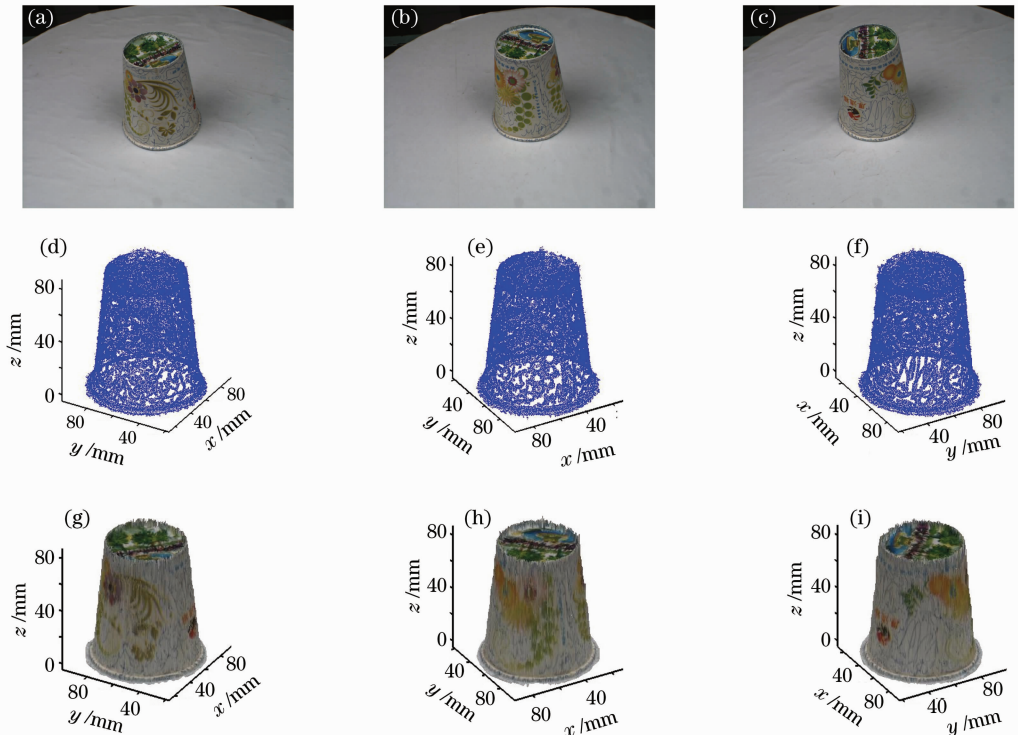


图 5 静止彩色纸杯的三维形貌重建实验

Fig. 5 Experimental results for static color cup

稠密表面点云(d)~(f)和特征融合模型(g)~(i)。对变形的彩色打印纸,使用两台摄像机以 7 frame/s 的速率采集纸张变形过程的立体序列,并重建其时变的三维形变过程。图 6 分别显示了第 0, 20, 40, 60, 80 和 100 frame 时的原始左图像(a)~(f)和形变面貌(g)~(l)。同时,为实现对比和定性评估,还采用平方和差的相似性测量函数<sup>[12]</sup>重建了静止彩色纸杯表面的稠密点云和变形纸张的时变形貌。为直观地对比,表 3 中给出了静止彩色纸杯重建过程中的一些结果。在三维重建中,运用 VC++6.0 软件编程,并使用 inter(R) core(TM) duo CPU T8100 @2.10 GHz 计算平台,静止的彩色纸杯的三维重建总时间为 456 s,变形的彩色打印纸第一帧

的重建时间为 632 s,后续每帧重建的平均时间为 318 s。

表 3 三维重建实验结果对比

Table 3 Experimental results for three-dimensional vision reconstruction

| Similarity function                         | Monogenic features | SSD   |
|---|--------------------|-------|
| Feature points each image                   | 1632               | 1632  |
| PBD   | 0.18               | 0.38  |
| Seed patches                                | 4699               | 3578  |
| Patches after first expansion               | 102417             | 8796  |
| Patches after first filter                  | 99336              | 8169  |
| Percentage of non-reconstructed patches / % | 7.26               | 16.32 |

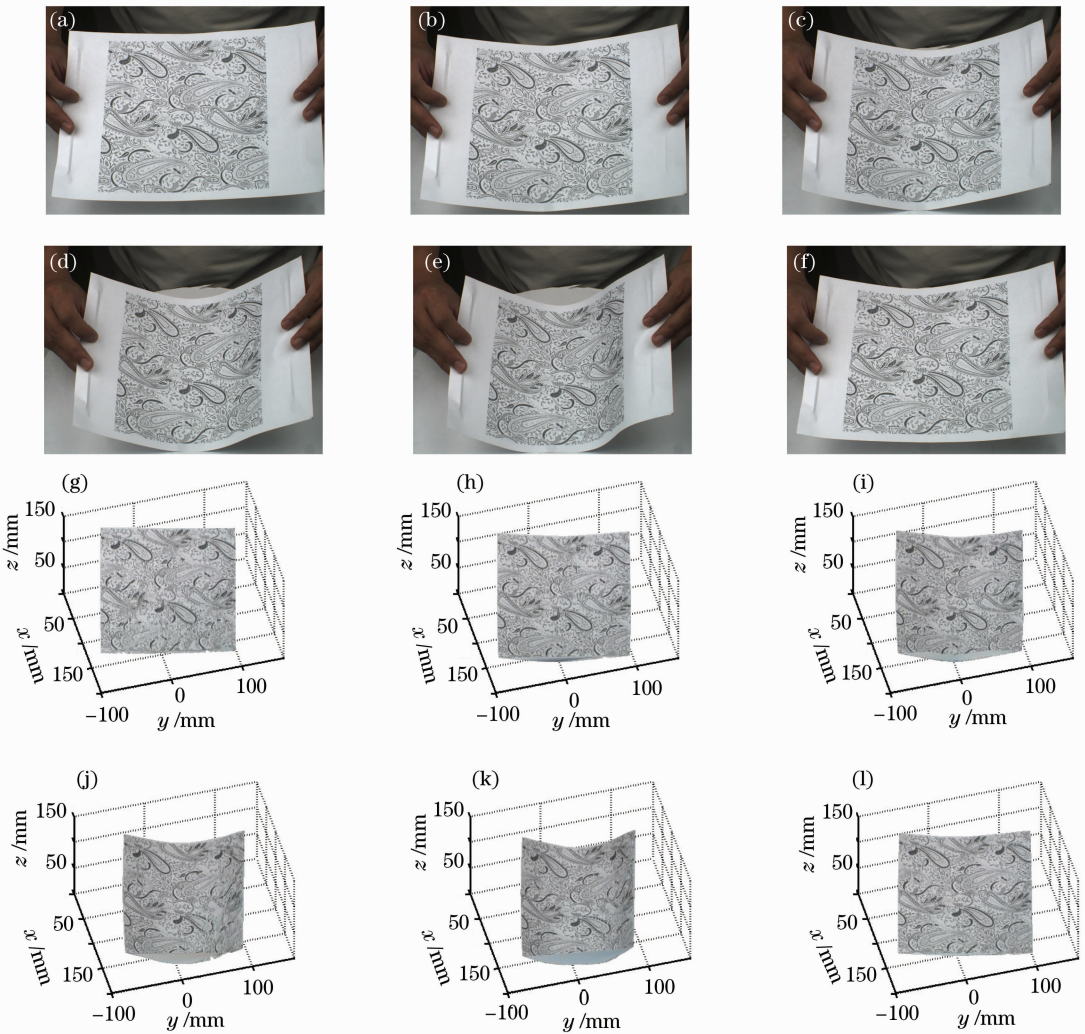


图 6 变形纸张的三维形变重建实验

Fig. 6 Experimental results for deformed paper

## 5 结 论

提出了一种基于多尺度单演特征检测与匹配的

三维视觉测量新方法。该方法首先采用泊松滤波算子和 Riesz 变换提取图像的多模式单演特征,包括



局部幅值、局部方向、局部相位和彩色相位;然后模拟人眼视觉功能建立多模态的局部特征描述子及特征空间的相似性距离测量函数;最后将该相似性测量应用于三维视觉重建的立体匹配中,并进行了定性和定量评估。实验结果表明,该方法较好地克服了传统三维视觉测量易受亮度变化、尺度变化、旋转变化及噪声影响的缺点,得到了高精度和稳健性的立体匹配和三维重建结果。

### 参 考 文 献

- 1 Zhang Xiubin, Ying Junhao. Visual Intelligent Measurement [M]. Beijing: Science Press, 2009  
张秀彬, 应俊豪. 视觉智能检测[M]. 北京: 科学出版社, 2009
- 2 Zhou Xiang, Zhao Hong. Three-dimensional profilometry based on Mexican hat wavelet transform [J]. *Acta Optica Sinica*, 2009, **29**(1): 198~202  
周翔, 赵宏. 基于 Mexican hat 小波变换的三维轮廓术[J]. 光学学报, 2009, **29**(1): 198~202
- 3 Zhou Xiang, Zhao Hong. Phase profilometry using discrete wavelet transform [J]. *Acta Optica Sinica*, 2009, **29**(6): 1562~1569  
周翔, 赵宏. 使用离散小波变换的相位轮廓术[J]. 光学学报, 2009, **29**(6): 1562~1569
- 4 Wu Di, Lu Naiguang, Ouyang Jing. A structured light three-dimensional measurement method based on unconstrained system configuration [J]. *Acta Optica Sinica*, 2008, **28**(3): 482~486  
吴迪, 吕乃光, 欧阳京. 基于无约束系统的结构光三维测量方法[J]. 光学学报, 2008, **28**(3): 482~486
- 5 Liu Tianliang, Luo Limin. Segmentation-based stereo matching algorithm with variable support and disparity estimation [J]. *Acta Optica Sinica*, 2009, **29**(4): 1002~1006  
刘天亮, 罗立民. 一种基于分割的可变权值和视差估计的立体匹配算法[J]. 光学学报, 2009, **29**(4): 1002~1006
- 6 Zhang Weiguang, Zhao Hong, Zhang Qi *et al.*. Calibration method for three-dimensional measurement system based on linear-structure light [J]. *Chinese J. Lasers*, 2009, **36**(1): 182~188  
张维光, 赵宏, 张琦等. 线结构光三维轮廓测量系统的标定方法[J]. 中国激光, 2009, **36**(1): 182~188
- 7 吴立德. 计算机视觉[M]. 上海: 复旦大学出版社, 1993
- 8 R. Szeliski. Computer Vision: Algorithms and Applications [M]. Berlin Heidelberg: Springer-Verlag; 2009
- 9 J. Li, H. Zhao, X. Zhou *et al.*. Robust stereo image matching using a two-dimensional monogenic wavelet transform [J]. *Opt. Lett.*, 2009, **34**(22): 3514~3516
- 10 L. Wietzke, G. Sommer. The signal multi-vector [J]. *J. Math. Imaging Vis.*, 2010, **37**(2): 132~150
- 11 M. Felsberg, G. Sommer. The monogenic signal [J]. *IEEE Trans. Signal Process.*, 2001, **49**(12): 3136~3144
- 12 J. Vince. The Geometric Product, Geometric Algebra for Computer Graphics [M]. London: Springer, 2008. 79~124
- 13 G. Demarcq, L. Mascarilla, P. Courtellemont. The color monogenic signal: a new framework for color image processing [C]. IEEE International Conference on Image Processing (ICIP), 2009, 481~484
- 14 N. Kruger, M. Felsberg. An explicit and compact coding of geometric and structural image information applied to stereo processing [J]. *Pattern Recognition Letters*, 2004, **25**(18): 849~863
- 15 K. J. Yoon, I. S. Kweon. Adaptive support-weight approach for correspondence search [J]. *PAMI*, 2005, **28**(4): 650~656
- 16 www.middlebury.edu/stereo/data/scenes2006/FullSize/Cloth1/
- 17 D. Scharstein, R. Szeliski. A taxonomy evaluation of dense two-frame stereo correspondence algorithms [J]. *International J. Computer Vision*, 2002, **47**(7): 1~35
- 18 A. Bhatti, S. Nahavandi. Stereo image matching using wavelet scale-space representation [C]. Proceedings of the International Conference on Computer Graphics, Imaging and Visualisation (CGIV'06), 2006: 267~272
- 19 Y. Xu, X. Yang. Cooperative stereo matching using quaternion wavelets and top-down segmentation [C]. IEEE ICME, 2007, 1954~1957
- 20 H. Bay, A. Ess, T. Tuytelaars *et al.*. Speeded-up robust features [J]. *Computer Vision and Image Understanding*, 2008, **110**(3): 346~359