一种改进的图像低维表示方法

曾步衢
（黄淮学院，驻马店 463000）

摘要：目的 解决当前方法需要对图像中的相应点手动标记界标，且局限于特定对象或形状变形的问题。**方法** 提出一种可以同时实现图像颜色、外观和形态的图像低维表示算法。**结果** 该算法通过将形态和外观的流形约束到低维子空间上，进一步降低了流形学习的采样复杂性。**结论** 文中方法的性能远优于目前典型的稳健型光流算法和 SIFT 流算法。在图像编辑和关节学习关任务中取得了令人满意的定性结果。

关键词： 图像低维表示；L_2 范数；稳健型光流算法；SIFT 流算法
中图分类号： TP393 **文献标识码：** A **文章编号：** 1001-3563(2017)09-0230-06

An Improved Method of Low-Dimensional Representation of Images

ZENG Bu-qu
(Huanghuai University, Zhumadian 463000, China)

ABSTRACT: The work aims to solve the problem that the existing solutions either require manually specified landmarks for corresponding points in the images, or are restricted to specific objects or shape deformations. A low-dimensional representation of images for simultaneously recovering color, appearance and shape was proposed. The proposed algorithm further reduced sample complexity of manifold learning as the manifolds of shape and appearance were restricted to low-dimensional subspaces. The proposed method significantly outperformed the current typical methods of robust optical flow and SIFT flow. Our qualitative results in some related tasks such as image deformation and joint learning are encouraging.

KEY WORDS: low-dimensional representation of images; L_2 norm; robust optical flow; SIFT flow

计算机视觉中的一个基本问题就是利用某种参数化方法来描述图像的“空间”特征[1—2]，其中，参数对应于目标数据的相关变化。这种图像表示方法对于图像形态转换或姿态变换、目标检测和识别等应用具有重要作用[3—6]。图像的表示问题从本质上来说属于自然图像的流形学习范畴。有 2 种类型的流形方法可用于这一领域。第 1 种方法是进行局部流形分析[7—8]。这种方法总是忽略图像的空间结构，而图像的空间结构对目标级别的图像表示具有重大意义。第 2 种方法是将流形分析应用于整个图片[9—10]，它们要求对图片空间进行非常密集的采样，但是，密集采样约束的限制性很强。例如，需要大量图像才能密集采样到图像的形态和颜色的所有缓慢变化。

针对以上方法的不足，文中将图像低维表示问题建模为基于 L_2 范数的组合优化问题，通过采取嵌套和组合策略，考虑了颜色、外观和形态方面的变化，进而提出一种可以同时实现图像颜色、外观和形态的图像低维表示方法。此外，通过将形态和外观的流形约束到低维子空间上，进一步降低了流形学习的采样复杂性。文中方法与文献[11]中的主动外观模型和文献[12]中的 Morphable 模型类似。文献[11]中的算法要求人工提供界标，而文献[12]中的方法需要图像预先对准。相反，文中方法为完全无监督模型。另外，文献[13]中的变换成份分析法(TCA)通过排除形态形变，实现低维外观子空间采集。TCA 方法必须对整个（离散）参数空间进行穷尽式处理，所以对高度复杂运动的可拓展性较差。相反，对于利用密集运动域表示的任意运动，文中方法均可处理。

收稿日期： 2016-03-31
基金项目： 河南省教育厅重点科技攻关项目(13A520786)
作者简介： 曾步衢（1979—），男，硕士，黄淮学院讲师，主要研究方向为图像处理、大数据处理。

1　系统模型

1.1　相关定义

为了便于阐述，下面首先给出了文中用到的相关符号的含义。其中，χ 为集合，$\boldsymbol{x}$ 为向量，$\boldsymbol{X}$ 为矩阵。$f(\ )$为标量函数，$\boldsymbol{f}(\ )$为矢量函数，$F[\ ]$为泛函。$\mathfrak{X}$ 为子空间，$\equiv$ 表示相等。集合 $\{x_1\cdots x_n\}$ 为 $\{x_k\}_{k=1}^n$，符号$\|\ \|$表示$\|\ \|_2$，$|\chi|$表示 χ 离散集合时 χ 的元素数量，当 χ 是连续集合时定义 $|\chi|\equiv\int_\chi \mathrm{d}\boldsymbol{x}$。运算符 $\mathrm{vec}(f(\boldsymbol{x}),\ \chi')$定义为关于 f: $\chi\to\mathbf{R}$ 的映射并在均匀分布的网格 $\chi'\subset\chi$ 上求值。它将所有值串联起来返回一个向量 $\boldsymbol{f}$。该向量的尺寸显然为$|\chi'|$。通过在 $\boldsymbol{f}$ 的元素间做插值处理，运算符 $\mathrm{unvec}(\boldsymbol{f},\chi)$可近似表示域 χ 上的初始映射 $\boldsymbol{f}$。如果 $\boldsymbol{f}(\boldsymbol{x})$是由 γ 个成份$(f_1(\boldsymbol{x})\ldots f_r(\boldsymbol{x}))$构成的多值函数，则定义 $\mathrm{vec}(f(\boldsymbol{x}),\chi')=(\mathrm{vec}(f_1(x),\chi')\ldots\mathrm{vec}(f_\gamma(x),\chi'))$，即将 $f(x)$的所有单值成份串联起来。此时，运算符 $\mathrm{unvec}(f,\chi)$的定义与单值映射类似，唯一区别就是它通过在 f 的元素间做插值处理后可以近似原来的多值函数 f。

1.2　问题描述

设 $\chi\subset\mathbf{R}^2$ 和 $C\subset\mathbf{R}^3$ 分布表示图像空间的位置和颜色。已知一组输入颜色图像 $F\equiv\{\boldsymbol{f}_k(\boldsymbol{x})\}_{k=1}^n$，其中每个图像 $\boldsymbol{f}_k$: $\chi\to C$。文中采用函数组合来模拟每个输入图像 $\boldsymbol{f}_k(\boldsymbol{x})\in F$：

$$\boldsymbol{f}_k(\boldsymbol{x})\approx\boldsymbol{h}_k(\boldsymbol{g}_k(\boldsymbol{u}_k(\boldsymbol{x}))),\ k=1\ldots n \tag{1}$$

其中映射为 $\boldsymbol{h}_k$:$C\to C$, $\boldsymbol{g}_k$:$\chi\to C$ 且 $\boldsymbol{u}_k$:$\chi\to\chi$。这种嵌套形式为图像 F 的生成添加了一种具体的隐性结构。根据每个映射的性质，可以将它们分别看成是光度转换、外观转换和几何转换。文中的目标是根据已知数据 F 估计出这些映射。采用 L_2 范数来定量描述近似误差，因此需要求解如下优化问题：

$$\min_{\{(\boldsymbol{h}_k,\boldsymbol{g}_k,\boldsymbol{u}_k)\}_{k=1}^n}\sum_{k=1}^n\int_\chi\left\|\boldsymbol{h}_k\left(\boldsymbol{g}_k(u_k(\boldsymbol{x}))\right)-\boldsymbol{f}_k(\boldsymbol{x})\right\|^2\mathrm{d}x \tag{2}$$

式（2）是一种病态问题[2]，为了求解该问题，文中通过选择性能相同的映射 $\{\boldsymbol{h}_k(\mathrm{c}),\boldsymbol{g}_k(\boldsymbol{x}),\boldsymbol{u}_k(\boldsymbol{x})\}_{k=1}^n$ 来近似 F。文中首先对解空间大小进行约束，即根据少量的基函数构建映射 $\{\boldsymbol{g}_k(\boldsymbol{x})\}_{k=1}^n$ 和 $\{\boldsymbol{u}_k(\boldsymbol{x})\}_{k=1}^n$ 来控制这些映射的容量。此外，通过对几何变换施加空间平滑度约束来对问题进行正规化处理，还对每个颜色变换函数 $\boldsymbol{h}_k$ 施加平移和旋转约束，即 $\boldsymbol{h}_k(\boldsymbol{c})\equiv\boldsymbol{A}_k\boldsymbol{c}+\boldsymbol{b}_k$。此时，旋转矩阵 $\boldsymbol{A}_k$ 为 3×3 矩阵，向量 $\boldsymbol{b}_k$ 为 3×1 向量，因此，正规化问题表述为：

$$\min_{\{(\boldsymbol{A}_k,\boldsymbol{b}_k,\boldsymbol{g}_k,\boldsymbol{u}_k)\}_{k=1}^n}\sum_{k=1}^n\left(\int_\chi\left\|\boldsymbol{A}_k\boldsymbol{g}_k\left(\boldsymbol{u}_k(\boldsymbol{x})\right)+\boldsymbol{b}_k-\boldsymbol{f}_k(\boldsymbol{x})\right\|^2\mathrm{d}\boldsymbol{x}+R[\boldsymbol{u}_k]\right) \tag{3}$$

条件：$\forall k;\boldsymbol{g}_k(\boldsymbol{x})\in G,\boldsymbol{u}_k(\boldsymbol{x})\in U,\boldsymbol{A}_k\in SO(3)$

式中：R 为对其自变量的空间变化进行惩罚的泛函。空间 G 和 U 具有如下属性：存在 d_g 个和 d_u 个基函数 $\{\varnothing_p(\boldsymbol{x})\}_{p=1}^{d_g}$ 及 $\{\psi_p(\boldsymbol{x})\}_{q=1}^{d_u}$，满足 $G=\mathrm{span}\left(\varnothing_1(\boldsymbol{x}),\ldots,\varnothing_{d_g}(\boldsymbol{x})\right)$ 且 $U=\mathrm{span}\left(\psi_1(\boldsymbol{x}),\ldots,\psi_{d_u}(\boldsymbol{x})\right)$。

2　问题求解

式（3）为非凸问题，因此求解该问题的难度很大。文中通过分别求解各个子问题来获得该问题的近似解。具体而言，式（3）中有 3 组变量，即颜色变换 $\{(\boldsymbol{A}_k,\boldsymbol{b}_k)\}_{k=1}^n$，外观 $\{\boldsymbol{g}_k(\boldsymbol{x})\}_{k=1}^n$，及形态形变 $\{\boldsymbol{u}_k(\boldsymbol{x})\}_{k=1}^n$。通过一次只关注一种类别（固定其他 2 项），可以获得约简目标函数（3）的高效更新规则。这些更新规则要么具有闭合形式的表达式，要么存在高效的求解算法。在下文中，将逐个分析这些子问题，最终将各部分综合起来给出一个完整的求解算法。

2.1　求解 $\{\boldsymbol{g}_k(\boldsymbol{x})\}_{k=1}^n$

根据式（1）可知，当 $k=1\ldots n$ 时，文中方法希望满足 $\boldsymbol{f}_k(\boldsymbol{x})\approx\boldsymbol{A}_k\boldsymbol{g}_k(\boldsymbol{u}_k(\boldsymbol{x}))+\boldsymbol{b}_k$。其中，旋转矩阵 $\boldsymbol{A}_k$ 显然是可逆矩阵。假设形态形变可逆，或者形态形变的逆可被近似。可将所提方法要求表示为：$\boldsymbol{g}_k(\boldsymbol{x})\approx\boldsymbol{A}_k^{-1}(\boldsymbol{f}_k(\boldsymbol{u}_k^{-1}(\boldsymbol{x}))-\boldsymbol{b}_k)$。因为 $\{(\boldsymbol{A}_k,\boldsymbol{b}_k)\}_{k=1}^n$ 和 $\{\boldsymbol{u}_k(\boldsymbol{x})\}_{k=1}^n$ 固定，所以优化子问题可表述如下：

$$\min_{\{\boldsymbol{g}_k\}_{k=1}^n}\sum_{k=1}^n\int_\chi\left\|\boldsymbol{g}_k(\boldsymbol{x})-\boldsymbol{A}_k^{-1}\left(\boldsymbol{f}_k\left(\boldsymbol{u}_k^{-1}(\boldsymbol{x})\right)-\boldsymbol{b}_k\right)\right\|^2\mathrm{d}\boldsymbol{x} \tag{4}$$

条件：$\forall k;\boldsymbol{g}_k(\boldsymbol{x})\in G$

假设 $z(\boldsymbol{x})\equiv\boldsymbol{A}_k^{-1}\left(\boldsymbol{f}_k\left(\boldsymbol{u}_k^{-1}(\boldsymbol{x})\right)-\boldsymbol{b}_k\right)$，且与 $\boldsymbol{g}_k(\boldsymbol{x})$无关。问题成为子空间的秩为 d_g 时使 L_2 重建误差 $\sum_{k=1}^n\int_\chi\|\boldsymbol{g}_k(\boldsymbol{x})-z(\boldsymbol{x})\|^2\mathrm{d}\boldsymbol{x}$ 最小化。该问题的解与协方差运算符 $\{z_k(\boldsymbol{x})\}_{k=1}^n$ 的特征函数相关。

为了求解式（4），文中利用 1.1 节定义的 vec 运算符，将谱分解用于离散化 $\boldsymbol{g}_k(\boldsymbol{x})$和 $z_k(\boldsymbol{x})$的协方差。设置 $\tilde{z}_k\equiv\mathrm{vec}(z_k(\boldsymbol{x}))$ 且 $\bar{z}\equiv\frac{1}{n}\sum_{k=1}^n\tilde{z}_k$。于是，可以计算出 $\sum_{k=1}^n(\tilde{z}_k-\bar{z})(\tilde{z}_k-\bar{z})^T$ 的前 d_g 个特征向量，即 $(\varnothing_1\cdots\varnothing_{d_g})$。于是，有 $\tilde{\boldsymbol{g}}_k=\bar{z}+\sum_{p=1}^{d_g}\varnothing_p\varnothing_p^T(\tilde{z}_k-\bar{z})$。

考虑到 $\boldsymbol{u}_k(\boldsymbol{x})$并不一定可逆，可能存在部分位置 x，使得 $\boldsymbol{u}_k^{-1}(\boldsymbol{x})$ 未被定义或者具有多个值，因此，文中没有采用 $\boldsymbol{u}_k^{-1}$，而是定义一个近似 $\boldsymbol{u}_k^{-1}$ 的函数 $\boldsymbol{u}_k^{\#}:\chi\to\chi$。利用式(5)来解决 $\boldsymbol{u}_k^{-1}$ 的多值问题。

$$\forall\boldsymbol{x}\in\chi;\ \left|\boldsymbol{u}_k^{-1}(\boldsymbol{x})\right|\geqslant 2\Rightarrow\boldsymbol{u}_k^{\#}(\boldsymbol{x})=\arg\min_{\hat{\boldsymbol{u}}\in\boldsymbol{u}_k^{-1}(\boldsymbol{x})}\|\hat{\boldsymbol{u}}-\bar{\boldsymbol{u}}_{\mathrm{x}}\| \tag{5}$$

式中：$\bar{\boldsymbol{u}}_{\boldsymbol{x}}\equiv\frac{1}{|U_{\boldsymbol{x}}|}\sum_{\tilde{\boldsymbol{u}}\in U_{\boldsymbol{x}}}\tilde{\boldsymbol{u}}$，$U_{\boldsymbol{x}}\equiv\bigcup_{\boldsymbol{y}\in N(\boldsymbol{x})}\{\boldsymbol{u}_k^{-1}(\boldsymbol{y})\}$，$N(\boldsymbol{x})$为 $\boldsymbol{x}$ 的一个小型邻域。处理完多值点后，开始处理 $\boldsymbol{u}_k^{-1}(\boldsymbol{x})$ 没有值的点 $\boldsymbol{x}$。可以利用具有数值的剩余 $\boldsymbol{u}_k^{\#}$

来插值，进而填充这些点。在下文中用 $\boldsymbol{u}^{-1}$ 表示 $\boldsymbol{u}^{\#}$，因为 $\boldsymbol{u}^{-1}$ 更能表示逆概念。

2.2 求解 $\{\boldsymbol{u}_k(\boldsymbol{x})\}_{k=1}^n$

因为 $\{(\boldsymbol{A}_k,\boldsymbol{b}_k)\}_{k=1}^n$ 和 $\{\boldsymbol{g}_k(\boldsymbol{x})\}_{k=1}^n$ 固定，所以式（3）优化问题可简化为：

$$\min_{\{\boldsymbol{u}_k\}_{k=1}^n}\sum_{k=1}^n\left(\int_\chi\left\|\boldsymbol{A}_k\boldsymbol{g}_k(\boldsymbol{u}_k(\boldsymbol{x}))+\boldsymbol{b}_k-\boldsymbol{f}_k(\boldsymbol{x})\right\|^2\mathrm{d}\boldsymbol{x}+R[\boldsymbol{u}_k]\right) \quad (6)$$

条件：$\forall k;\boldsymbol{u}_k(\boldsymbol{x})\in U$

为了提高效率，文中将优化问题松弛为 2 个更为简单的子问题。首先求解没有子空间约束的问题（6），然后将问题的解投影到子空间上。具体而言，首先计算：

$$\hat{\boldsymbol{u}}_k(\boldsymbol{x})\equiv\arg\min_{\boldsymbol{u}_k}\int_\chi\left\|\boldsymbol{A}_k\boldsymbol{g}_k(\boldsymbol{u}_k(\boldsymbol{x}))+\boldsymbol{b}_k-\boldsymbol{f}_k(\boldsymbol{x})\right\|^2\mathrm{d}\boldsymbol{x}+R[\boldsymbol{u}_k] \quad (7)$$

然后求解如下问题：

$$\min_{\{\boldsymbol{u}_k\}_{k=1}^n}\sum_{k=1}^n\int_\chi\left\|\boldsymbol{u}_k(\boldsymbol{x})-\hat{\boldsymbol{u}}_k(\boldsymbol{x})\right\|^2 \quad (8)$$

条件：$\forall k;\boldsymbol{u}_k(\boldsymbol{x})\in U$

式（7）中的问题属于标准的光流问题，当前有多种高效的光流问题求解算法。文中采用文献[14]中的光流算法。该算法利用 Charbonnier 惩罚函数 $\sqrt{x^2+\epsilon^2}$ 来构建一个稳健的空间正规化泛函 $R[\ ]$。式（8）表示的子问题仍然是在子空间秩约束下，使 L_2 重建误差最小化。与 2.1 节类似，通过谱分解可求解该问题。

2.3 求解 $\{(\boldsymbol{A}_k,\boldsymbol{b}_k)\}_{k=1}^n$

对于颜色变换 $\{(\boldsymbol{A}_k,\boldsymbol{b}_k)\}_{k=1}^n$ 而言，其子问题为：

$$\min_{\{\boldsymbol{A}_k,\boldsymbol{b}_k\}_{k=1}^n}\sum_{k=1}^n\int_\chi\left\|\boldsymbol{A}_k\boldsymbol{g}_k(\boldsymbol{u}_k(\boldsymbol{x}))+\boldsymbol{b}_k-\boldsymbol{f}_k(\boldsymbol{x})\right\|^2\mathrm{d}\boldsymbol{x} \quad (9)$$

条件：$\boldsymbol{A}_k\in SO(3)$

该问题具有闭合形式的解，称为 Kabsch 解[15]。将 T_k 定义为：

$$T_k\equiv\int_\chi\left(\boldsymbol{f}_k(\boldsymbol{x})-\bar{\boldsymbol{f}}_k\right)\left(\boldsymbol{g}_k(\boldsymbol{u}_k(\boldsymbol{x}))-\bar{\boldsymbol{g}}_k\right)^{\mathrm{T}}\mathrm{d}\boldsymbol{x} \quad (10)$$

式中：$\bar{\boldsymbol{f}}_k\equiv\frac{1}{|\chi|}\int_\chi\boldsymbol{f}_k(\boldsymbol{x})$ 且 $\bar{\boldsymbol{g}}_k\equiv\frac{1}{|\chi|}\int_\chi\boldsymbol{g}_k(\boldsymbol{u}_k(\boldsymbol{x}))$。

将 $T_k=USV^T$ 定义为 3×3 矩阵 $\boldsymbol{T}_k$ 的奇异值分解。则：

$$\begin{aligned}\boldsymbol{A}_k&=\boldsymbol{U}\boldsymbol{V}^T\\ \boldsymbol{b}_k&=\bar{\boldsymbol{f}}_k-\boldsymbol{A}_k\bar{\boldsymbol{g}}_k\end{aligned} \quad (11)$$

2.4 算法

文中构建了一种图像数据低维表示算法，可以同时实现颜色、外观和形态的同步低维表示，算法伪码为算法 1，计算图像集的低维表示为：

Input: 图像集 $\{\mathbf{f}_1(\mathrm{x}),\ldots,\mathbf{f}_n(\mathbf{x})\}$

$\mathbf{u_k}(\mathbf{x})=\mathbf{x}$ for k=1,…,n

$A_k=I$, $b_k=0$ for k=1,…,n

repeat

for k=1,…,n **do**

$\tilde{f}_k(x)=A_k^{-1}\left(f_k\left(u_k^{-1}(x)\right)-b_k\right)$

$\tilde{f}_k=\mathrm{vec}\left(\tilde{f}_k(x)\right)$

end for

$\bar{f}=\frac{1}{n}\sum_{k=1}^n\tilde{f}_k$

$\left(Ø_1,\ldots,Ø_{d_g}\right)\leftarrow\sum_{k=1}^n(\tilde{f}_k-\bar{f})\left(\tilde{f}_k-\bar{f}\right)^T$ 的前 d_g 个特征向量

for k=1,⋯,n do

$\tilde{g}=\bar{f}+\sum_{p=1}^{d_g}Ø_pØ_p^T\left(\tilde{f}_k-\bar{f}\right)$

$g_k(x)=\mathrm{unvec}(\tilde{g})$

$\hat{u}_k(x)=\mathrm{argmin}_u\int_X\left(\left\|A_kg_k(u_k(x))+b_k-f_k(x)\right\|^2+R(u)\right)dx$

$\hat{u}_k=\mathrm{vec}\left(\hat{u}_k(x)\right)$

end for

$\bar{u}=\frac{1}{n}\sum_{k=1}^n\hat{u}_k$

$\left(\psi_1\cdots\psi_{d_u}\right)\leftarrow\sum_{k=1}^n(\hat{u}-\bar{u})(\hat{u}-\bar{u})^T$ 的前 d_u 个特征向量 ψ

for k=1,⋯,n do

$\tilde{u}=\bar{u}+\sum_{p=1}^{d_u}\psi_p\psi_p^T(\hat{u}-\bar{u})$

$u_k(x)=\mathrm{unvec}(\tilde{u})$

$\bar{f}_k=\frac{1}{|\chi|}\sum_{x\hat{\imath}\chi}f(x)$

$\bar{g}_k=\frac{1}{|\chi|}\sum_{x\hat{\imath}\chi}g(x)$

$T=\sum_{x\hat{\imath}\chi}\left(f_k(x)-\bar{f}_k\right)\left(g_k(x)-\bar{g}_k\right)^T$

$(\mathbf{U},\mathbf{S},\mathbf{V})=\mathrm{svd}(\mathbf{T})$

$\mathbf{A}_k=\mathbf{U}\mathbf{V}^T$

$b_k=\bar{f}_k-A_k\bar{g}_k$

end for

until 收敛

Output:

$\{g_1(\mathbf{x}),\ldots,g_n(\mathbf{x})\}$, $\{u_1(\mathbf{x}),\ldots,u_n(\mathbf{x})\}$, $\{Ø_1(x),\ldots,Ø_{d_g}(x)\}$，及 $\left\{\psi_1(x),\ldots,\psi_{d_u}(x)\right\}$

3 定量评估

根据文中算法估计的形变域 $\{\boldsymbol{u}_k(\boldsymbol{x})\}_{k=1}^n$ 的精度，对文中算法进行定量评估。结合稳健光流算法[14]和 SIFT 流算法[16]来比较这些域的质量。在 4 种对象类别上测试文中算法：马、蘑菇、花和鸟。对于第一种类别，采用由 327 个彩色图像构成的 Weizmann 公开数据集。对于其他各种类别，利用 Google 图像搜索采集了 120 个图像。在所有 4 种集合中，图像边缘基本与对象的定界框重合。每个数据集 F 均有一个真实

图像集 B。即对每个 $\boldsymbol{f}(\boldsymbol{x})\in F$，对应存在一个映射 $b(x)\in B$，其中 $b:\chi\rightarrow\{0,1\}$。

各个算法均生成一个形变域 $\boldsymbol{u}_{jk}(\boldsymbol{x})$，确定从图像 $\boldsymbol{f}_j$ 中位置到图像 $\boldsymbol{f}_k$ 中位置的密集对应关系。对于光流算法和 SIFT 流算法，需要提供($\boldsymbol{f}_j, \boldsymbol{f}_k$)作为输入来获得输出 $\boldsymbol{u}_{jk}(\boldsymbol{x})$，然后对每一对 i,j 运行稳健光流算法和 SIFT 流算法。在评估算法时考虑 2 个性能指标[14,16]：基于区域的性能指标和基于边界的性能指标。每个指标均是基于 2 二元映射的泛函 z，可衡量各对的相似度或差异度，它具有形式 $z(b_j(), b_k())$。因为基于区域的性能指标和基于边界的性能指标均针对图像对，所以对整个数据集的性能定义为所有可能图像对的均值。

$$\bar{z}\equiv\frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n}z(b_j(\boldsymbol{u}_{jk}()),b_k()) \tag{11}$$

对于一对二元映射，基于区域的性能指标定义为：

$$z_R(b_j(),b_k())\equiv\frac{\sum_x b_j(\boldsymbol{x}),\ b_k(\boldsymbol{x})}{\sum_x b_j(\boldsymbol{x})} \tag{12}$$

式（12）从本质上描述了 b_j 中的哪些图像点与 b_k 中的图像点重合。当 $b_j=b_k$ 时，$z_R(b_j(), b_k())=1$，因此，z_R 值较大时表示重叠度较高，则 z_R 值越大越好。

基于边界的性能指标 z_B 可衡量边界位移误差。用 Q_k 表示 $b_k(x)$中的边界坐标集合 x。z_B 定义为：

$$z_B(b_j(),b_k())\equiv\frac{1}{|Q_j|}\sum_{x\in Q_j}\min_{y\in Q_k}\|x-y\|$$

表 2 和表 3 总结了各算法的性能。从 2 个表中可以看到，文中算法在边界位移误差方面的性能优于其他算法，且在基于区域的性能指标方面也取得了优异性能。

表 2　基于区域的性能指标 $\bar{z}_R$（越高越好）

Tab.2 Region-based performance criterion $\bar{z}_R$ (higher is better)

	蘑菇	花	鸟	马
光流	0.76	0.61	0.69	0.62
SIFT 流	0.70	0.62	0.67	0.59
文中方法	0.73	0.68	0.71	0.64

表 3　基于边界的性能指标 $\bar{z}_B$（越小越好）

Tab.3 Boundary–based performance criterion $\bar{z}_B$ (lower is better)

	蘑菇	花	鸟	马
光流	11.54	15.34	14.27	9.62
SIFT 流	13.73	15.39	14.31	9.72
文中方法	5.69	5.65	6.10	4.61

4　定性评估

4.1　图像变形

图像变形的目的是逐渐地将图像 $\boldsymbol{f}_1(\boldsymbol{x})$变换到另一图像 $\boldsymbol{f}_2(\boldsymbol{x})$中，使变换过程看起来非常自然，没有对象被破坏或者显现出其他人为痕迹。高质量变形有两大条件。首先，$\boldsymbol{f}_1(\boldsymbol{x})$像素和 $\boldsymbol{f}_2(\boldsymbol{x})$像素之间必须存在有意义的对应关系。这决定着 $\boldsymbol{f}_1(\boldsymbol{x})$中的每个像素落于 $\boldsymbol{f}_2(\boldsymbol{x})$中的哪个位置。另一个重要考虑是中间图像必须显得自然。这需要已知 $\boldsymbol{f}_1(\boldsymbol{x})$和 $\boldsymbol{f}_2(\boldsymbol{x})$所在图像的空间，以便保持位于该空间内的中间图像。

如果 $\boldsymbol{f}_1(\boldsymbol{x})$和 $\boldsymbol{f}_2(\boldsymbol{x})$不相似，则难以满足上述 2 个条件。文中方法有助于满足这 2 个条件。首先，利用一组图像 F 来获得低维估计的表示，将会把变换空间约束到构建 F 所需要的图像。其次，当 2 个不同图像间的对应关系难以确定或者非常不明确时，$\mathcal{F}$ 中元素的共同关系可以确定 $\boldsymbol{f}_1(\boldsymbol{x})$和 $\boldsymbol{f}_2(\boldsymbol{x})$的关联方式。考虑到 $\boldsymbol{g}_k(\boldsymbol{x})=\boldsymbol{A}_k^{-1}\boldsymbol{f}_k(\boldsymbol{u}_k(\boldsymbol{x}))-\boldsymbol{b}_k$，因此对于任意 $\boldsymbol{x}$，$\boldsymbol{f}_j$ 中的位置 $\boldsymbol{u}_j^{-1}(\boldsymbol{x})$ 对应于 $\boldsymbol{f}_k$ 中的位置 $\boldsymbol{u}_k^{-1}(\boldsymbol{x})$。通过引入范围在 0 到 1 之间的时间参数 t，可以使该对应关系成为渐进式关系：

$$\boldsymbol{f}_{\text{morph}}\left((1-t)\boldsymbol{u}_1^{-1}(\boldsymbol{x})+t\boldsymbol{u}_2^{-1}(\boldsymbol{x})\right)=(1-t)\boldsymbol{f}_1(\boldsymbol{x})+t\boldsymbol{f}_2(\boldsymbol{x}) \tag{13}$$

考虑到 t=0 时，有 $\boldsymbol{f}_{\text{morph}}\left(\boldsymbol{u}_1^{-1}(\boldsymbol{x})\right)=\boldsymbol{f}_1(\boldsymbol{x})$，$t$=1 时有 $\boldsymbol{f}_{\text{morph}}\left(\boldsymbol{u}_2^{-1}(\boldsymbol{x})\right)=\boldsymbol{f}_2(\boldsymbol{x})$。通过将 t 的范围限定为 0 到 1 之间，在实现图像变形的同时，利用子空间受限的学习变形 $\{\boldsymbol{u}_k(\boldsymbol{x})\}_{k=1}^{n}$ 构建出需要的变形。

假设需要将图像 I_1 的形态转移到图像 I_2 上（见图 1a, b）。这可能需要已知 2 个完全不同对象之间的密集对应场，稳健性光流[14]和 SIFT 流[15]等经典方法无法满足这一要求。这 2 种方法均无法获得“蘑菇空间”的外观。这些算法解中的“蘑菇空间”见图 1。相反，文中方法在蘑菇图像的颜色、外观和形态方面以蘑菇图像空间的学习模型为基础，可以实现有效的形态转移。

4.2　关节学

假设在一组图像 F 中有一个对象含有多个关节且以不同姿态呈现。关节学习的目的是提供可以用于

a I_1　b I_2　c 文中算法　d SIFT 流算法　e 光流算法

图 1　采用不同算法时的图像形态转移

Fig.1 Image morphological transition with different algorithms used

操纵该对象的部分参数。证明文中方法可用于实现这一目的且质量较高。使用 Youtube 网上的一个木偶视频作为输入。提取出一段视频，从中采样 100 个帧构成集合 F（见图 2）。

通过将文中算法应用于上述集合，可以获得集合 $\{\boldsymbol{u}_k(\boldsymbol{x})\}_{k=1}^{n}$，逆映射 $\{\boldsymbol{u}_k^{-1}(\boldsymbol{x})\}_{k=1}^{n}$，以及基函数 $\{\psi_q(\boldsymbol{x})\}_{q=1}^{d_u}$。利用学习子空间 U 获得木偶的几何形变，因此，系数 c_q 和各个基函数 $\psi_q(\boldsymbol{x})$ 作为参数，可以控制木偶沿序列内的主要形变方向运动。现在，假设希望控制集合 F 中第 k 个图像的木偶运动，即图 2 框中的木偶。利用式（14）可实现上述目的。

$$\boldsymbol{f}_k^{\dagger}\left(\boldsymbol{x};c_1\ldots c_{d_u}\right)=\boldsymbol{f}_k\left(\boldsymbol{u}_k^{-1}\left(\boldsymbol{u}_k(\boldsymbol{x})+\sum_{q=1}^{d_u}c_q\psi_q(\boldsymbol{x})\right)\right) \quad (14)$$

式中：$\boldsymbol{f}_k^{\dagger}$ 为 $\boldsymbol{f}_k$ 更新后的姿态；c_q 决定了第 q 个基函数的分布。对所有 $q=1\cdots d_u$，原点处 $c_q=0$，且 $\boldsymbol{f}_k^{\dagger}\left(\boldsymbol{x};c_1,\ldots,c_{d_u}\right)=\boldsymbol{f}_k(\boldsymbol{x})$。图 3 为通过这种方法进行合成的图像（图 3a 为原始图像，图 3b, c 为基于 $\psi_1(\boldsymbol{x})$ 的控制，图 3d, e 为基于 $\psi_2(\boldsymbol{x})$ 的控制。图 3f,i 为相应的运动域），可以获得手部的上/下和左/右运动。合成后的图像为原始集合中并不存在的新图像。

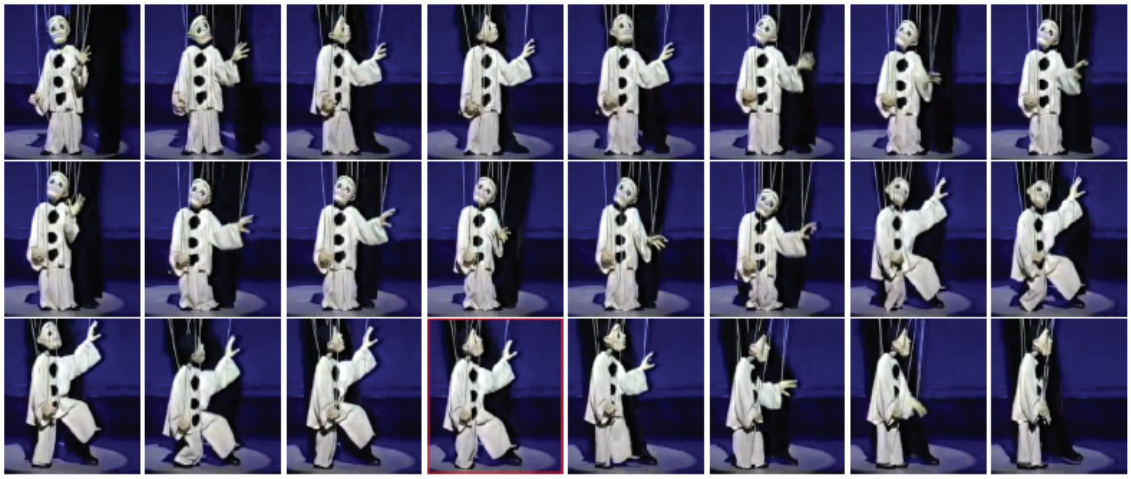

图 2　木偶序列的部分图像

Fig.2 Some images from the puppet sequence

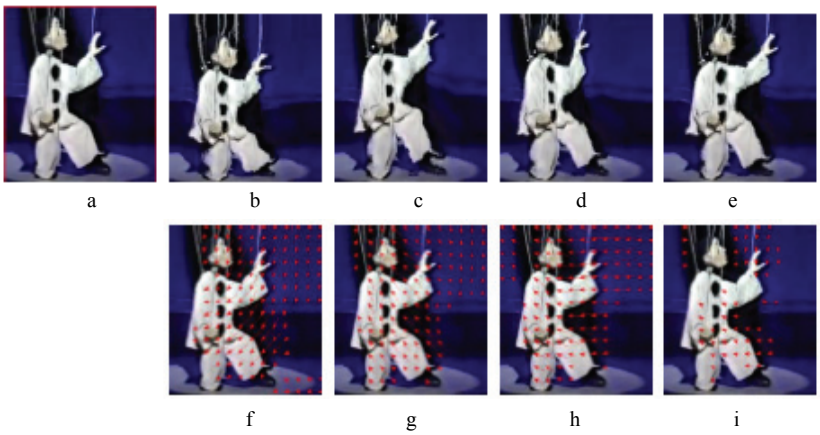

图 3　新合成的姿态

Fig.3 Newly synthesized poses

5　结语

针对现有图像表示方法的不足，文中提出一种可以同时实现图像颜色、外观和形态的图像低维表示算法。该方法通过采取嵌套和组合策略，考虑了颜色、外观和形态方面的变化，其性能要远优于目前典型的稳健型光流算法和 SIFT 流算法。此外，文中方法在图像编辑和图像浏览等相关任务中也取得了令人满意的定性结果。在下一步工作中，将对图像识别中光照，姿态等变化带来的识别困难进行分析，为了提高稀疏表示图像识别的鲁棒性，研究一种基于 Gabor 特征和字典学习的高斯混合稀疏表示图像识别算法。

参考文献：

[1]　王向阳，胡峰丽．一种基于位平面综合特征的彩色

图像检索方案[J]. 计算机研究与发展, 2015, 44(5): 867—872.
WANG Xiang-yang, HU Feng-li. An Efficient Color Image Retrieval Technique Based on Multi-Features of Bit-Plane[J]. Journal of Computer Research and Development, 2015, 44(5): 867—872.

[2] 胡正平, 李静. 基于低秩子空间恢复的联合稀疏表示人脸识别算法[J]. 电子学报, 2013, 41(5): 987—991.
HU Zheng-ping, LI Jing. Face Recognition of Joint Sparse Representation Based on Low-Rank Subspace Recovery[J]. Acta Electronica Sinica, 2013, 41(5): 987—991.

[3] WANG J, LIU Z, WU Y, et al. Learning Actionlet Ensemble for 3D Human Action Recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36(5): 914—927.

[4] DAI Y, LI H, HE M. A Simple Prior-Free Method for Non-rigid Structure-from-motion Factorization[J]. International Journal of Computer Vision, 2014, 107(2): 101—122.

[5] DING C, XU C, TAO D. Multi-task Pose-Invariant Face Recognition[J]. IEEE Transactions on Image Processing, 2015, 24(3): 980—993.

[6] BRIVIO P, TARINI M, CIGNONI P. Browsing Large Image Datasets Through Voronoi Diagrams[J]. IEEE Transactions on Visualization and Computer Graphics, 2015, 16(6): 1261—1270.

[7] OLSHAUSEN B A. Emergence of Simple-cell Receptive Field Properties by Learning a Sparse Code for Natural Images[J]. Nature, 2013, 381(6583): 607—609.

[8] PENG Y, GANESH A, WRIGHT J, et al. RASL: Robust Alignment by Sparse and Low-rank Decomposition for Linearly Correlated Images[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(11): 2233—2246.

[9] ZHANG Z, WANG J, ZHA H. Adaptive Manifold Learning[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(2): 253—265.

[10] YE D H, DESJARDINS B, HAMM J, et al. Regional Manifold Learning for Disease Classification[J]. IEEE Transactions on Medical Imaging, 2014, 33(6): 1236—1247.

[11] CHEN X, UDUPA J K, BAGCI U, et al. Medical Image Segmentation by Combining Graph Cuts and Oriented Active Appearance Models[J]. IEEE Transactions on Image Processing, 2012, 21(4): 2035—2046.

[12] JONES M J, POGGIO T. Multidimensional Morphable Models: A Framework for Representing and Matching Object Classes[J]. International Journal of Computer Vision, 2014, 29(2): 107—131.

[13] SEILER C, PENNEC X, REYES M. Capturing the Multiscale Anatomical Shape Variability with Polyaffine Transformation Trees[J]. Medical Image Analysis, 2012, 16(7): 1371—1384.

[14] SUN D, ROTH S, BLACK M J. Secrets of Optical Flow Estimation and Their Principles[C]// 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) San Francisco, CA, USA IEEE Press, 2010: 2432—2439.

[15] LIU C, YUEN J, TORRALBA A, et al. Sift Flow: Dense Correspondence Across Different Scenes[C]// IEEE 12th European Conference on Computer Vision (ECCV) Florence, Italy IEEE Press, 2008: 28—42.

[16] KABSCH W. A Solution for the Best Rotation to Relate Two Sets of Vectors[J]. Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography, 2006, 32(5): 922—923.