

文章编号 1004-924X(2016)04-0882-10

基于 CAD 模型的单目六自由度位姿测量

宋 薇*, 周 扬

(上海大学 机电工程与自动化学院, 上海 200072)

摘要:为实现单目摄像头对一般工业零件的六自由度位姿测量,研究了离线模板库自动建立,图像处理以及优化匹配 3 个方面的问题。首先,由 STL 文件格式的 CAD 文件建立不同观察方位下的匹配模板,定义了一种方便模板库调用并有图像意义的位姿描述参数。然后,对 Chamfer Match 方法进行改进,将距离映射图按边缘倾角分层,结合模板探索点的倾角进行匹配,在不增加运算量的情况下提升匹配度函数的灵敏性。最后,使用加入了退火算法的遗传算法进行最优搜索,从而提高匹配效率。实验结果表明:基于文中的方法完成一次位姿测量耗时在 500 ms 以内,搜索可以获到正确的结果。在 330 mm 观察距离上,平移误差和深度方向误差分别控制在 1 mm、2 mm 以内,转角误差为 1°左右。得到的结果基本满足单目视觉六自由度位姿快速匹配自动测量的要求。

关键词:单目视觉;位姿测量;CAD 模型;Chamfer 距离;优化匹配

中图分类号:TP391.4;TP242.6 **文献标识码:**A **doi:**10.3788/OPE.20162404.0882

Estimation of monocular vision 6-DOF pose based on CAD model

SONG Wei*, ZHOU Yang

(School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 200072, China)

* Corresponding author, E-mail: song_wei@shu.edu.cn

Abstract: To automatically and continuously estimate the 6-DOF pose of industry parts with a monocular camera, three related problems, building an off-line template library, image processing, and optimizing match, were researched. Firstly, the matching templates under different orientations were established by the CAD model in a STL(Standard Template Library)format, and a new pose parameter was defined to provide the call for a template library. Then, Chamfer Matching method was improved and the sensitivity of fitness function was improved without more time spending by layering the distance mapping map and matching the angle of template searching points. Finally, simulated annealing algorithm together with genetic algorithm was applied to the optimum searching to improve the matching efficiency in real time. The experiment result indicates that the whole match progress spends less than 500 ms, and the correct results are obtained. The translation error and the depth direct error have been kept by less than 1 mm and 2 mm respectively, and the rotate error is about 1° in a viewing distance of 330 mm. The result satisfies the requirements of the monocular vision system for rapid and automatic pose estimation.

Key words: monocular vision; pose estimation; CAD model; Chamfer distance; Optimal matching

收稿日期:2015-12-04;修订日期:2016-01-14.

基金项目:国家自然科学基金资助项目(No. 61105029)

1 引言

在工业自动化和智能设备中,目标物体的位姿是需要反馈给执行机构的一项重要信息。视觉测量以其快速性和免接触的特点,被广泛用于视觉测量领域。目前基于视觉的位姿测量问题主要分为以下两类:一类是根据特征点的投影变换关系求出两坐标系的变换关系,即求解 PnP 问题。这类方法通常使用角点特征^[1],或椭圆、直线这样的几何特征进行匹配^[2]。这种方法为正向求解过程,计算量小速度快,并且可以达到很高的精度。可用于传送带上的零件分拣和空间对接等领域。但该方法要求待测目标要有明显的特征点,对于形状复杂多变的工业零件而言,特征各不相同,有时还需要人为设定特征点^[3],因此该方法的通用性较差。另一类是通过搜索的方法求得最优匹配位姿。这种方法需要确定假定位姿下的探索模板与实际图像的匹配度函数,以匹配度最大时的假定位姿作为最终结果。该类方法是基于图像的边缘或灰度信息的,从而避免了寻找目标的几何特征,故通用性强。Changhyun Choi 等提出离线建立探索模板序列的方法^[4],其模板方位离散,平面外转角的精度依赖于离线模板的离散程度。在匹配函数的确定问题上,刘玉明等提出了以倾角匹配度作为 Chamfer Distance 权值的思想^[5],得到了很好的匹配结果,但计算图像上的边缘倾角时很耗时。

本文研究了一种实现工业零件的位姿快速自动测量方法,其具有较强通用性,并能用于自动抓取机械手的单目视觉伺服系统中以选择搜索匹配。以下从模板离线建立、匹配函数确定、搜索优化 3 个方面展开讨论。

2 离线探索模板的建立

2.1 虚拟观察点的分布和三维模板序列

本文所述的模板是基于边缘的,需与实际图像上的边缘点进行匹配。对于同一个模型,不同的观察方位会导致可见的边缘形状发生变化,因此需要讨论 CAD 模型的哪些边线在映射后会成为图像上的边缘的问题。由于物体的可见边缘仅与观察方位有关,所以可以在空间中设立一系列

不同方位的虚拟观察点来预先建立探索模板库,在后续的匹配中,可以直接调用,从而可以节省时间。

虚拟观察点应分布于以模型为中心的球面上,它们到物体中心的距离,可根据实际观察距离取一个合适的值 d_{est} 。在球面上设置一个或多个中心观察点,并以中心观察点为中心在一定范围内做经纬网格划分,网格的每一个交点代表一个虚拟观察点的位置。如图 1 所示。

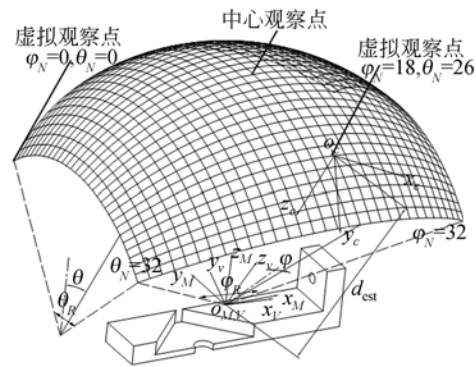


图 1 虚拟观察点的布置

Fig. 1 Arrangement of virtual observation points

定义虚拟观察点经度、纬度方向的实际角度分别为 φ, θ ; 网格数分别为 Ψ_N, θ_N , 角度范围分别为 φ_R, θ_R 。图 1 中, $\Psi_N = \theta_N = 32^\circ$, $\varphi_R = \theta_R = 90^\circ$ 。对于每一个虚拟观察点,由 CAD 模型投影得到可见边缘线,并经过采样得到探索模板点集。

本文使用的 CAD 文件格式为 STL 格式, STL 文件格式以 3 个点坐标表示一个三角网格,并附带该三角网格的法向量信息,用一系列封闭的三角网格来描述三维模型的形状^[6],如图 2 所示。

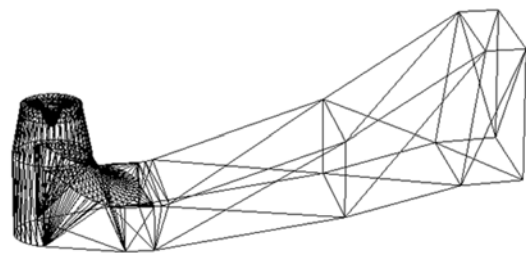


图 2 STL 文件中的三角网格

Fig. 2 Triangular mesh in STL file

从三角网格划分到三维探索模板的建立,可按以下几个步骤进行:

(1) 将 STL 文件中的三角网格转为线段表示。

(2) 提取出可能呈现为图像边缘的线段,此类线段分为两类,棱边线段和切边线段:棱边线段为两个呈一定夹角的面的共边线段;切边线段为在某观察点下,呈现于模型最边缘处的线段。

(3) 投影面上等间距提取步骤(2)所得线段的采样点,注意短线段需合并后再进行等间距采样。

(4) 参考 OpenGL 深度测试原理,消隐掉被遮挡的采样点。



图 3 边缘探索模板点集

Fig. 3 Template point set for edge exploration

经过上述 4 个步骤,得到图 2 所示的点集就是模型在某观察点下的可见边缘点点集,在模型坐标系下表示为 $\{(M\mathbf{P})_n\}$ 。同时还需要记录每个点所在线段的方向信息,记为 $\{(M\mathbf{v})_n\}$ 。 $\{(M\mathbf{p}, M\mathbf{v})_n\}$ 就是某虚拟观察点下的三维探索模板。对每一个虚拟观察点建立三维探索模板,就得到了 $\Psi_N \times \theta_N$ 个三维探索模板,称为探索模板序列,记为 $\{(M\mathbf{p}, M\mathbf{v})_n\}_p, p \in [0, \Psi_N \times \theta_N]$ 。

定义观察点的经度纬度方向序号分别为 φ_N, θ_N , 在匹配搜索时,就可以直接根据 φ_N, θ_N 从离线建立的模板库中提取数据 $\{(M\mathbf{p}, M\mathbf{v})_n\}_p$ (p 由 φ_N, θ_N 决定),从而节省了匹配模板建立时间。

为使方位值连续,引入参数 φ_n, θ_n , 它们的整数部分分别等于 φ_N, θ_N , 而细化小数部分,从而解决了用 φ_N, θ_N 表述时产生的方位离散问题:

$$\varphi_n = [\varphi_n], \varphi_n \in (0, \Psi_N), \quad (1)$$

$$\theta_n = [\theta_n], \theta_n \in (0, \theta_N). \quad (2)$$

这样一来,离线模板序列虽然是离散的,但变量 φ_n, θ_n 是连续的。在投影时依据 φ_n, θ_n 取最近方位上的三维模板进行投影变换。由于三维模板

上的采样点始终是原模型上的点,因此一般情况下,离线模板离散建立不会影响精度。

φ_n, θ_n 表达了相机相对于物体的方位,决定了探索模板即可见边缘的形状。

2.2 位姿表示法和三维模板投影变换

2.1 节中定义了表达观察方位的变量 φ_n, θ_n 。为确定相机坐标系(C)和模型坐标系(M)的变换关系,本文进一步引入以下 4 个参数:

x_l 为模型中心在虚拟相机图像上投影的水平位置(以比例表示,最左为 -1,最右为 1, y_l 同)。

y_l 为模型中心在虚拟相机图像上投影的垂直位置。

d_l 为模型坐标系中心点在图像坐标系中的深度与估计观察距离 d_{est} 之比。

ω 为模型坐标系 X 轴在图像上的投影与图像水平方向的夹角。

以上 6 个参数 ($\varphi_n, \theta_n, x_l, y_l, d_l, \omega$) 可以唯一确定一个物体相对于相机的位姿,以这种方式表达位姿,其优点在于:可直接根据 φ_n, θ_n 从预建立的模板序列中调取模板,节省了三维模型处理的时间。另外,这 6 个参数在图像上均有独立意义,耦合性小。改变其中一个参数,不会影响探索模板的其他意义,从而不会导致其他参数的峰值有较大改变,使搜索的归一性好。现在由以上 6 个参数推导模型坐标系(M)中的探索点,生成相机成像平面上的探索点。

观察点的实际经纬度 φ, θ 可由式(3)(4)求得:

$$\varphi = \frac{\varphi_n \times \varphi_R - \varphi_R}{\Psi_N - 2}, \quad (3)$$

$$\theta = \frac{\theta_n \times \theta_R - \theta_R}{\theta_N - 2}, \quad (4)$$

因此虚拟观察点在 M 坐标系中的方位向量为:

$${}^M\mathbf{V} = (\sin \varphi \cdot \mathbf{x}_M, \cos \varphi \sin \theta \cdot \mathbf{y}_M, \cos \varphi \cos \theta \cdot \mathbf{z}_M)^T, \quad (5)$$

定义一个观察坐标系(V)作辅助过渡,见图 1,其原点与模型坐标系重合, Z 轴指向虚拟观察点, X 轴为模型坐标系 X 轴 (\mathbf{x}_M) 沿 V 坐标系 Z 轴的投影,即:

$${}^M\mathbf{z}_V = {}^M\mathbf{V}, \quad (6)$$

$${}^M\mathbf{x}_V = (\mathbf{x}_M / {}^M\mathbf{V})_1, \quad (7)$$

Y 轴与 Z 轴、X 轴正交。得出 M 坐标系到 V 坐标系的转换矩阵为:

$${}^V\mathbf{T}_M = \begin{bmatrix} [(\mathbf{x}_M/{}^M\mathbf{V})_1 \quad {}^M\mathbf{V} \times (\mathbf{x}_M/{}^M\mathbf{V})_1] \quad {}^M\mathbf{V}^{-1} & 0 \\ \mathbf{0}^T & 1 \end{bmatrix}, \quad (8)$$

其中,定义运算 $v_1/v_2 = v_1 - \frac{v_1 \cdot v_2}{|v_2|^2} v_2$, $(v)_1 = \frac{1}{|v|} v$,下同。

在相机小孔成像模型中,以 R_X, R_Y 分别作为 X, Y 方向成像尺寸的一半与焦距之比(R_X, R_Y 与相机内参数有关),则由 x_l, y_l, d_l 可得模型中心即 V 坐标系中心在相机坐标系(C)中的坐标,将其记为向量 ${}^c\mathbf{t}$:

$${}^c\mathbf{t} = (x_l d_l d_{est} R_X, y_l d_l d_{est} R_Y, d_l d_{est})^T, \quad (9)$$

由于 V 坐标系 Z 轴指向相机坐标系中心,所以:

$${}^c\mathbf{z}_v = (-{}^c\mathbf{t})_1, \quad (10)$$

又由 ω 的定义可知, V 坐标系的 X 轴沿 Z 轴的投影与 C 坐标系的 X 轴夹角为 ω ,因此可得:

$${}^c\mathbf{x}_v = ((\cos \omega, \sin \omega, 0)^T / {}^c\mathbf{t})_1, \quad (11)$$

Y 轴与 Z 轴、X 轴正交。得 V 坐标系到 C 坐标系的转换矩阵为:

$${}^C\mathbf{T}_V = \begin{bmatrix} \begin{bmatrix} ((\cos \omega, \sin \omega, 0) / {}^c\mathbf{t}^T)_1 \\ (-{}^c\mathbf{t}^T)_1 \times ((\cos \omega, \sin \omega, 0) / {}^c\mathbf{t}^T)_1 \\ (-{}^c\mathbf{t}^T)_1 \end{bmatrix} & {}^c\mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}, \quad (12)$$

2.1 小节中得到的探索模板点集和方向为 $\{({}^M\mathbf{p}, {}^M\mathbf{v})_n\}$,则可得到探索点在相机坐标系中的坐标和方向 $\{({}^c\mathbf{p}, {}^c\mathbf{v})_n\}$:

$${}^c\mathbf{p}_i = {}^c\mathbf{T}_V {}^V\mathbf{T}_M {}^M\mathbf{p}_i, \quad (13)$$

$${}^c\mathbf{v}_i = {}^c\mathbf{T}_V {}^V\mathbf{T}_M {}^M\mathbf{v}_i, \quad (14)$$

相机标定后,就可以使用标定相机的参数对相机坐标系中的点和方向 $\{({}^c\mathbf{p}, {}^c\mathbf{v})_n\}$ 进行投影,得到图像平面上的探索点和点所在线段的方向 $\{({}^l\mathbf{p}, {}^l\mathbf{v})_n\}$ 。由于镜头的成像畸变对测量结果的影响较大,故在标定时应计算镜头的畸变参数,另外,在投影变换时,加入畸变校正以使相机逼近针孔成像模型。

2.3 转换为常规位姿表示

物体的位姿通常以常规位姿表示,即沿三轴

平移 (t_x, t_y, t_z) 和沿三轴旋转 (α, β, γ) 6 个参数。本文自定义的位姿描述参数 $(\varphi_n, \theta_n, x_l, y_l, d_l, \omega)$ 用于搜索和匹配,当得到最优匹配后,需要转化为 $(t_x, t_y, t_z, \alpha, \beta, \gamma)$ 作为最终结果,方法如下:

由式(3)-(8),利用 φ_n, θ_n 求得 ${}^V\mathbf{T}_M$,由式(9)-(12)可知,由 x_l, y_l, d, ω 可求得 ${}^c\mathbf{T}_V$,它们相乘可得模型坐标系到相机坐标系的转换矩阵 ${}^c\mathbf{T}_M$ 的值:

$${}^c\mathbf{T}_M = {}^c\mathbf{T}_V {}^V\mathbf{T}_M, \quad (15)$$

对应的数值矩阵和表达式矩阵 t_x, t_y, t_z 分别等于 $({}^c\mathbf{T}_M)_{1,4}, ({}^c\mathbf{T}_M)_{2,4}, ({}^c\mathbf{T}_M)_{3,4}$,而 α, β, γ 也可由下式得到:

$$\alpha = \arctan2(-({}^c\mathbf{T}_M)_{2,3}, ({}^c\mathbf{T}_M)_{3,3}), \quad (16)$$

$$\beta = \arcsin(({}^c\mathbf{T}_M)_{1,3}), \quad (17)$$

$$\gamma = \arctan2(-({}^c\mathbf{T}_M)_{1,2}, ({}^c\mathbf{T}_M)_{1,1}), \quad (18)$$

从而利用求得的 $(\alpha, \beta, \gamma, t_x, t_y, t_z)$ 就可以表示 CAD 模型坐标系(M)相对于相机坐标系(C)的常规位姿。

3 模板与图像的匹配

3.1 匹配度函数

匹配度函数是衡量探索模板与边缘图像的匹配程度的函数。本文以探索点的平均 Chamfer Distance 作为模板和图像的匹配度,这在一些文献中也称为平均 Hausdorff 距离^[9]。Chamfer Distance 提供了一种快速求取图像上一点到最近边缘点距离的方法。该方法通过正向和反向两次平移 Chamfer 模板,将原本二值化的边缘图变为 Chamfer 距离映射图,图上任一点的灰度值就代表其最近的边缘点距离^[7]。它以所有探索点到图像上最近边缘点的平均距离来衡量匹配的优劣。

在杂乱的环境下,普通的 Chamfer 距离映射图中,边缘线较多,模板上的探索点也很容易落到不正确的边缘线旁,使得平均 Chamfer 距离偏小,匹配度误差增大。有文献提出在匹配边缘点处引入角度差异作为权值^[5],也有学者提出按图像倾角分层方法^[8]。

本文采用倾角分层匹配的方法。其主要依据为:当一个模板点正确匹配时,其理论边缘倾角应与图像上的边缘倾角相符。根据这一特点,模板

点在 Chamfer 距离映射图上取灰度时,模板点的倾角必须与图像上边缘点的倾角在一定范围内,以此降低错误的匹配,突出正确匹配时的峰值。

本文将原图边缘点按其所在直线的倾角分入 4 幅图中,它们的倾角范围分别为 $-22.5^\circ \sim 22.5^\circ, 22.5^\circ \sim 67.5^\circ, 67.5^\circ \sim 112.5^\circ, 112.5^\circ \sim 157.5^\circ$,即边缘线倾角分别在 $0^\circ, 45^\circ, 90^\circ, 135^\circ$ 附近。分别对它们做 Chamfer 距离映射图,如图 4。

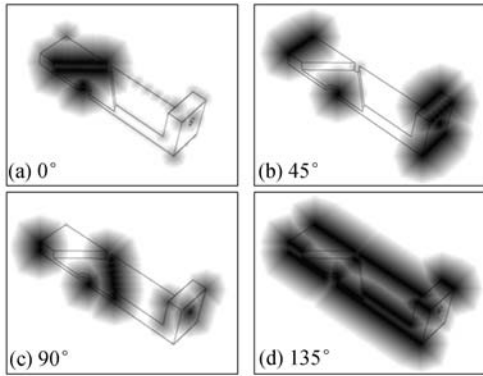


图 4 边缘倾角在 4 个角度附近的 Chamfer 映射图
Fig. 4 Chamfer distance mapping when edge inclination angles in the vicinity of 4 certain angles

定义 $D_n(x, y)$, n 为 $0 \sim 3$ 表示 4 幅图的灰度分布。然后计算匹配模板上的点(探索点)所在直线的倾角 μ_i 。 μ_i 可由 2.3 小节中得到的边缘线方向 ${}^l v_i$ 求得,即:

$$\mu_i = \text{atan2}({}^l v_i(y), {}^l v_i(x)), \quad (19)$$

根据每个探索点的 μ_i ,在相应的距离映射图上取灰度值,用公式表述为: $D_{n_i}({}^l P_i)$,其中 $n_i = \text{mod}\left(\left\lceil \frac{\mu_i + 22.5^\circ}{45^\circ} \right\rceil, 4\right)$ 。所以匹配度函数表示为:

$$F(\varphi_n, \theta_n, x_i, y_i, d_i, \omega) = 255 - \sum_i^N D_{n_i}({}^l P_i) / N, \quad (20)$$

引入最高灰度 255 是为了符合数值越高匹配越高的习惯。该匹配量度下,最高匹配度为 255。

下面对比分析普通 Chamfer 距离映射和加入倾角分层的 Chamfer 映射的匹配度函数差异。选取一张单目标图像,先匹配到其最优解,然后分别只变动 $(x_i, y_i), (d_i, \omega)$ 检查匹配度函数的分布。如图 5、6 所示。

图 5、图 6 中,左边为基于普通 Chamfer 距离的匹配度函数分布图,右边为改进后的。可以看

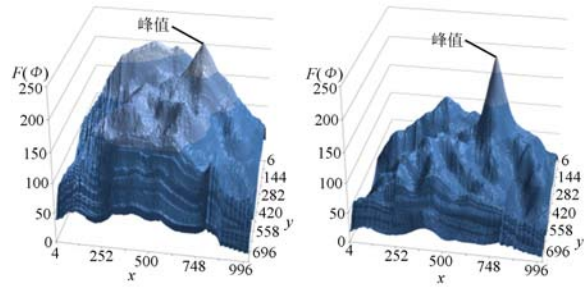


图 5 变量参数为 x_i, y_i 的匹配度函数分布图
Fig. 5 Matching function distribution when x_i, y_i are variables

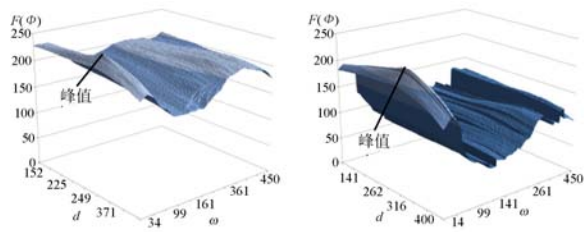


图 6 变量参数为 d_i, ω 的匹配度函数分布图
Fig. 6 Matching function distribution when d_i, ω are variables

出,增加倾角匹配信息后,只有在正确的目标附近,匹配度值才更高,目标峰值体现得才更加明显。

3.2 图像分层处理方法

3.1 节中提到,将原图边缘线按倾角分为 4 层,但由边缘图像求得每个边缘点的倾角十分耗时,因此,本文通过分析开源图像处理库 OpenCV 中的 Canny 源码,提出一种在提取 Canny 边缘的同时做分层处理的方案。

Canny 图像上的点有两种产生方式,其一是在低阈值下进行非极大值抑制得到的点,其二是高阈值图中可以与低阈值图中的点相连的点。第一类点提供了该点的灰度梯度方向,第二类点提供了与该点相连的边缘点的方向。OpenCV 中的 Canny 函数根据这两种方式,在一个已有的边缘点周围的 8 个像素中生成一个新的边缘点,新边缘点生成时,可以知道该点相对于原来边缘点的方向,根据这个生成方向,将新边缘点放入相应的倾角分层图中。这样一来,在整图 Canny 边缘提取完成时,边缘按倾角分层的工作也已经完成,减少了多余的时间开销。

但是,在实际拍摄的图像中存在干扰边缘,利

用 Canny 边缘连接时也不能保证连接方向的一致性。若直接将边缘点按倾角分散在 4 个区间中,则会出现错误的孤立点。在二值化的边缘图像中,即便是一个孤立的边缘点也会造成大范围的 Chamfer 距离(如图 7),对后面的匹配造成不良影响。

为解决此问题,本文提出以灰度代替二值化边缘图来衡量边缘点的孤立程度。具体方法为:沿着区间倾角方向统计某点附近的边缘点数(图 8)(彩图见期刊电子版),若能连续则设定该点灰度值为 255,否则按比例减小灰度值。淡化孤立点并进行 Chamfer 距离映射的结果见图 9。

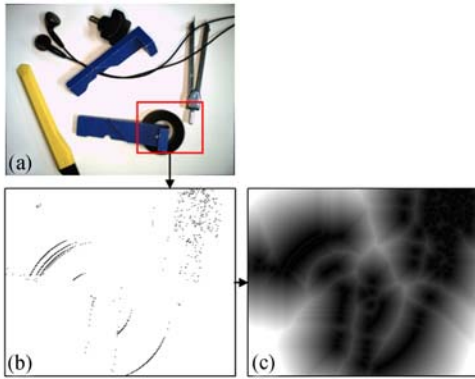


图 7 普通 Chamfer 距离映射结果((a)原图像;(b)边缘图像;(c) Chamfer 距离映射)

Fig. 7 Mapping results of Chamfer distance((a) Original image; (b) Edge detected image; (c) Chamfer Distance image)

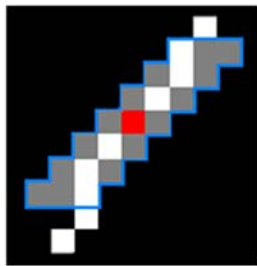
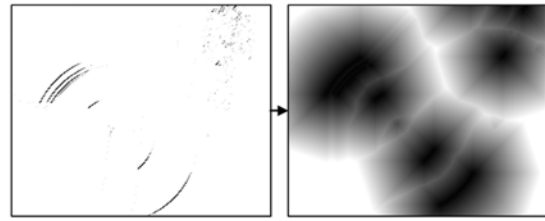


图 8 45°方向上 7×3 区域内边缘点个数

Fig. 8 Number of edge points in a 7×3 region along 45° direction

可以看出,随着孤立点灰度值的淡化,其产生的 Chamfer 映射范围也减小了,这样一来,探索模板上的点就更倾向于与原图中角度连续的直线或弧线匹配,从而较少地受到错误孤立点的影响。



(a)淡化边缘点后的边缘图像 (b)淡化孤立点后的距离映射图
(a) Edge detected image after isolation fading (b) Chamfer distance image after isolation fading

图 9 改进的 Chamfer 距离映射结果

Fig. 9 Mapping results of improved Chamfer distance

4 匹配搜索方法

4.1 遗传算法全局寻优

遗传算法(Genetic Algorithm,GA)是一种自适应全局优化概率搜索算法,它通过模拟自然进化过程来搜索问题的最优解。GA 求解时,需要确定一个匹配度函数,并将待解参数编码为二进制或格雷码,模拟为基因个体。匹配度函数用来评价每个基因个体的正确程度,GA 选择优良的基因个体进行交叉繁殖,而淘汰劣质基因,从而使种群趋于最优解^[10]。遗传算法的步骤如下:

- (1)确定基因编码格式;
- (2)随机产生初始种群 $\{\Phi_i^1\}$;
- (3)计算每个个体的匹配度 $F(\Phi)$;

(4)按匹配度从大到小的排序,按一定比例在本代个体中选择前面的个体进行交叉变异操作,从而产生新的个体,淘汰未被选择的个体;

(5)循环第 4 步直至匹配度达到一定阈值时,输出结果。

本文采用二进制法编码基因。待解参数为 $\Phi = (\varphi_n, \theta_n, x_l, y_l, d_l, \omega)$ 。 x_l, y_l 的位长依据图像的像素尺寸取 8~10 位, $d_l, \omega, \varphi_n, \theta_n$ 则依据结果精度要求取适度位长。

GA 是一种智能搜索算法,其能在大范围内快速找到优秀的群体,其搜索间距能随着群体聚拢而自动缩小。在匹配度达到一定程度之前,可使用 GA 找出较优秀的群体,但由于其具有随机性,难以找到极值^[11]。在种群收敛到一定水平后,需要一种有向的极值搜索方法。

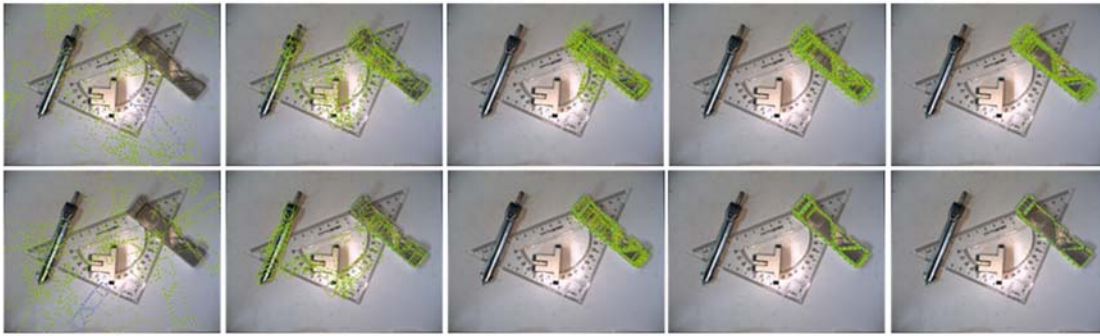


图 10 单遗传算法(上)和加入退火算法后的遗传算法(下)的搜索收敛结果比较

Fig. 10 Comparison of searching convergence between Genetic Algorithm(upper) and GA+SA(lower)

4.2 模拟退火算法

为达到局部最优解,最简单的做法是在最优值附近作线性搜索^[12](即爬山法)。但匹配度函数并不是完全光滑的,在最优值附近仍然会存在局部小峰值,但简单的线性搜索很容易使结果陷入局部极值。模拟退火算法(Simulated Annealing, SA)作为线性搜索的一个改进方法,是以一定概率接受比当前匹配度值更低的解,以跳出局部最小。SA 的具体步骤如下:

对于 Φ_x 附近的节点 Φ_{x+1} ,若 $F(\Phi_{x+1}) \geq F(\Phi_x)$,则替换 Φ_x 为 Φ_{x+1} ,否则,以概率 $P = \exp(-dE/(kT))$ 接受 Φ_x 替换为 Φ_{x+1} 。其中, $dE = F(\Phi_x) - F(\Phi_{x+1})$, k 为大于 0 的常数, T 为退火时间,这里可直接取为运算次数。

本文有 6 个相互独立的参数待优化,且各参数的变动范围也不同,因此各参数退火算法的节点间距各自独立,在参数值随退火算法改变了 2 次或 3 次时,将步长变为原来的 1/2,直至达到最小步长,以保证基因个体的最小精度。终止条件定为:某次退火算法结束后,没有任何一个参数改变。

模拟退火算法被安排在遗传算法个体排序后进行,对于每一个待繁殖的个体,在其产生后代前,先利用模拟退火算法求出其附近的局部最优值。为了得到全局较优解,可设置一个匹配度阈值 F_H ,仅当遗传算法最优个体的匹配度大于 F_H 时,才加入模拟退火算法。整个过程如图 11 所示。

分别使用单遗传算法和加入退火算法后的遗传算法匹配目标,大致间隔 100 世代记录一次前

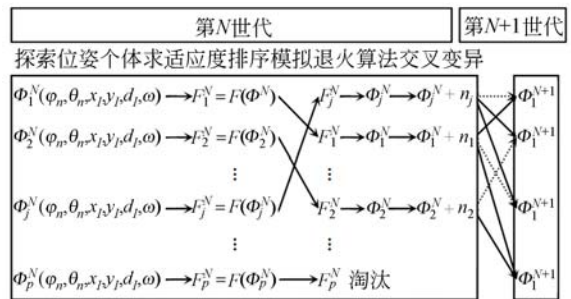


图 11 加入模拟退火算法的遗传算法

Fig. 11 Genetic Algorithm by adding Simulated Annealing

20 个个体的位姿情况,结果如图 10 所示。可以看出,在搜索初期,是否加入退火算法并无太大区别,依靠 GA 就可以使种群趋于目标。但接近目标后,GA 难以使群体归于最优解,匹配度值很难提升,但加入退火算法后,种群可以很快趋于单一,最终得到最优解。

5 实验

5.1 加入退火算法对收敛速度的提升实验

为衡量模拟退火算法对匹配速度的优化程度,本文选取图 10 作为实验对象,使用 GA 和 GA+SA(取 $F_H=200$)两种方法做搜索匹配。记录适应度值随世代数增长的变化情况,并做 5 次实验,取适应度变化的平均值,结果如图 12 所示。加入退火算法后,适应度超过 F_H 的个体会被做一次退火运算,因此,与单遗传算法相比每一世代的耗时有所增加,为对比两种方法的耗时,分别记录两种方法下,适应度值首次超过 230(GA+

SA)和 220(单 GA,因 300 世代下 GA 可能无法达到 230 的匹配度)时的耗时。同样取 5 次实验的平均值,结果如表 1。

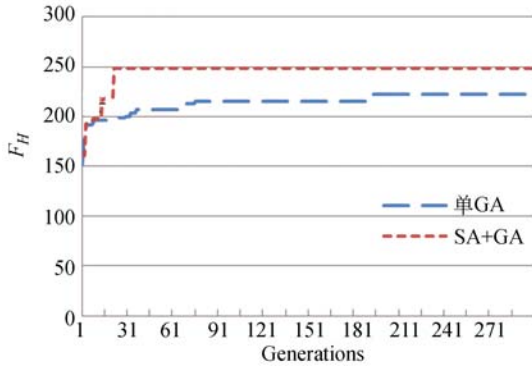


图 12 两种方法的匹配度提升情况对比

Fig. 12 Matching improvement contrast of two methods

表 1 两种方法耗时对比

Tab. 1 Contrast of consuming time of two method (ms)

	GA(>220)	GA+SA(>230)
耗时	1 141.693 6	317.142 3

可以看出,加入退火优化后的遗传算法,适应度函数上升迅速,且能达到较高的匹配度。实际上,虽然最终 GA 的匹配度大于 200,但匹配结果有较大误差,相比 GA+SA 而言,鲁棒性不强。从耗时上也可以看出,虽然对 GA 降低了要求,但其耗时仍然比 GA+SA 长一些。一次图像处理的耗时在 50 ms 左右,因此使用 GA+SA 估计一次位姿匹配的耗时在 500 ms 以内。综上可知,退火算法对遗传算法的优化作用明显。

5.2 位姿测量结果验证

为了验证位姿测量结果的正确性,本小节安排了 4 组实验:将物体以确定的位姿(部分或全部自由度已知)摆放,然后用本文方法测量其位姿,并与已知位姿作对比。

4 组实验包括:(1)平面内移动实验一;(2)平面内移动实验二;(3)深度方向移动实验;(4)平面外旋转实验。

物体的位姿都在世界坐标系(W)下描述,相机放置在世界坐标系原点上方 330 mm 处。图 13 给出了全 0 位姿下物体的状态。

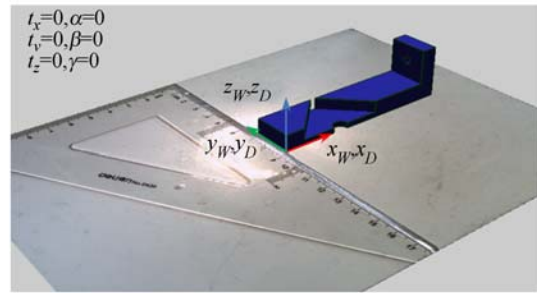


图 13 世界坐标系(W)的定义

Fig. 13 Definition of world coordination

图 14~17 是 4 组实验的匹配结果,目标物体的放置位姿放在图的左侧,匹配得到的位姿标于图上。实验(3)、(4)同时给出了普通相机拍摄的实验场景。

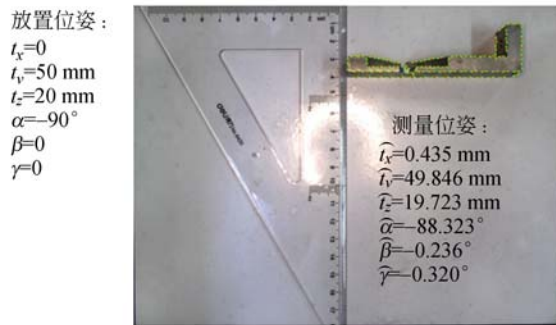


图 14 平面内移动之实验一

Fig. 14 In-plane moving experiment one

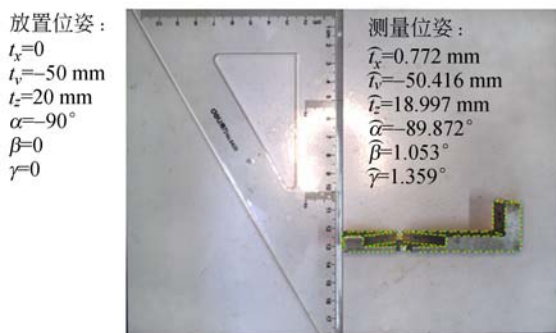


图 15 平面内移动之实验二

Fig. 15 In-plane moving experiment two

计算 4 组实验得到的测量值与理论值之差,并统计各自由度误差的最大值,记录于表 2 中:

分析本测量装置的系统误差:在 $1\ 024 \times 768$ 的图像分辨率下,镜头内参数 $f_x=1\ 407.423$, $f_y=1\ 407.956$ 。在理想的针孔模型下,将一个像素

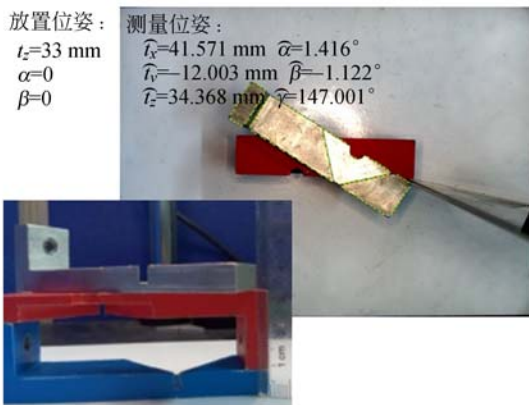


图 16 深度方向移动实验

Fig. 16 Experiment of moving along depth direction

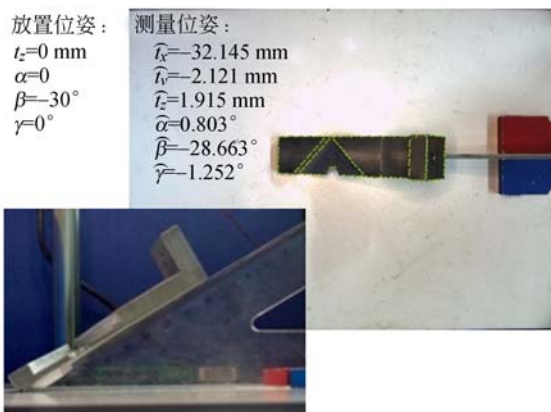


图 17 平面外旋转实验

Fig. 17 Experiment of rotating out-of-plane

可以区分的位姿变动作为系统误差,可得到平移误差为 $E(t_x) = E(t_y) = 0.24 \text{ mm}$; 对于长度在 50 mm 左右的工件, z 轴方向的误差为 $E(t_z) = 1.78 \text{ mm}$; 绕三轴的转角误差 $E(\alpha) = E(\beta) = E(\gamma) = 0.54^\circ$ 。对照表 2 可知,除转角误差较大外,实验误差基本接近系统误差。造成转角误差较大的因素有零件本身尺寸误差、零件棱边磨损、镜头畸变不能完全消除等。

参考文献:

- [1] 黄风山, 刘恩福, 方忆湘, 等. 基于智能三坐标测量机的零件位姿单目立体视觉识别[J]. 光学精密工程, 2013, 21(5):1326-1332.
HUANG F SH, LIU E F, FANG Y X, *et al.*. Single camera stereo vision recognition for parts'

表 2 4 组实验测量误差统计

Tab. 2 Error statistics of 4 groups of experiments

	实验 1.	实验 2.	实验 3.	实验 4.	最大值
	平面移动一	平面移动二	深度移动	面外旋转	
$\Delta t_x/\text{mm}$	0.435	0.772	-	-	0.772
$\Delta t_y/\text{mm}$	0.154	0.416	-	-	0.416
$\Delta t_z/\text{mm}$	0.277	1.003	1.368	1.915	1.915
$\Delta \alpha/(\text{degrees})$	1.687	0.128	1.416	0.803	1.687
$\Delta \beta/(\text{degrees})$	0.236	1.053	1.122	1.337	1.337
$\Delta \gamma/(\text{degrees})$	0.310	1.359	-	1.252	1.359

6 结 论

针对基于 CAD 模型的目标匹配和位姿测量问题,本文提出了一套方法,包括 CAD 文件的分析,探索模板的离线建立,适应度函数的定义和优化搜索方法。在探索模板建立中,提出一种新的模型位姿表示方法,该方法以其中的两个参数作为变量离线建立探索模型,减少了实时匹配中的模型生成运算,提高了匹配速度。该表示方法中各参数在图像上均有独立意义,相互间的耦合性小,有利于各参数的独立收敛。定义适应度函数时,加入了边缘倾角匹配,使其适应度函数值能更好地反映匹配程度,一定程度上提高了遗传算法的搜索效率。匹配搜索时,引入模拟退火算法,使其既能快速地从全局范围缩小到局部较优范围,也能在局部范围内准确找到最优解。整个算法耗时短,匹配效率高。

实验结果表明:完成一次位姿测量的耗时在 500 ms 以内,搜索可以得出正确的结果。在 330 mm 的观察距离上,平移误差控制在 1 mm 以内,深度方向误差均在 2 mm 以内,转角误差在 2° 以内,接近于理论误差,能够满足一般机械手抓取任务的精度要求。

pose based on intelligent three coordinate measuring machine[J]. *Opt. Precision Eng.*, 2013, 21(5): 1326-1332. (in Chinese)

- [2] 刘凌云, 钱新恩, 罗敏. 确定对象单目视觉位姿检测算法的研究[J]. 中国机械工程, 2010, 21(23): 2825-2829.

LIU L Y, QIAN X N, LUO M. Study on pose de-

- tection algorithm for specific object based on monocular vision [J]. *China Mechanical Engineering*, 2010, 21(23):2825-2829. (in Chinese)
- [3] 耿明超, 赵铁石, 边辉, 等. 基于并联理论的单目视觉位姿估计[J]. *光学精密工程*, 2013, 21(10):2617-2626.
GENG M CH, ZHAO T SH, BIAN H, *et al.*. Pose estimation of monocular vision based on parallel theory [J]. *Opt. Precision Eng.*, 2013, 21(10):2617-2626. (in Chinese)
- [4] CHOI CH, CHRISTENSEN H I. 3D textureless object detection and tracking: An edge-based approach[C]. 2012 *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, IEEE, 2012:3877-3884.
- [5] LIU M Y, TUZEL O, VEERARAGHAVAN A, *et al.*. Fast directional Chamfer Matching[C]. *IEEE Conference on Computer Vision & Pattern Recognition (CVPR)*, IEEE Computer Society, 2010:1696-1703.
- [6] 张征宇, 丁玉成, 洪军. STL模型分割截面的三角剖分算法[J]. *计算机辅助设计与图形学学报*, 2005, 17(6):1240-1245.
ZHANG ZH Y, DING Y CH, HONG J. A triangulating algorithm for cutting cross-section of STL model [J]. *Journal of Computer-Aided Design and Computer Graphics*, 2005, 17(6):1240-1245. (in Chinese)
- [7] 王红涛, 傅卫平, 康业娜. 工件图像识别的边缘匹配方法研究[J]. *仪器仪表学报*, 2008, 29(5):986-991.
WANG H T, FU W P, KANG Y N. Study of edge matching approach to workpiece image recognition [J]. *Chinese Journal of Scientific Instrument*, 2008, 29(5):986-991. (in Chinese)
- [8] SHOTTON J, BLAKE A, CIPOLLA R. Multi-scale categorical object recognition using contour fragments[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2008, 30(7):1270-1281.
- [9] 孙冰岩, 曹琦, 王星. 基于Hausdorff距离的目标跟踪方法研究[J]. *现代防御技术*, 2010, 38(5):127-130.
SUN B Y, CAO Q, WANG X. Target tracking algorithm based on Hausdorff distance[J]. *Modern Defence Technology*, 2010, 38(5):127-130. (in Chinese)
- [10] 边霞, 米良. 遗传算法理论及其应用研究进展[J]. *计算机应用研究*, 2010, 27(7):2425-2429.
BIAN X, MI L. Development on genetic algorithm theory and its applications [J] *Application Research of Computer*, 2010, 27(7):2425-2429. (in Chinese)
- [11] 牛向阳, 倪前月, 高成修. 基于遗传算法和模拟退火算法的混合算法[J]. *昆明理工大学学报:理工版*, 2008, 33(2):25-28.
NIU X Y, NI Q Y, GAO CH X. A hybrid algorithm based on genetic algorithm and simulated annealing algorithm[J]. *Journal of Kunming University of Science and Technology (Science and Technology)*, 2008, 33(2):25-28. (in Chinese)
- [12] 赵磊, 刘书桂. 基于改进遗传算法实现柔性三坐标测量机参数标定[J]. *光学精密工程*, 2011, 19(11):2753-2758.
ZHAO L, LIU SH G. Implementation of parameter calibration for flexible coordinate measurement machine based on improving genetic algorithm[J]. *Opt. Precision Eng.*, 2011, 19(11):2753-2758. (in Chinese)

作者简介:



宋 薇(1981—),女,辽宁铁岭人,博士,副研究员,2006年、2009年于日本福井大学分别获得硕士、博士学位,主要从事机器视觉及精密测量方面的研究。E-mail: song_wei@shu.edu.cn



周 扬(1990—),男,河南南阳人,硕士研究生,2013年于河南科技大学获得学士学位,主要从事机器视觉方面的研究。E-mail: actzhy@126.com