

## 基于韦伯感知和导引滤波分层聚合快速立体图像匹配

刘天亮<sup>\*①②</sup> 戴修斌<sup>③</sup> 朱秀昌<sup>①②</sup> 罗立民<sup>④</sup>

<sup>①</sup>(江苏省图像处理与图像通信重点实验室 南京 210003)

<sup>②</sup>(南京邮电大学通信与信息工程学院 南京 210003)

<sup>③</sup>(南京邮电大学地理与生物信息学院 南京 210003)

<sup>④</sup>(东南大学影像科学与技术实验室 南京 210096)

**摘要:** 该文利用韦伯定律和导引滤波提出基于代价分层聚合的快速立体匹配方法。首先提取立体图像对各彩色通道Weber描述符并初始化匹配代价。利用导引滤波增强匹配代价并提取视差候选;利用候选子集联合空间离散采样与自适应支持权重实现分层代价聚合;据此快速择优选取初始视差。视差求精中采用改进型双边滤波和对称映射后处理有效改善初始视差图中歧义区域。实验表明,该文方法能有效消除匹配歧义,获得分段平滑、高精度稠密视差;结构简单、快速高效且对光照变化具有鲁棒性。

**关键词:** 图像处理; 立体视觉; 韦伯定律; 导引滤波; 分层聚合

中图分类号: TN911.73

文献标识码: A

文章编号: 1009-5896(2012)04-0992-05

DOI: 10.3724/SP.J.1146.2011.00993

## Hierarchical Aggregation Fast Stereo Image Matching Based on Weber Perception and Guided Filtering

Liu Tian-liang<sup>\*①②</sup> Dai Xiu-bin<sup>③</sup> Zhu Xiu-chang<sup>①②</sup> Luo Li-min<sup>④</sup>

<sup>①</sup>(*Jiangsu Provincial Key Lab of Image Processing and Image Communication, Nanjing 210003, China*)

<sup>②</sup>(*College of Telecommunications and Information Engineering, Nanjing University  
of Posts and Telecommunications, Nanjing 210003, China*)

<sup>③</sup>(*College of Geographic and Biologic Information, Nanjing University  
of Posts and Telecommunications, Nanjing 210003, China*)

<sup>④</sup>(*Lab of Image Science and Technology, Southeast University, Nanjing 210096, China*)

**Abstract:** This paper presents a hierarchical cost aggregation-based fast stereo image matching method based on Weber's law and guided filtering. Weber local descriptors for each color channel are firstly extracted from stereo pairs, and raw matching costs between the images are initialized by the descriptors. The matching costs are enhanced with guided filtering to extract the subsets of disparity candidates. Joint spatial discrete sampling and adaptive support weight are utilized to implement hierarchical cost aggregation on the candidate subsets. Then initial disparities from the subsets are selected fast and optimally. Modified bilateral filtering and symmetric warping-based post-processing are sequentially exploited in disparity refining to improve effectively ambiguous regions of initial disparity maps. The experimental results indicate that this proposed technique can obtain piecewise smooth, accurate and dense disparity map while eliminating effectively matching ambiguity. Being concise, fast and high efficiency, and it is robust to illumination change.

**Key words:** Image processing; Stereo vision; Weber's law; Guided filtering; Hierarchical aggregation

### 1 引言

双目立体匹配是实现从2维图像对3维视觉感知的基础,在机器视觉、无人车自主导航、3维建模及

3维视频编码等领域应用广泛。文献[1]全面概述并分类总结了一些代表性的立体匹配技术及评价体系,并将其分为局部匹配和全局匹配两大类。局部方法通常容易实现、效率较高,但也存在兼顾鲁棒性和计算效率、利用局部特征和代价聚合消除匹配歧义等问题<sup>[2-11]</sup>。利用截断阈值<sup>[1]</sup>和特征描述<sup>[2]</sup>初始化匹配代价,适度改善匹配性能。DAISY描述显著特征较鲁棒<sup>[2]</sup>。文献[12]利用韦伯准则构造鲁棒光照局部描述。稳健高效的代价聚合有利于局部匹配。传统自适应权重<sup>[3]</sup>及带分割变种<sup>[4]</sup>,需要稠密计算或图像

2011-09-21 收到, 2012-01-06 改回

国家青年科学基金(61001152), 国家自然科学基金(61071166, 61071091, 61172118), 江苏省自然科学基金(BK2010523), 江苏省高校自然科学基金(11KJB510012), 南京邮电大学科研基金(NY210053/NY210069/NY210073/NY211030)和江苏高校优势学科建设工程资助课题

\*通信作者: 刘天亮 liutl@njupt.edu.cn

预分割，匹配效果虽然较好，但由于算法复杂度和内存容量的限制<sup>[13]</sup>，速度较慢，难处理大视差搜索范围、高分辨率图像对。文献[5]利用导引滤波<sup>[6]</sup>快速增强匹配代价；而文献[13]回顾了代价聚合中减少计算冗余方法。传统的双边滤波基于四邻域搜索去除视差噪声、保存边缘<sup>[7]</sup>；但该方法改善效果欠佳。兼顾实时性，可以牺牲精度<sup>[8]</sup>来实现高效匹配；也可借助硬件(图形加速硬件<sup>[9]</sup>或嵌入式系统<sup>[10]</sup>)实现。

设基准图像 $I_b$ 和待匹配图像 $I_m$ 是极线校正彩色图像对；视差范围 $R_d=[d_{\min}, d_{\max}]$ ，其总个数 $N_d$ 为 $d_{\max}-d_{\min}+1$ 。受上述研究启发，提出一种新颖的基于导引滤波分层代价聚合快速立体图像匹配方法。

## 2 初始匹配

### 2.1 基于韦伯感知初始化匹配代价

根据韦伯定律<sup>[12]</sup>，将邻域亮度变化 $v_{00}$ 与当前像素 $x$ 的亮度 $v_{01}$ 之间的比值，视为差分激励 $\xi(x)$ 。

$$\xi(x) = \arctan\left[\alpha \frac{v_{00}}{v_{01}}\right] = \arctan\left[\alpha \sum_{i=0}^7 \left(\frac{x_i - x}{x}\right)\right] \quad (1)$$

其 $\arctan(x)$ 防止输入较大(较小)输出响应增加(减少)过快， $\alpha$ 调整邻域间差分响应。 $x_i$ 为八邻域亮度。

首先，提取立体图像中各像素 $x_p$ 的各色彩通道分量 $c$ 的差分激励 $\xi_c(x_p)$ 。 $\xi(x_p) \in [-\pi/2, \pi/2]$ 。 $c \in \{R, G, B\}$ 。接着，连接RGB色彩三色通道差分激励 $\xi_c(x_p)$ ，形成整体局部特征描述向量 $\mathbf{WLD}(x_p)$ 。

$$\mathbf{WLD}(x_p) = (\xi_R(x_p), \xi_G(x_p), \xi_B(x_p)) \quad (2)$$

然后，图像 $I_b$ 中某一像素 $p_b$ (坐标为 $(u, v)$ )和待匹配图像 $I_m$ 中极线上关于视差 $d$ 的像素 $q_{m,d}$ (坐标为 $(u-d, v)$ )的相关程度，利用 $p_b, q_{m,d}$ 两个像素的 $\mathbf{WLD}$ 描述和RGB色彩之间绝对值差异关于各自阈值 $T_{\mathbf{WLD}}$ 和 $T_{\text{RGB}}$ 来线性表达。则初始匹配代价 $C(u, v, d)$ 为

$$\left. \begin{aligned} C(u, v, d) &= \beta C_{\mathbf{WLD}}(u, v, d) + (1 - \beta) C_{\text{RGB}}(u, v, d) \\ C_{\mathbf{WLD}}(u, v, d) &= \min \left\{ \sum_c |\xi_c(u, v) - \xi_c(u-d, v)|, T_{\mathbf{WLD}} \right\} \\ C_{\text{RGB}}(u, v, d) &= \min \left\{ \sum_c |I_c(u, v) - I_c(u-d, v)|, T_{\text{RGB}} \right\} \end{aligned} \right\} \quad (3)$$

合适选择 $T_{\mathbf{WLD}}$ 和 $T_{\text{RGB}}$ 约束差异测度可剔除宿外点。为了降低光照变化<sup>[12]</sup>影响，平衡因子 $\beta$ 取值接近1。

### 2.2 分层代价聚合和快速视差选择

**2.2.1 基于导引滤波提取视差候选** 为避免重复遍历所有视差假设<sup>[13]</sup>，从较大视差范围 $R_d$ 提取 $D_c(D_c \ll N_d)$ 个次优候选；形成小规模视差候选集 $S_d$ 并作为代价分级聚合和视差选择的对象实现加速匹配。

首先，初始匹配代价 $C(u, v, d)$ 中关于视差值 $d$ 的 $N_d$ 个切片，基于以当前像素为中心小尺度 $r \times r$ 匹配窗 $w_k$ 局部色彩信息导引滤波抑制噪声<sup>[6]</sup>。逐个

像素 $p_b=(u, v)$ 稳健增强匹配代价为

$$C_{gf}(u, v, d) = \sum_j W_{b,j}(p_b, p_j) C_j(u, v, d) \quad (4)$$

其中 $C_{gf}(u, v, d)$ 表示滤波后匹配代价， $p_b$ 和 $p_j$ 分别表示当前像素和匹配窗 $w_k$ 的邻域像素。权重化平均的滤波权重 $W_{b,j}(p_b, p_j)$ 取决于彩色导引图像 $I_b$ <sup>[6,9]</sup>为

$$W_{b,j}(p_b, p_j) = 1/|w_k|^2 \cdot \sum_{(b,j) \in \omega_k} \left(1 + (I_b - \mu_k)^T (\Sigma_k + \varepsilon U)^{-1} (I_j - \mu_k)\right) \quad (5)$$

其中 $I_b, I_j$ 和 $\mu_k$ 为图像 $I_b$ 中像素 $p_b$ 和 $p_j$ 及局部矩形窗 $w_k$ 内的色彩 $3 \times 1$ 均值向量；协方差矩阵 $\Sigma_k$ 和单位矩阵 $U$ 都是 $3 \times 3$ 矩阵。参数 $\varepsilon$ 控制权重平均的加权力度。权重滤波响应能保持彩色导引图像边缘。其次，逐像素遍历整个视差搜索空间 $R_d$ ，将导引滤波稳健增强后的匹配代价按大到小顺序依次排列。然后，选取匹配代价最小的 $D_c(D_c \ll N_d)$ 个视差，构成分层代价聚合候选视差子集 $S_d$ ，缩小搜索范围。

**2.2.2 空间离散采样联合式代价聚合** 为减少计算冗余，提出基于结构化空间离散周围对称聚合采样改进自适应权重匹配窗，作为匹配代价高效分级聚合的有效支撑。代价聚合结构是一组基于极坐标依赖当前参考像素周围对称空间离散采样点构成的结构化支撑模板(如图1所示)。分层联合式代价聚合为

$$C_{\text{aggr}}(p_b, d)|_{d \in S_d} = \sum_{l=0}^L w_{b,l}(p_b, p_l) C_{gf}(p_l, d)|_{d \in S_d} \quad (6)$$

$$w_{b,l}(p_b, p_l) = \exp\left(-\left(\Delta C_{p_b p_l} / \gamma_c + \Delta G_{p_b p_l} / \gamma_g\right)\right) \quad (7)$$

其中 $p_l(l=1, 2, \dots, L)$ 表示该支撑模板中第 $s$ 个同心圆环上第 $t$ 个代价累积采样点， $t=1, 2, \dots, 8, s=1, 2, 3$ 。代价分层聚合中当前层所对应的视差值 $d$ 来自上一层提取的视差候选集 $S_d$ 。权重 $w_{b,l}(p_b, p_l)$ 表示 $p_b$ 和 $p_l$ 关于RGB色彩差异 $\Delta C_{p_b p_l}$ 和几何距离 $\Delta G_{p_b p_l}$ 的自适应权重<sup>[3]</sup>， $\gamma_c$ 和 $\gamma_g$ 为相应权重系数。

联合式代价聚合中，忽略归一化求和自适应权重本身；仅保留聚合前匹配代价内积相应权重。仅保留内积操作的改进不会影响匹配代价聚合的精度<sup>[13]</sup>，且基于空间采样有利于匹配代价快速聚合。

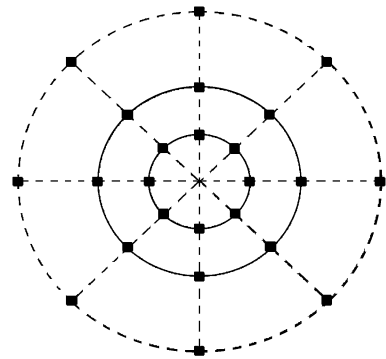


图1 空间离散采样结构化支撑模板

**2.2.3 利用候选子集快速选择视差** 采用优胜者全选法, 逐像素利用视差候选子集  $S_d$  加速两个视图选择初始视差。以基准图像  $I_b$  为例, 初始匹配视差选择:

$$D_{b,init}(p_b) = \arg \min_{d \in S_d} C_{aggr}(p_b, d) \quad (8)$$

### 3 视差求精

#### 3.1 改进型视差双边滤波

色彩不连续伴随深度不连续。基于四邻域双边滤波有助于视差不连续保持<sup>[7]</sup>; 但“\*”字型滤波窗口空间覆盖滤波范围更大。利用八方向邻域(见图2), 提出改进型双边滤波改善视差质量。设滤波半径为  $W_{bf}$ , 以当前像素  $p=(x, y)$  为中心滤波窗口的八邻域视差值集合为  $d_p$ 。设  $u_p, v_p, u'_p$  和  $v'_p$  是“\*”字型滤波窗口各方向路径: 水平  $u_p=[x-W_{bf}, \dots, x+W_{bf}]$ , 垂直  $v_p=[y-W_{bf}, \dots, y+W_{bf}]$ ,  $u'_p$  是  $45^\circ$  和  $225^\circ$  路径;  $v'_p$  是  $135^\circ$  和  $315^\circ$  路径。则顺序更新视差:

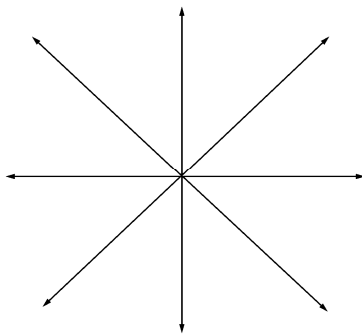


图2 “\*”字型滤波窗口

$$\left. \begin{aligned} D_{bf}(x, y) &= \arg \min_{d \in d_p} \{f_d\} \\ f_d &= \sum_{u \in u_p} \sum_{v \in v_p} \sum_{u' \in u'_p} \sum_{v' \in v'_p} W(u, v) C_{bf}(u, v, d) \\ C_{bf}(u, v, d) &= \min(\lambda_d N_d, |D_{init}(u, v) - d|^2) \end{aligned} \right\} \quad (9)$$

其中  $C_{bf}(u, v, d)$  反映邻域视差波动幅度, 常数  $\lambda_d$  控制去除宿外点;  $W(u, v)$  是基于色彩和空间相似类似式(7)的自适应权重<sup>[3]</sup>。权重化滤波也仅保留内积操作, 省去不必要的求和归一化操作<sup>[13]</sup>。改进双边滤波能保留视差图像的边缘细节, 并有效剔除视差局外点。

#### 3.2 利用对称映射后处理视差

首先, 利用含不可靠阈值  $T_{occ}$  的左右对称一致性约束检测不可靠歧义区域(比如误匹配及遮挡等):

$$|D_{b,bf}(x_l, y_l) - D_{m,bf}(x_l - d_{l2r}, y_l)| > T_{occ} \quad (10)$$

其中  $(x_l, y_l)$  为初始视差  $D_{b,bf}$  的坐标。不满足该约束, 则剔除该不可靠视差。接着, 将同一扫描极线上距离最近的可靠视差分配填充给当前的不可靠视差像素点。然后, 利用水平极线约束和左右2个视差图

的潜在冗余, 组合另一个视图基于逆向映射产生当前视图下的视差值, 消除窄遮挡的歧义。最后, 采用  $3 \times 3$  中值滤波平滑当前视图, 得到最终视差图。

### 4 实验结果与性能分析

本文方法保持相同参数设置测试 Middlebury 网站<sup>[14]</sup>上4组标准立体图像对 Tsukuba, Venus, Teddy 和 Cones 验证匹配性能。表1为与该网站上部分结果比较(2011年9月20日)并基于错误匹配率的评价表(误匹配阈值为1)。  $R_o$ ,  $R_D$  和  $R_A$  分别为非遮挡区域、接近深度不连续区域及未知像素除外总的错误百分比 PBP (Percent of Bad Pixels); 各列整体平均 AvgPBP。复杂度低、速度快的本方法, 稍微逊色于复杂度高的文献[3,4]等方法。整体匹配效果看, 本方法的 AvgPBP 为 8.11 (非常接近并仅次于文献[4]算法的 8.07)。本方法稍微好于文献[9]算法; 同时, 明显优于来自文献[8], 文献[11]和文献[7]等算法。

图3直观呈现本文方法匹配上述4组测试图像的相关结果。其中第1、第2行分别是各组立体对中的基准图像及其标准视差; 第3、第4行分别给出相应的初始视差和最终视差; 而第5行是最终视差与其标准视差之间的差异(误差阈值  $> 1$ )。可见, 初始匹配能得到较可靠的初始视差; 视差求精能有效改善歧义区域。图4给出本文方法处理 Middlebury 网站<sup>[14]</sup>提供的一组不同光照下真实芦荟立体图像的视差结果; 自左向右各列代表各组不同光照条件; 第1和第2行分别表示相应的基准  $I_b$  和待匹配  $I_m$ , 而第3行为相应的各基准图像视差结果。可见, 即使光照条件发生较大变化, 本文算法匹配性能几乎不受光照条件影响。

表2给出了未做代码优化和并行加速处理下本文方法与其他算法测试4组立体对的执行时间。与文献[3]方法相比, 本文方法运行时间降低了近一个数量级; 其整体计算复杂度不高, 适合匹配较高分辨率立体图像对。结合图3, 图4定性显示和表1, 表2定量评估的综合评价及其执行时间, 验证了本文方法在平衡精度和效率及真实光照变化上的优越性。

### 5 结束语

为了快速消除立体匹配歧义, 本文提出一种基于韦伯描述符和导引滤波由粗到细分层代价聚合的快速局部立体图像匹配方法。实验表明, 本文方法能获得高质量的稠密视差; 复杂度低, 匹配速度快; 光照鲁棒性好。面向高分辨率多视点/立体视频编码等新媒体应用, 计划研究更高质量实时立体匹配。

表1 Middlebury网站标准测试对的评估结果

算法	Tsukuba			Venus			Teddy			Cones			Avg. PBP
	$R_o$	$R_A$	$R_D$	$R_o$	$R_A$	$R_D$	$R_o$	$R_A$	$R_D$	$R_o$	$R_A$	$R_D$	
文献[3]	1.38	1.85	6.90	0.71	1.19	6.13	7.88	13.3	18.6	3.97	9.79	8.26	6.67
文献[4]	1.74	2.11	9.23	0.41	0.94	3.97	8.08	14.3	19.8	7.07	12.9	16.3	8.07
<b>本文算法</b>	<b>2.04</b>	<b>2.39</b>	<b>8.26</b>	<b>0.69</b>	<b>1.08</b>	<b>6.35</b>	<b>9.80</b>	<b>14.5</b>	<b>18.5</b>	<b>7.06</b>	<b>12.8</b>	<b>13.9</b>	<b>8.11</b>
文献[9]	1.92	2.45	9.66	1.03	1.65	6.89	8.48	14.2	18.7	6.56	12.7	14.4	8.21
文献[8]	2.25	3.08	11.6	0.92	1.31	7.53	10.7	15.8	23.6	8.25	13.5	16.6	9.59
文献[11]	3.62	5.52	14.6	3.15	4.20	20.4	11.5	18.2	23.2	4.93	13.0	11.7	11.2
文献[7]	2.00	4.17	10.5	1.48	3.11	17.7	11.1	20.2	27.5	5.98	16.5	16.0	11.4

表2 不同算法的执行时间(s)

图像对( $W \times H \times N_d$ )	Tsukuba(384×288×16)	Venus(434×383×20)	Teddy(450×375×60)	Cones(450×375×60)
本文算法( $D_c=6$ )	5.36	8.21	19.52	19.49
文献[4]	13.37	31.04	76.47	87.50
文献[3]	49.43	85.56	201.74	198.22

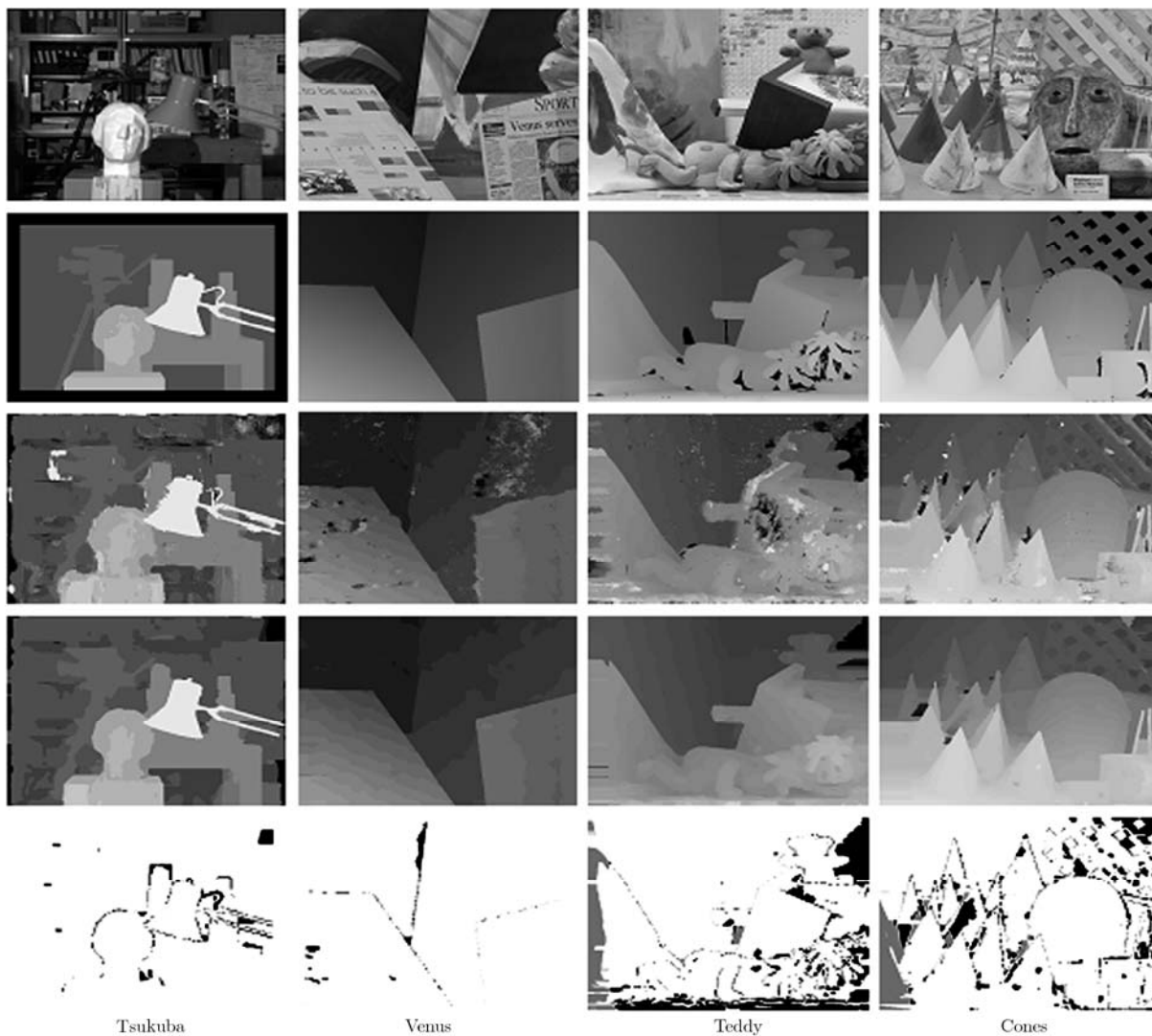


图3 立体图像对‘Tsukuba’, ‘Venus’, ‘Teddy’ 和‘Cones’的稠密视差结果

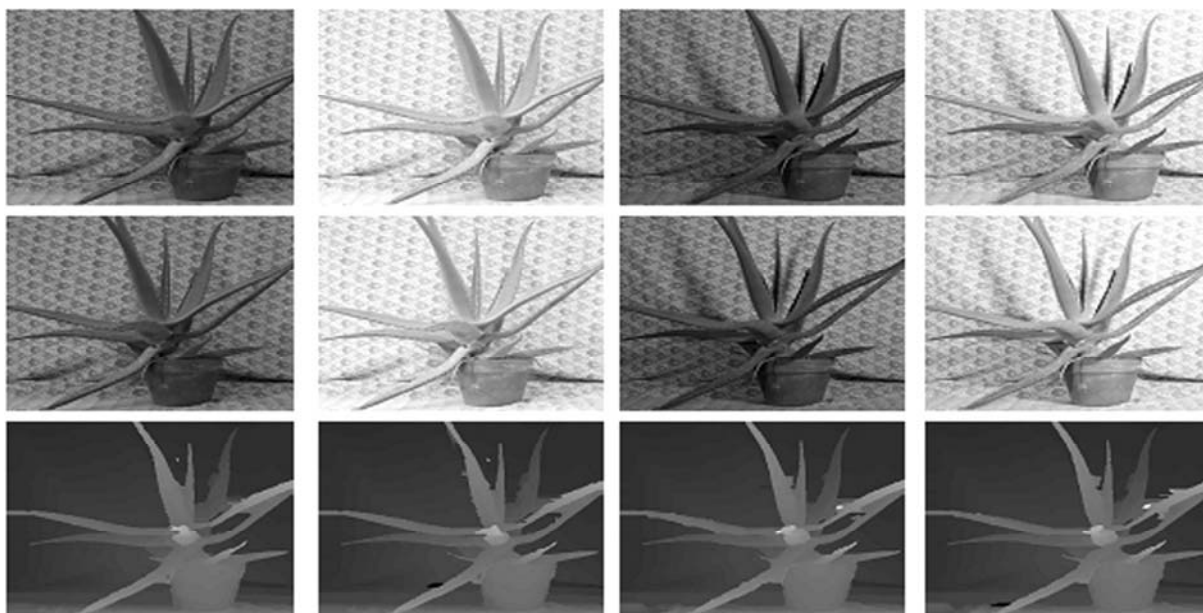


图4 不同光照条件下芦荟立体图像对的稠密视差

### 参考文献

- [1] Scharstein D and Szeliski R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms [J]. *International Journal of Computer Vision*, 2002, 47(1/2/3): 7-42.
  - [2] Tola E, Lepetit V, and Fua P. DAISY: an efficient dense descriptor applied to wide-baseline stereo[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32(5): 815-830.
  - [3] Yoon K J and Kweon I S. Adaptive support-weight approach for correspondence search[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, 28(4): 650-656.
  - [4] Liu T L, Zhang P Z, and Luo L M. Dense stereo correspondence with contrast context histogram, segmentation-based two-pass aggregation and occlusion handling[C]. Proceedings of the 3rd Pacific-Rim Symposium on Image and Video Technology, Tokyo, 2009: 449-461.
  - [5] Rhemann C, Hosni A, Bleyer M, *et al.* Fast cost-volume filtering for visual correspondence and beyond [C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Colorado Springs, USA, 2011: 3017-3024.
  - [6] He K M, Sun J, and Tang X O. Guided image filtering [C]. Proceedings of 11th European Conference on Computer Vision, Crete, Greece, 2010: 1-14.
  - [7] Yang Q X, Wang L, and Ahuja N. A constant-space belief propagation algorithm for stereo matching[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, San Francisco, USA, 2010: 1458-1465.
  - [8] Gupta R K and Cho S Y. A correlation-based approach for real-time stereo matching[C]. Proceedings of 6th International Symposium on Visual Computing, Las Vegas, USA, 2010: 129-138.
  - [9] Yu W, Chen T, Franchetti F, *et al.* High performance stereo vision designed for massively data parallel platforms [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2010, 20(11): 1509-1519.
  - [10] 丁菁汀, 杜歆, 周文晖, 等. 基于FPGA的立体视觉匹配的高性能实现[J]. *电子与信息学报*, 2011, 33(3): 597-603.  
Ding J T, Du X, Zhou W H, *et al.* High performance implementation of stereo vision matching based on FPGA[J]. *Journal of Electronics & Information Technology*, 2011, 33(3): 597-603.
  - [11] Nalpantidis L and Gasteratos A. Stereo vision for robotic applications in the presence of non-ideal lighting conditions [J]. *Image and Vision Computing*, 2010, 28(6): 940-951.
  - [12] Chen J, Shan S G, He C, *et al.* WLD: a robust local descriptor [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32(9): 1705-1720.
  - [13] Min D B, Lu J B, and Do M N. A revisit to cost aggregation in stereo matching: how far can we reduce its computational redundancy?[C]. Proceedings of 13th International Conference on Computer Vision, Barcelona, Spain, 2011: 1567-1574.
  - [14] Scharstein D and Szeliski R. Middlebury Stereo Evaluation — Version 2. <http://vision.middlebury.edu/stereo/eval>, 2011.
- 刘天亮: 男, 1980年生, 博士, 讲师, 研究方向为计算机视觉、图像处理与多媒体通信。
- 戴修斌: 男, 1980年生, 博士, 副研究员, 研究方向为医学图像重建、图像处理和模式识别。
- 朱秀昌: 男, 1947年生, 教授, 博士生导师, 研究方向为图像处理与多媒体通信。
- 罗立民: 男, 1956年生, 博士, 教授, 长江学者, 博士生导师, 研究方向为图像处理、计算机辅助医学诊断与治疗。