

# 基于仿射参数估计的迭代点匹配算法

谭志国 孙即祥 滕书华

(国防科技大学电子科学与工程学院 长沙 410073)

**摘要** 本文提出了一种新的迭代点匹配算法。算法建立点集间仿射映射关系,把匹配问题转化为函数优化问题,通过点集间匹配对应关系和仿射变换参数的反复迭代最终求出问题的解。文中提出了构造虚拟点对和最小方差两种仿射参数估计方法,并利用改进最近点原则求解点集匹配关系,且证明了算法的收敛性。本文算法较好地解决了由仿射带来的非刚性形变点集匹配问题,且有很好的抗噪声和点性能。实验证明了算法的有效性和鲁棒性。

**关键词** 图像匹配,点匹配,虚拟点对仿射估计,仿射变换,迭代算法

## A Novel Iterative Point Matching Algorithm Based on Affine Parameter Estimation

TAN Zhi-Guo SUN Ji-Xiang TENG Shu-Hua

(College of Electronic Science and Engineering, National University of Defense Technology, Changsha 410073)

**Abstract** A novel algorithm is proposed in this paper to solve the point matching problem. We suppose the point sets are in the relationship of affine transformation, and give out two affine parameters estimating methods, virtual pairs and least square methods. Based on this estimation utilizing the modified closest point rule, the algorithm iterates for the final exact affine parameters and corresponds. In the paper, the convergence of the algorithm is proved. Experiments on both synthetic point sets with varying degrees of affine deformation, noise and outliers, and on real image point-sets fully demonstrate the robustness of the algorithm.

**Keywords** Image registration, Point matching, Virtual point pairs estimation, Affine transformation, Iterative algorithm

## 1 引言

点匹配方法是图像特征匹配中的一个最基本和重要的方法,在物体识别、运动检测、目标跟踪等方面都得到广泛运用。但由于噪声、形变、出格点等因素的影响始终难以完全解决,因而在各个领域一直都受到广泛的关注和研究。解决点匹配问题传统方法主要有:基于特征比较的算法,如形状描述算子(shape context)算法<sup>[1]</sup>和特征向量匹配算法<sup>[2]</sup>等。该方法是在若干假设或约束下通过某种手段获取点集中点的局部或全局特征,比较得到匹配关系;基于点集变换参数估计的方法,如迭代最近点(ICP)算法<sup>[3]</sup>,基于粒子群优化(PSO)算法<sup>[4]</sup>,基于确定性退火匹配算法<sup>[5]</sup>等。该方法通过对变换参数与匹配关系的估计和优化函数的评估,迭代获得最优的变换参数和匹配关系。与第一类方法比较,第二类方法具有更好的鲁棒性。

就匹配对象而言,目前主要是针对刚性变换和相似变换来进行研究,技术较为成熟。而对于更具一般性的仿射变换匹配问题研究不多,主要的方法有:Chang<sup>[6]</sup>等人利用二维聚类的匹配算法;Huttenlocher<sup>[7]</sup>等给出的基于对准的点模式匹配算法;Gold<sup>[8]</sup>等人提出的利用软指派(soft assign)方法以及Zhang<sup>[9]</sup>等人提出的基于遗传算法的匹配方法等。上述都可归类于特征比较和参数估计两种方法。这些算法普遍存在鲁棒性较差,复杂度较高的缺点。

本文提出一种新的基于仿射参数估计的迭代点匹配算

法。算法基本思想是通过估计变换参数,来变换待匹配点集,从而寻找匹配关系,由匹配关系再计算变换参数,通过反复迭代获得最终的匹配结果和点集的变换关系。算法操作简单,计算复杂度小,鲁棒性好。

## 2 点匹配问题描述

考虑两个二维特征点集  $P = \{p_1, p_2, \dots, p_m\}$ ,  $Q = \{q_1, q_2, \dots, q_n\}$ 。假设  $Q$  是  $P$  通过仿射变换得到,则对于其中某一对应匹配点我们可用下式表示:

$$q_i = F p_j + t \quad (1)$$

其中,  $F = \begin{bmatrix} f_1 & f_2 \\ f_3 & f_4 \end{bmatrix}$  为仿射变换矩阵,  $t = (t_x, t_y)'$  为平移矢量。

按最小方差原则,可把点集匹配转化为如下函数优化问题:

$$e(F, t, m(q_i \leftrightarrow p_j)) = \sum_i \|q_i - F p_j - t\|^2 \quad (2)$$

其中,  $m(q_i \leftrightarrow p_j)$  表示点与点之间的匹配关系。由式(2)知,点集匹配主要解决两个关键问题:寻找点与点之间的匹配关系和求解点集间变换关系。通常这两个问题相辅相成,获得匹配关系即可计算得到变换关系,反之亦然,因而可通过求解一个问题来估计另一问题的解。

## 3 仿射参数估计

### 3.1 点集归一化

谭志国 博士生,研究方向为计算机视觉和图像处理;孙即祥 博士生导师,研究方向为计算机视觉、图形图像处理、模式识别及智能信息处理;滕书华 博士生,研究方向为计算机视觉、智能信息处理。

考虑完全匹配情况。点集一一对应,满足如下关系式:

$$Q=FP+th \quad (3)$$

其中,  $t=(t_x, t_y)'$ ,  $h=(1, 1, \dots, 1)$ 。设  $P$  与  $Q$  点集质心分别为  $p_{oid}=(\bar{x}, \bar{y})'$ ,  $q_{oid}=(x, y)'$ 。对所有匹配点由式(1)展开, 求和平均得:

$$\begin{cases} x=f_1\bar{x}+f_2\bar{y}+t_x \\ y=f_3\bar{x}+f_4\bar{y}+t_y \end{cases} \quad (4)$$

如果点集质心都在坐标原点时, 即  $p_{oid}=(0, 0)'$ ,  $q_{oid}=(0, 0)'$  则  $t=(0, 0)'$ 。把点集质心移至同一坐标原点, 初值估计时则可不考虑平移矢量, 有:

$$Q=FP \quad (5)$$

### 3.2 构造虚拟点对

匹配关系未知时估计仿射变换参数比较困难, 但可通过线性操作来构建点集匹配关系进行估计。下面介绍构造虚拟点对方法, 建立这种匹配关系。

经 3.1 处理后, 两点集通常在图像坐标系的 4 个象限中都有点分布。假设  $Q$  中的点与  $P$  中匹配对应点处于同一象限。由式(5)可得

$$\begin{cases} \frac{1}{k_N} \sum_{p_i \in V_N} x_i = \frac{1}{k_N} (f_1 \sum_{q_j \in V_N} \bar{x}_j + f_2 \sum_{q_j \in V_N} \bar{y}_j) \\ \frac{1}{k_N} \sum_{p_i \in V_N} y_i = \frac{1}{k_N} (f_3 \sum_{q_j \in V_N} \bar{x}_j + f_4 \sum_{q_j \in V_N} \bar{y}_j) \end{cases} \quad (6)$$

其中  $N=1, 2, 3, 4, V_N$  表示第  $N$  象限,  $k_N$  为第  $N$  象限中匹配点对数目。从而得到 4 对匹配点(4 个象限子点集质心):

$$\begin{aligned} \tilde{p}_N &= \left( \frac{1}{k_N} \sum_{q_j \in V_N} \bar{x}_j, \frac{1}{k_N} \sum_{q_j \in V_N} \bar{y}_j \right)' \\ \tilde{q}_N &= \left( \frac{1}{k_N} \sum_{p_i \in V_N} x_i, \frac{1}{k_N} \sum_{p_i \in V_N} y_i \right)' \end{aligned}$$

只要满足假设, 点集之间的仿射变换关系可通过这 4 对虚拟点直接计算。实际中, 假设一般不成立, 为减小误差, 下面给出已知匹配关系时, 最小方差仿射参数估计法。

### 3.3 最小方差仿射参数估计

不仅限于上述虚拟点对间的仿射参数估计, 更广泛地来讨论该问题。对上述  $P$  和  $Q$  点集已知其间匹配关系, 在最小方差意义下寻找尽可能满足式(3)的仿射变换参数。不失一般性, 考虑齐次坐标下变换表达式, 则式(3)可用式(7)表述

$$\tilde{Q}=\tilde{F}\tilde{P} \quad (7)$$

其中

$$\tilde{Q}=\begin{bmatrix} Q \\ 1 \end{bmatrix}, \tilde{P}=\begin{bmatrix} P \\ 1 \end{bmatrix}, \tilde{F}=\begin{bmatrix} F & t \\ 1 & 1 \end{bmatrix}$$

问题可转化为式(8)的极小化问题:

$$\tilde{\mathcal{L}}(\tilde{F})=\|\tilde{Q}-\tilde{F}\tilde{P}\|^2 \quad (8)$$

极小化  $\tilde{\mathcal{L}}$  的  $\tilde{F}$  应满足  $\frac{\partial \tilde{\mathcal{L}}}{\partial \tilde{F}}=2QP'-2\tilde{F}PP'=0$ , 则有  $\tilde{F}=QP'(PP')^{-1}$ 。该方法相当于用伪逆矩阵来求解式(7)中  $\tilde{F}$ 。

由上可得, 对未知匹配关系的情况可分三步来求解(虚拟点对参数估计法):

- 1) 按 3.1 方法消去平移参数的影响;
- 2) 构造两点集在坐标中的虚拟匹配点对;
- 3) 利用最小方差估计法获得  $F, t$ 。

对已知匹配关系的情况可以直接利用最小方差估计法。虚拟点对参数估计法用于初值估计, 在假设条件不满足时该估计方法存在误差, 而这种误差可在算法后续迭代中逐步校正。

## 4 点集的匹配

点集匹配通常可采取最近点原则<sup>[3]</sup>, 该方法主要思想为: 设点集  $Q$  中点的数目小于点集  $P$ , 定义点集  $Q$  中的一点  $q_i$  到点集  $P$  的最近距离为:

$$d(q_i, P)=\min_{p_j \in P} \|q_i - p_j\| \quad (9)$$

由最近距离决定点集  $Q$  中每一点在点集  $P$  中的对应点, 若  $Q$  中有两个点对应  $P$  中同一点, 则选择距离最小的点为对应点。

考虑  $P$  与  $Q$  为不完全匹配, 且两者都存在出格点, 当  $P$  中真实匹配点都找到相应最近的对应点时,  $P$  中出格点由于没有真实匹配点, 则会寻找  $Q$  中出格点, 从而导致误差。在这种情况下, 须对匹配关系进行控制。本文设定阈值  $\sigma$ , 当最近点之间的距离大于该阈值时则作为出格点处理, 此时  $d(q_i, P)=\sigma$ 。

设  $m \leq n$ ,  $P$  与  $Q$  有  $k$  对匹配点, 通过上述处理, 点匹配可转化为求解如下问题:

$$e=\min_{F, t, m} \left[ \sum_i \|q_i - Fp_i - t\|^2 + (n-k)\sigma^2 \right] \quad (10)$$

其中, 第二项表示  $Q$  中出格点的惩罚项。

## 5 迭代算法

### 5.1 算法步骤

考虑不完全匹配,  $m \leq n$ 。算法步骤如下:

1) 获取初值。首先按 3.1 节, 进行归一化; 然后采用虚拟点对参数估计法获得  $F_0$ , 变换点集  $P$  获得  $\bar{P}_0$ 。最后, 利用第 4 节方法获取一一对应大小为  $k_0$  的点集  $P_0$  和  $Q_0$ , 由式(10)计算  $e_0$ 。

2) 若满足收敛条件  $e_t < \omega$  或达到最大迭代步数, 则迭代停止; 否则进入 3)。

3) 按 3.3 节方法由  $P_{t-1}$  和  $Q_{t-1}$  来估计  $F_t, t_t$ 。

4) 由  $\bar{P}_{t-1}$  通过  $F_t, t_t$  变换获得点集  $\bar{P}_t$ , 利用第 4 节中方法建立对应点集  $P_t$  和  $Q_t$ 。根据式(10)计算匹配误差, 返回 2)。

### 5.2 收敛性质

算法收敛关键在于  $e_t$  是否收敛, 以下证明  $e_t$  的收敛性质。

对于  $\forall q_i \in Q$ , 在第  $t-1$  次迭代时, 可分为两类, 即匹配点和未匹配点。匹配点的集合为  $Q_{t-1}$ , 设未匹配点集合则为  $\bar{Q}_{t-1}$ , 有

$$\begin{cases} d(q_i, \bar{P}_{t-1}) = \min \|q_i - \bar{p}_{t-1}\| < \sigma, & q_i \in Q_{t-1} \\ d(q_i, \bar{P}_{t-1}) = \sigma, & q_i \in \bar{Q}_{t-1} \end{cases} \quad (11)$$

因而

$$e_{t-1} = \sum_{q_i \in Q_{t-1}} d^2(q_i, \bar{P}_{t-1}) + \sum_{q_i \in \bar{Q}_{t-1}} d^2(q_i, \bar{P}_{t-1}) \quad (12)$$

因  $F_t, t_t$  是通过最小方差法估计获得, 变换后有

$$\begin{aligned} \sum_{i=1}^{k_{t-1}} \|q_i - F_t \bar{p}_{t-1} - t_t\| &\leq \sum_{i=1}^{k_{t-1}} \|q_i - \bar{p}_{t-1}\| \\ &= \sum_{q_i \in Q_{t-1}} d(q_i, \bar{P}_{t-1}) \end{aligned} \quad (13)$$

其中,  $q_i, \bar{p}_{t-1}$  为第  $t-1$  次迭代分别属于  $Q_{t-1}, P_{t-1}$  的匹配点对。又  $P_t$  是由  $Q_t$  通过 4 中原则获得, 则  $\forall q_i \in Q_{t-1}$ , 必有下列情况之一

$$\begin{cases} d(q_i, P_t) = \min \|q_i - \bar{p}_t\| \leq \|q_i - F_t \bar{p}_{t-1} - t_t\| < \sigma, & q_i \in Q_t \\ d(q_i, P_t) = \min \|q_i - \bar{p}_t\| < \sigma \leq \|q_i - F_t \bar{p}_{t-1} - t_t\|, & q_i \in \bar{Q}_t \end{cases} \quad (14)$$

而  $\forall q_i \in \bar{Q}_{t-1}$ , 必有下列情况之一

$$\begin{cases} d(q_i, \bar{P}_t) = \min \|q_i - \bar{p}_t\| < d(q_i, \bar{P}_{t-1}) = \sigma, q_i \in \bar{Q}_t \\ d(q_i, \bar{P}_t) = d(q_i, \bar{P}_{t-1}) = \sigma, q_i \in \bar{Q}_t \end{cases} \quad (15)$$

由(13), (14), (15)可得

$$\begin{aligned} e_{t-1} &= \sum_{q_i \in \bar{Q}_{t-1}} d^2(q_i, \bar{P}_{t-1}) + \sum_{q_i \in \bar{Q}_{t-1}} d^2(q_i, \bar{P}_{t-1}) \\ &\geq \sum_{q_i \in \bar{Q}_{t-1}} \|q_i - F_t \bar{p}_{t-1} - t_t\|^2 + \sum_{q_i \in \bar{Q}_{t-1}} d^2(q_i, \bar{P}_{t-1}) \\ &\geq \sum_{q_i \in \bar{Q}_{t-1}} d^2(q_i, \bar{P}_t) + \sum_{q_i \in \bar{Q}_{t-1}} d^2(q_i, \bar{P}_t) \\ &= \sum_{q_i \in \bar{Q}_t} d^2(q_i, \bar{P}_t) + \sum_{q_i \in \bar{Q}_t} d^2(q_i, \bar{P}_t) = e_t \end{aligned}$$

所以对  $\forall t$  有  $0 \leq e_t \leq e_{t-1} \leq n\sigma^2$ , 即  $\{e_t\}$  收敛。

## 6 实验结果与分析

为了验证算法的有效性和鲁棒性, 进行四组实验, 每组实验各进行 50 次。前三组用模拟点集进行匹配, 分别测试算法抗仿射形变能力, 抗噪性能和抗出格点性能。第四组对真实图像进行匹配。每组实验中, 与清华大学 Zhang<sup>[11]</sup> 等人提出的基于遗传算法(GA)的仿射点匹配方法进行比较。限于篇幅, 每组只给出其中一次实验的结果。

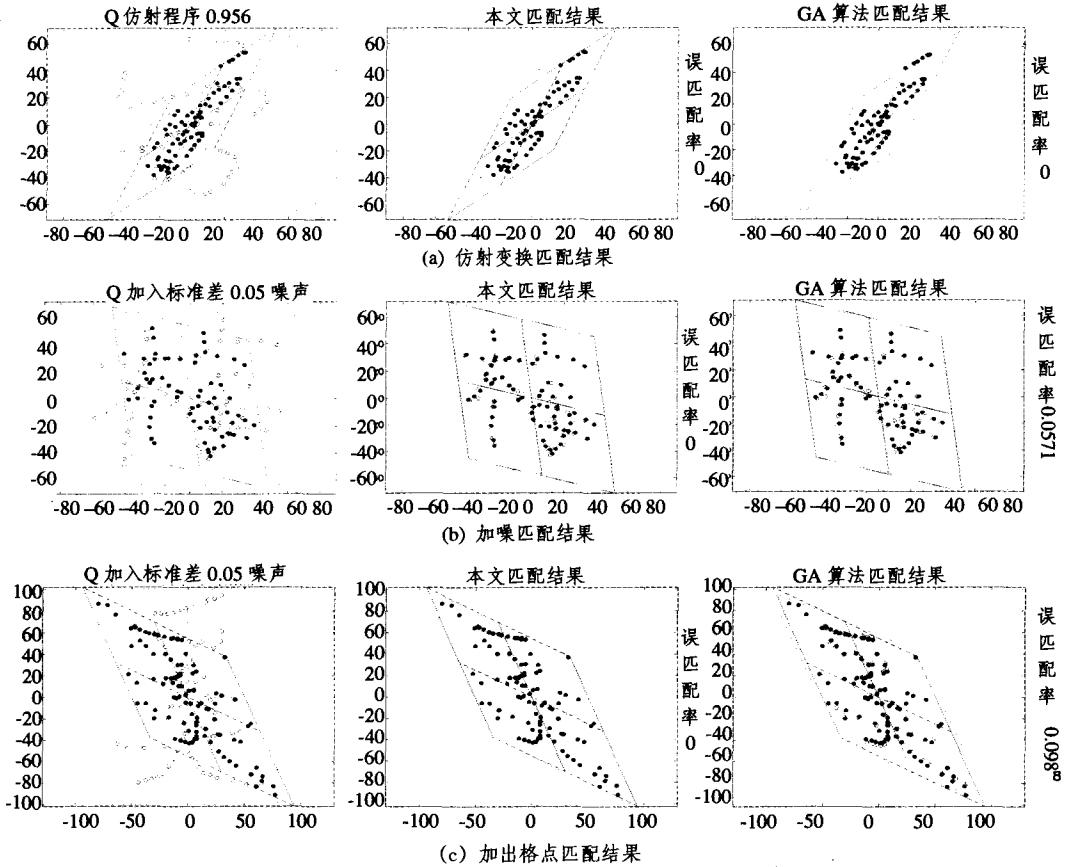


图 1 模拟点集匹配结果

### 6.1 实验结果

前三组实验中, 首先产生模拟点集  $P$ , 图中用空心点表示; 由  $P$  通过仿射变换、加噪、加出格点等方法产生  $Q$  点集, 图中用实心点表示。为直观, 图中都画出了两点集所在的仿

射坐标系, 浅色为点集  $P$  的, 深色为点集  $Q$  的。匹配结果的好坏, 可由两坐标系空间上的差异看出。如果两坐标系完全重叠, 则表示匹配效果好。

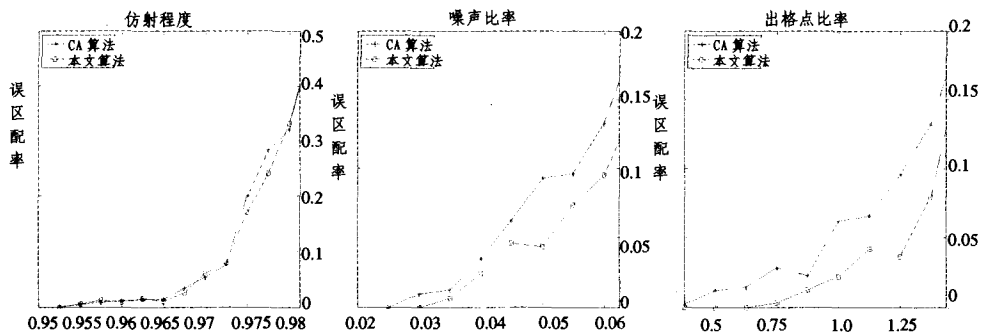


图 2 统计结果

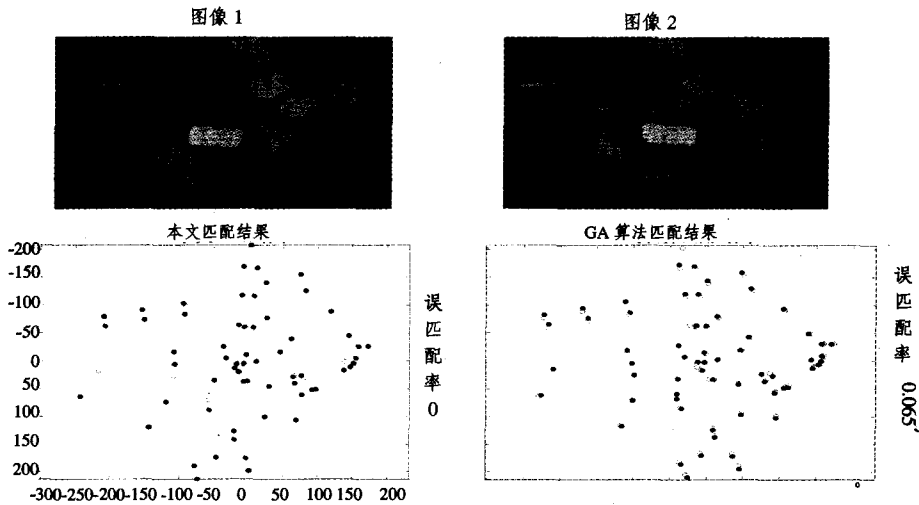


图3 真实图像匹配结果

在第一组实验中,分别用不同程度的仿射形变点集对算法进行测试,目的测试算法抗仿射形变能力。这里仿射形变程度用椭圆的离心率表示<sup>[12]</sup>,离心率越大,则表示仿射形变程度越大,图 1(a)给出了其中一次的实验结果。

第二组实验,用于测试算法的抗噪性能,我们在一定程度的仿射变换情况下,分别加入了不同标准差的高斯噪声,图 1(b)为加入标准差为 0.05 高斯噪声的匹配情况。

第三组实验用于测试算法抗出格点性能,点集在一定程度的仿射变换后分别加入不同比率的出格点,图 1(c)为加入 0.5 比率出格点情况。图 2 为对上述几组实验的统计结果。

第四组实验,用多组真实照片,分别从各组图像中提取特征点集进行匹配,共进行 50 次实验,平均匹配正确率为 96.64%。图 3 为其中一次实验的结果。

### 6.2 分析

#### 1) 迭代过程

实验中,对算法进行跟踪获得迭代过程,如图 4。图中左上角的坐标系分别表示两点集在该时刻所处的仿射坐标系,浅色为点集 P,深色为点集 Q。坐标系的变化表示的是点集的变化,也是整个算法迭代过程的体现。

由图 4 第一行到第二行,点集 P 有一个较大的变化过程。这是通过构造虚拟点对,初始参数估计得到的变换。它在整个算法中是非常关键的一步。当未知匹配关系时,直接利用 3.3 节方法不能获得准确或近似的变换关系。通过构造虚拟点对来估计仿射参数,能较好估计点集之间的变换情况,为寻找匹配关系和估计仿射参数搭建良好平台。在此基础上,由第二行根据 4 中的方法获得匹配关系再次估计仿射参数,变换 P 得到第三行的结果,可以看出该步的变换在向真实的变换趋近。随着估计仿射参数趋于真实值,反过来作用于匹配关系的寻找。如此反复,直至最后达到收敛精度。

#### 2) 初值估计

初值估计是算法的关键之一。文中提出了利用构造虚拟点对估计初值的方法,然而当点集之间有较大旋转时,该估计方法失效。为此对该方法进行改进。

点集的旋转是一种线性变换,不会改变点集中心在空间的分布顺序,因而,在获取点集的虚拟点  $\tilde{p}_n, \tilde{q}_n$  后,建立匹配关系时不仅仅按照第 3 节中方法象限——对应,同时也建立

另外三组匹配关系,  $\tilde{p}_{n+1} \leftrightarrow \tilde{q}_{(n \oplus i)+1}, n=0,1,2,3, i=1,2,3, \oplus$  表示模 4 和。由此,得到四组参数估计值,取最优作为算法初值。

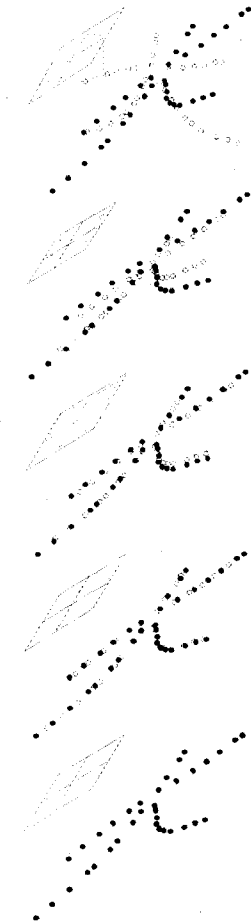


图4 迭代过程

#### 3) 阈值 $\sigma$ 的选取

阈值  $\sigma$  的选取对算法有一定的影响,选值适当则可减少算法的迭代次数,增加算法精度。由于迭代初期误差较大,后期随着参数和匹配关系逐步趋向正确值,误差减小,则匹配条件相应更为严格。因此我们可以采取确定性退火策略利用退火温度  $T$  来不断减小  $\sigma$  值,则  $e_i$  变为式(19):

$$e_i = \sum_{j=1}^{k_i} \|q_i - \bar{p}_j\|^2 + (n - k_i)(T\sigma)^2 \quad (19)$$

通过退火温度的控制可使  $e_i$  以较快速度收敛。初始阈值  $\sigma$  可选择为点集中点之间最小距离,  $T$  为 10~20 之间, 退火速率为 0.8~0.95 之间。

#### 4) 比较分析

实验统计结果表明, 本文算法与基于 GA 的算法的抗仿射形变能力不相上下, 但抗噪性能和抗出格点性能则更胜一筹。真实图像实验也能说明, 本文算法更具实用性。就时间复杂度来说, 基于 GA 的算法每一次迭代复杂度为  $O(mn) \times N$ , 其中  $N$  为遗传种群的规模, 取值为 100~1000。本文算法每次迭代复杂度至多为  $O(mn)$ , 因而从复杂度上小于基于 GA 的算法。此外, 基于 GA 的算法需要较多的迭代步数(60~100)才能达到收敛, 本文算法只需较少的迭代(5~20)则完全收敛。

**结论** 本文提出一种新的基于仿射参数估计的迭代点匹配算法, 很好地解决了由仿射带来的非刚性形变点集匹配问题。该算法把点匹配基本问题转化为函数优化问题, 通过对点集间变换关系和匹配关系的反复估计迭代得到问题的最优解。文中给出了在匹配未知和已知两种情况下的仿射参数估计方法。在点集间匹配关系未知情况下, 我们通过构造虚拟点对估计仿射参数, 该方法简单有效, 为算法提供了良好的初值。在匹配关系已知时, 利用最小方差法进行参数估计, 它确保了算法的收敛性。利用改进的最近点原则求解点集的匹配关系, 减小了算法匹配误差, 加快了收敛速度。文章还对虚拟点对初值估计方法和阈值  $\sigma$  的选取问题进行了讨论, 并给出

了改进办法。通过实验, 并与基于 GA 的匹配方法进行比较, 说明算法具有较强的鲁棒性和实用性。

## 参考文献

- 1 Belongie S, Malik J, Puzicha J. Shape Matching and Object Recognition Using Shape Contexts. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2002, 24(4): 509~522
- 2 Shapiro L S, Brady J M. Feature-Based correspondence: an eigenvector approach. *Image and Vision Computing*, 1992, 10(5): 283~288
- 3 Besl P J, McKay N D. A method for registration of 3-D shapes. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 1992, 14(2): 8~14
- 4 Talbi H, Batouche M C. Particle Swarm Optimization for Image Registration. In: 2004 International Conference on Information and Communication Technologies: From Theory to Applications, 2004, 4: 397~398
- 5 Chui H, Rangarajan A. A New Algorithm for Non-Rigid Point Matching. In: Proceedings IEEE Conference on Computer Vision and Pattern Recognition, 2000, 2: 44~51
- 6 Chang S, Cheng F, Hsu Wen-hsing, et al. Fast algorithm for point pattern matching: invariant to translations, rotations and scale changes. *Pattern Recognition*, 1997, 29(1): 11~16
- 7 Huttenlocher D P, Ullman S. Object recognition using alignment. In: Proceeding of the First International Conference on Computer Vision, London, 1987. 102~111
- 8 Gold S, Rangarajan A, Lu C, et al. New algorithms for 2D and 3D point matching pose estimation and correspondence. *Pattern Recognition*, 1998, 31(8): 1019~1031
- 9 Zhang L, Xu W, Chang C. Genetic algorithm for affine point pattern matching. *Pattern Recognition Letters*, 2003, 24(1-3): 9~19
- 10 孙焱, 王秀坤, 邵刚, 冯林, 贺明峰. 二维点模式图像的仿射变换配准. *计算机辅助设计与图形学学报*, 2005, 7(17): 1497~1503

(上接第 203 页)

- 9 Masreliez C J, Martin R D. Robust Bayesian estimation for the linear model and robustifying the Kalman filter. *IEEE Transactions on Automatic Control*, 1977, AC 22(3): 361~371
- 10 Hawkins D. Identification of Outliers. London: Chapman and Hall, 1980
- 11 Knorr E M, Ng R T. A unified notion of outliers: Properties and computation. In: Proc. KDD 1997, 1997. 219~222
- 12 Knorr E M, Ng R T. Algorithms for mining distance-based outliers in large datasets. In: Proc. VLDB 1998, 1998. 392~403
- 13 Knorr E M, Ng R T. Finding intentional knowledge of distance-based outliers. In: Proc. VLDB 1999, 1999. 211~222
- 14 Knorr E M, Ng R T, Tucakov V. Distance-based outliers: Algorithms and applications. *VLDB Journal*, 2000, 8: 237~253
- 15 Jiang F, Sui Y F, Cao C G. Outlier Detection Using Rough Set Theory. In: RSFDGrC 2005, LNAI, 2005. 79~87
- 16 Cao Feng, Ester M, Qian Weining, et al. Density-Based Clustering over an Evolving Data Stream with Noise. In: SDM'2006
- 17 John G H. Robust Decision Trees: Removing Outliers from Databases. In: Proceedings of the First International Conference on Knowledge Discovery and Data Mining, 1995. 174~179
- 18 Liu H C, Shah S, Jiang W. On-line outlier detection and data cleaning. *Computers and Chemical Engineering*, 2004, 28: 1635~1647
- 19 Jagadis H V, Koudas N, Muthukrishnan S. Mining Deviants in a Time Series Database. VLDB, Edinburgh: Morgan Kaufmann Publishers, 1999. 102~113
- 20 Han J W, Kamber M. Data mining: concepts and techniques. New York: Morgan Kaufmann Publishers, 2001
- 21 Davies L, Gather U. The identification of multiple outliers. *Journal of the American Statistical Association*, 1993, 88: 782~792
- 22 Otey M E, Ghoting A, Parthasarathy S. Fast Distributed Outlier Detection in Mixed-Attribute Data Sets. *Data Mining and Knowledge Discovery*, 2005
- 23 Perarson R K. Outliers in Process Modeling and Identification. *IEEE Transactions on Control Systems Technology*, 2002, 10: 55~63
- 24 Tsay R S. Outliers, level shifts, and variance changes in time series. *Journal of Forecasting*, 1998, 7: 1~20
- 25 Tsay R S. Time series model specification in the presence of outliers. *Journal of the American Statistical Association*, 1996, 81: 132~141
- 26 Martin R D, Thomson D J. Robust-resistant spectrum estimation. In: Proceeding of the IEEE 1982, 70: 1097~1115
- 27 Papadimitriou S, Kitawaga H, Gibbons P, et al. LOCI: Fast Outlier Detection Using the Local Correlation Integral. In: Proc of the International Conference on Data Engineering, 2003. 315~326
- 28 Muthukrishnan S, Shah R, Vitter J S. Mining Deviants in Time Series Data Streams. In: Proc. of the 16th Int'l Conf. on Scientific and Statistical Database Management. Santorini Island: IEEE Computer Society, 2004. 41~50
- 29 Ramaswamy S, Rastogi R, Kyuseok S. Efficient Algorithms for Mining Outliers from LargeData Sets. In: SIGMOD'00 2000. 427~438
- 30 Johnson T, Kwok I, Ng R T. Fast computation of 2-dimensional depth contours. In: Proc. KDD 1998, 1998. 224~228
- 31 Barnett V, Lewis T. Outliers in Statistical Data. John Wiley, 1994
- 32 Hodge V J, Austin J. A Survey of Outlier Detection Methodologies. *Artificial Intelligence Review*, 2004, 22: 85~126
- 33 Yang Y D, Sun Z H, Zhu Y Q, et al. A Fast Outlier Detection Algorithm for Data Streams Based on Dynamic Grids. *Journal of Software*, 2006, 17(8): 1796~1803
- 34 Tao Y F, Xiao X K, Zhou S G. Mining Distance-based Outliers from Large Databases in Any Metric Space. *KDD'06*, 2006. 394~403
- 35 Breuning M M, Kriegel H P, Ng R T, et al. Lof: Identifying density-based local outliers. In: Proc. ACM SIGMOD Conf 2000, 2000. 93~104
- 36 Cui Hongyin. Online Outlier Detection Over Data Streams. Paper for Master Degree. Smon Fraser University, 2005
- 37 He Z Y, Xu X F, Huang J Z, et al. A Frequent Pattern Discovery Method for Outlier Detection. *WAIM 2004*, LNCS 3219, 2004. 726~732
- 38 金澈清, 钱卫宁, 周傲英. 数据流分析与管理综述. *软件学报*, 2004, 15(8): 1172~1181