

# 基于深度信息的压缩感知人脸检测跟踪<sup>①</sup>

郑宪成<sup>1</sup>, 蒋敏<sup>1</sup>, 孔军<sup>1,2</sup>, 吴佼<sup>1</sup>

<sup>1</sup>(江南大学 轻工过程先进控制教育部重点实验室, 无锡 214122)

<sup>2</sup>(新疆大学 电气工程学院, 乌鲁木齐 830047)

**摘要:** 传统的压缩感知跟踪是基于彩色视频图像序列中的目标跟踪, 但在跟踪过程中可能会受到光照变化和旋转遮挡因素的影响, 从而导致复杂环境下跟踪结果的鲁棒性不足. 为了获得稳定的跟踪结果, 提出了一种基于深度信息的压缩感知人脸检测跟踪算法. 首先, 根据改进的质心分割算法确定首帧深度图中人脸的跟踪位置. 其次, 根据深度信息计算出深度图中每一点对应的平均曲率并形成平均曲率图. 然后, 基于平均曲率图随机提取压缩特征; 最后, 通过压缩降维, 目标邻域搜索, 迭代更新特征模板, 计算出平均曲率图中下一帧人脸的最优跟踪位置, 实现人脸跟踪. 实验结果表明, 将人脸的深度信息和压缩感知特征相结合在光照变化和旋转遮挡情况下具有很好的鲁棒性, 能更好的实现复杂背景下对多姿态人脸的跟踪.

**关键词:** 压缩感知; 平均曲率图; 深度信息; 人脸跟踪

## Depth Based Face Detection and Tracking Using Compressive Sensing

HENG Xian-Cheng<sup>1</sup>, JIANG Min<sup>1</sup>, KONG Jun<sup>1,2</sup>, WU Jiao<sup>1</sup>

<sup>1</sup>(Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), Jiangnan University, Wuxi 214122, China)

<sup>2</sup>(College of Electrical Engineering, Xinjiang University, Urumqi 830047, China)

**Abstract:** The traditional target tracking algorithms based on compressive sensing uses color video sequence. It is easily affected by rotation, occlusion and illumination. Thus it has poor robustness in complex environment. In order to obtain stable tracking results, this paper proposed a novel face detection and tracking algorithm using compressive sensing based on depth information. Firstly, we identify the position of face in the initial depth map automatically using the improved centroids segmentation algorithm. Secondly, we calculate the mean curvature graph with the depth map. Thirdly, the compression features are extracted based on the mean curvature graph. Finally, by compression and dimension reduction, the feature template is updated and the best position of face is estimated in the neighborhood space based on the mean curvature graph. The experiment results demonstrate that by combining the depth information of human face and the compressive sensing feature, the proposed algorithm has better robustness dealing with problems of rotation, illumination and occlusion.

**Key words:** compressive sensing; mean curvature graph; depth information; face tracking

目标跟踪在运动监督、目标分类、识别以及医疗图像等方面具有很广泛的应用. 传统 RGB 视觉传感器主要捕获色彩信息, 很难在光照剧烈变化和旋转遮挡发生时对目标实现准确的跟踪. 随着深度数据提取技术和深度摄像机的不断成熟与完善, 深度图像的重要性愈发凸显. 近年来许多学者将三维空间信息和深度

相机结合来解决一些复杂的问题. 例如, Gordon 等<sup>[1]</sup>使用立体照相机重构三维空间的信息来实现人体跟踪. 文献[2]使用 Kinect 获得深度图像并将前景附近区域与背景进行分割, 从而减小背景对目标的干扰来有效地实现目标的跟踪. Tang 等<sup>[3]</sup>提出了一种基于深度相机的联合的稀疏外观的模型的跟踪.

① 基金项目: 新疆维吾尔自治区自然科学基金(201233146-6)

收稿时间: 2015-01-12; 收到修改稿时间: 2015-03-18

Donoho 和 Candes 于 2006 年提出了压缩感知理论 (CS, Compressed Sensing)<sup>[4]</sup>, 它基于原信号稀疏特性和远小于 Nyquist 频率的随机采样获取原信号离散样本, 然后通过非线性重构算法构造出具有原信号特性的降维后的特征信号. 随着该理论在工程应用和数学等方面的广泛应用, 产生了 1-BIT CS<sup>[5]</sup>理论、Bayesian CS<sup>[6]</sup>理论、分布式 CS<sup>[7]</sup>理论等多个分支. 2012 年前后, 压缩感知算法被应用到跟踪领域<sup>[8,9]</sup>. 比较有代表性的两篇文章是 Zhang<sup>[9]</sup>等人把压缩感知理论拓展到检测的目标跟踪方面, 大大的提高了跟踪效果的实时性, 但缺点是目标描述单一, 在光照发生剧烈变化时跟踪目标将会发生脱靶; Li<sup>[8]</sup>等人基于压缩感知理论拓展 L1<sup>[10]</sup>跟踪算法, 使得跟踪速率在 L1 的基础上得到了大大的提高, 缺点是当遮挡与旋转发生时跟踪效果比较差.

针对文献[8-9]中传统压缩跟踪算法对彩色视频序列跟踪的不足, 本文对压缩感知跟踪算法提出了改进, 针对跟踪过程中复杂情况, 如光照发生剧烈变化和旋转遮挡等干扰, 提出了基于深度信息的人脸跟踪算法, 使得跟踪的效果更加稳定. 实验证明, 本文中提出的算法能够在光照发生剧烈变化和旋转遮挡发生情况下精确地实现人脸的跟踪, 算法的鲁棒性显著提升.

### 1 压缩感知跟踪算法

设为  $t$  帧的跟踪人脸位置. 在以  $O$  为圆心, 半径为  $\alpha$  的圆内采集正样本; 以  $\alpha$  ( $\gamma > \alpha$ ) 为内半径,  $\beta$  ( $\beta > \gamma$ ) 为外半径采集负样本.

每一个样本  $Z \in R^{w \times h}$  ( $R$  为实数集,  $w$  和  $h$  分别为样本  $z$  的宽度和长度), 为了解决幅度的问题, 使用多尺度滤波器  $\{k_{1,1}, k_{1,2}, \dots, k_{w,h}\}$  与  $z$  进行卷积, 其中

$$k_{i,j}(x,y) = \begin{cases} 1, & 1 \leq x \leq i, 1 \leq y \leq j \\ 0, & \text{其它} \end{cases} \quad (1)$$

其中  $i$  和  $j$  分别代表矩形滤波器的宽度和长度. 从而, 每一个尺度图像可以用  $R^{w \times h}$  中的任意一个列向量表示, 把这些列向量串联起来可以得到一个多尺度图像  $x = (x_1, \dots, x_m)^T \in R^m$  向量, 其中  $m = (wh)^2$ .  $m$  一般从  $10^6$  到  $10^{10}$ , 维数非常高. 为了有效的降低维数, Zhang<sup>[9]</sup>等人借助一个随机投影矩阵进行降维处理, 其中中的每一个元素为

$$r_{i,j} = \sqrt{\rho} \times \begin{cases} 1 & \text{概率为 } \frac{1}{2\rho} \\ 0 & \text{概率为 } 1 - \frac{1}{\rho} \\ -1 & \text{概率为 } \frac{1}{2\rho} \end{cases} \quad (2)$$

其中通过平均概率在 2~4 中随机选取.

在实际应用中为了降低特征的冗余性, 压缩感知算法是对多尺度向量进行压缩投影操作来得到低维向量, 具体压缩过程如公式(3).

$$v = Rx \quad (3)$$

其中  $v \in R^{n \times 1}$  为压缩后的低维向量,  $x \in R^{m \times 1}$  ( $n \ll m$ ) 为压缩前的原始信号. 根据压缩感知理论可知, 从低维向量  $v$  中我们能以很高的概率得到多尺度向量  $x$ .

其后的目标跟踪就是提取样本压缩后的特征, 将压缩后的正负样本特征输入朴素贝叶斯分类器训练, 得到最大的似然度即认为是目标区域, 从而确定下一帧中目标的跟踪位置, 见公式(4).

$$K(v) = \log \frac{\prod_{i=1}^k p(v_i|y=1)p(y=1)}{\prod_{i=1}^k p(v_i|y=0)p(y=0)} = \sum_{i=1}^k \log \left( \frac{p(v_i|y=1)}{p(v_i|y=0)} \right) \quad (4)$$

采样过程中, 将趋于目标的样本视为正样本, 远离目标的样本视为负样本. 这里的  $y$  为随机变量且满足,  $p(v_i|y=1) = p(v_i|y=0) = 0.5$ , 假定条件分布  $p(v_i|y=1)$  与  $p(v_i|y=0)$  满足高斯分布, 见公式(5), 其中  $\mu, \sigma$  分别为样本的均值和标准差.

$$\begin{cases} p(v_i|y=1) \sim N(\mu_i^1, \sigma_i^1) \\ p(v_i|y=0) \sim N(\mu_i^0, \sigma_i^0) \end{cases} \quad (5)$$

最后选取  $K(v)$  值最大样本位置作为下一帧的目标位置, 并对相关参数进行更新以适应目标与背景的更新, 目标参数的更新公式为:

$$\begin{cases} \mu_i^1 = \lambda \mu_i^1 + (1-\lambda) \mu^1 \\ \sigma_i^1 = \sqrt{\lambda(\sigma_i^1)^2 + (1-\lambda)(\mu_i^1 - \mu^1)^2} \end{cases} \quad (6)$$

公式中表示对参数的更新速度且与参数更新的速度成负相关性, 本文中选取  $\lambda = 0.8$ .

## 2 面向深度图的压缩感知人脸定位跟踪算法

### 2.1 首帧图像中的人脸定位

为了确定首帧图像中人脸的位置,本文采用了颜色质心分割算法(CCS)<sup>[11]</sup>对首帧彩色图像中的人脸进行检测定位,将待选跟踪区域投影得到对应深度图像中的位置.进一步,分析各待选区域深度信息,基于人脸深度正样本确定人脸准确位置.具体过程如下:

第 1 步:建立 CCS 模型并基于其确定质心分割的阈值.

第 2 步:将 RGB 颜色空间中 RGB 成份转换到 2-D 极坐标系中,然后用第 1 步中得到的阈值分割质心区域.

第 3 步:根据得到的质心区域确定首帧彩色图像中的人脸的待选位置.

第 4 步:依据同一帧彩色与深度图像中各个点对应的关系,将待选区域投影到深度图像中.

第 5 步:依据人脸表面的平滑性与连续性特性,对待选区域进行二次分割.

第 6 步:比对人脸曲率正样本,基于最大似然度最终确定人脸跟踪区.如图 1 所示:

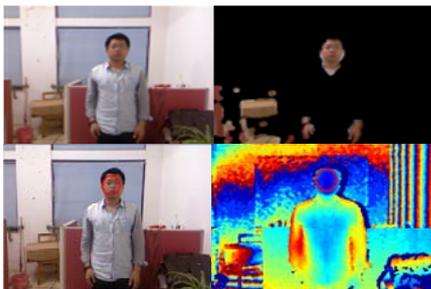


图 1 首帧人脸的定位过程

### 2.2 基于平均曲率图构造特征

假设深度图曲面  $f$  上面的任意一点坐标为  $(x, y)$ , 下面我们定义两个基本的矩阵形式

$$g = \begin{pmatrix} \left\| \frac{\partial f}{\partial x} \right\|^2 & \frac{\partial f}{\partial x} \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial x} \frac{\partial f}{\partial y} & \left\| \frac{\partial f}{\partial y} \right\|^2 \end{pmatrix} \quad h = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} \vec{n} & \frac{\partial^2 f}{\partial x \partial y} \vec{n} \\ \frac{\partial^2 f}{\partial x \partial y} \vec{n} & \frac{\partial^2 f}{\partial y^2} \vec{n} \end{pmatrix} \quad (7)$$

其中的  $\vec{n}$  为常规单位向量定义如下:

$$\vec{n} = \frac{\frac{\partial f}{\partial x} \times \frac{\partial f}{\partial y}}{\left\| \frac{\partial f}{\partial x} \times \frac{\partial f}{\partial y} \right\|} \quad (8)$$

从文献[12]中我们可以知道平均曲率与矩阵对角线的一半相等,即:

$$H = \frac{\text{tr}(hg^{-1})}{2} \quad (9)$$

依次计算深度图中的每个点将会得到它对应的平均曲率图.

本文算法的特征提取是面向平均曲率图实现的,为了更好的增强特征值的稳健性,对背景与目标进行了严格的分类.本文的特征提取的公式如下.

$$v_i = \sum_{n=1}^N \frac{p_{in} \text{Recls}_{in}}{\sqrt{N}} \quad (10)$$

其中  $\text{Recls}$  为目标候选区域随机选取的图像块,  $N$  为样本数目,此时得到的压缩特征  $v_i$  是由原始特征以  $p_{in}$  为权值的加权和.

### 2.3 跟踪算法流程

本文对压缩感知算法在彩色视频序列中的应用进行了改进,算法的流程如下:

第 1 步:针对首帧,基于深度与肤色信息,自动定位出目标的跟踪窗,同时计算出首帧的平均曲率图,并将跟踪窗投影到深度图上面;

第 2 步:面向深度图像,采集目标样本和背景样本,然后根据(2)式和(10)计算出训练样本的特征;

第 3 步:读取下一帧,计算出对应的平均曲率图,基于(2)式和(10)式抽取样本特征,代入(4)式,基于贝叶斯分类器,选取  $K_{\max(v)}$  所对应的位置为下一帧的目标的位置  $pos$ ;

第 4 步:分析位置  $pos$  原始深度信息,根据人脸表面的连续性与平滑性特性,进一步修正得到目标位置  $pos$ ;

第 5 步:如果图像序列未结束,转到第 3 步.

## 3 试验结果与分析

为验证本文算法的有效性,在真实场景中对移动的人脸目标进行跟踪.本文使用微软的 Kinect 作为测试用的深度和彩色摄像机,其深度图像分辨率为  $640 \times 480$  像素,帧率 30/s.最后将本文算法在深度序列上的跟踪效果与压缩感知算法在彩色序列上面的跟踪效果进行比较(两种算法是在同一时间同一帧上进行的).实验环境 intel (R) Core(TM)i5 CPU 760

@2.80GHz、4G 内存、操作系统 XP, 开发平台 VS2010, 结合 Matlab 库函数. 实验结果以定性与定量的跟踪效果图和定量的误差曲线图进行说明.



图 2 压缩感知跟踪算法对光照剧烈变化下彩色序列的跟踪效果

### 3.1 在光照剧烈变化情况下的实验对比

图 2 与图 3 分别是压缩感知算法对光照变化下序列跟踪效果, 图片大小为 640 像素 480 像素, 红色椭圆框为目标跟踪区域.

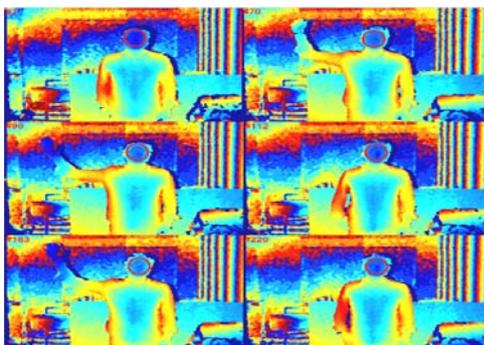


图 3 本文跟踪算法对光照剧烈变化下深度序列的跟踪效果

从图 2 中可以看出当实验中的光源剧烈变化时, 压缩感知算法跟踪效果不稳定. 例如#70 相对于#1 来说, 跟踪窗发生了明显的偏移, 甚至从#90 开始出现了脱靶的现象. 而本文提出的算法(见图 3)丝毫没受到光照影响.

### 3.2 在旋转与部分遮挡情况下的实验的对比

为了进一步验证本文算法的有效性, 利用压缩感知跟踪算法和本文算法对旋转与部分遮挡下的序列进行跟踪, 图片大小为 640 像素 480 像素, 实验者在整个过程中进行头部旋转与摘戴眼镜的操作, 将图 4 与图 5 对比我们发现当头部偏转与遮挡发生时压缩感知跟踪算

法对人脸的跟踪发生了偏移, 而本文的算法依然能准确的对人脸进行跟踪.



图 4 压缩感知跟踪算法对旋转与部分遮挡下彩色序列的跟踪效果

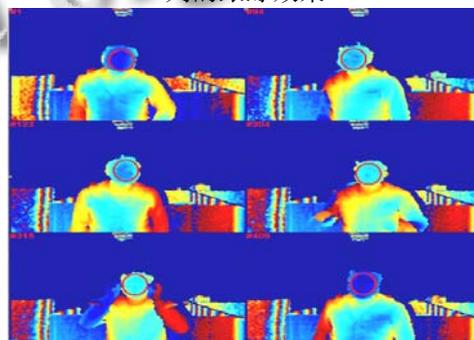


图 5 本文跟踪算法对旋转与部分遮挡下深度序列的跟踪效果

### 3.3 两种算法的误差分析

下面定量分析压缩感知跟踪算法和本文跟踪算法的跟踪精度对比, 图 6 与图 7 是两种算法针对光照和旋转部分遮挡序列的跟踪误差曲线图. 横轴表示帧号, 纵轴表示“跟踪窗左上角”与“真实目标左上角”偏离误差, 单位为像素, 按照 X 和 Y 方向分别描述. 纵坐标为零的点表示对应帧真实目标左上角位置, 上方曲线表示正误差, 下方曲线表示负误差. 红色线表示本文算法的误差曲线, 绿色线表示压缩感知算法误差曲线.

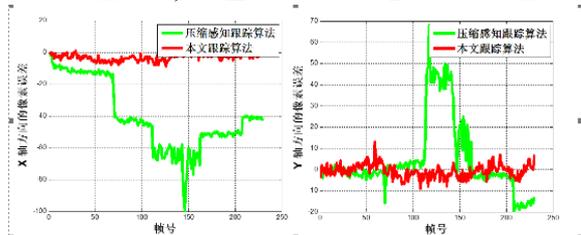


图 6 光照变化序列的跟踪误差曲线图(从左到右依次为 X 方向, Y 方向)

从图 6 可以看出压缩感知跟踪算法在上面 2 种情

况下面的误差曲线从首帧开始就较大幅度的偏离原点位置,尤其是当光照发生剧烈变化时发生了完全脱靶的现象.

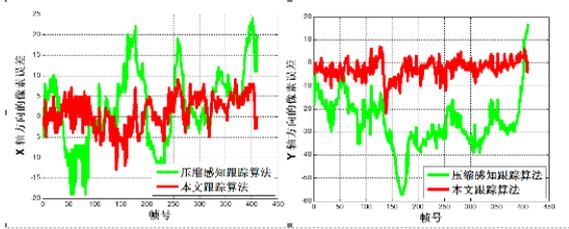


图 7 旋转与遮挡序列的跟踪误差曲线图(从左到右依次为 X 方向, Y 方向)

从图 7 中可以得到当旋转与遮挡发生时候,压缩感知跟踪算法的跟踪结果发生了大幅度的偏移,而本文的算法几乎能完美的进行了跟踪.

为了进一步的评价本文算法的跟踪的性能,与文献[7]类似,采用跟踪成功率  $p$  来衡量人脸跟踪的成功率. 首先,计算每一帧跟踪位置与真实位置的重合率

$$score = \frac{area(B_T \cap B_G)}{area(B_T \cup B_G)} \quad (11)$$

其中,  $B_G$  表示算法跟踪得到的人脸位置的椭圆形框;  $B_T$  表示真实的人脸所在位置的椭圆形框;  $area$  跟踪人脸的面积大小. 式中的  $score$  大小表示跟踪人脸位置的准确度的高低,如果  $score$  大于 0.5 则认为对该帧跟踪成功,反之跟踪失败. 最后根据跟踪成功的帧数除以总的帧数计算出跟踪的成功率. 对比结果见表 1.

表 1 两种算法的跟踪成功率对比(%)

| 序号    | 压缩感知算法 | 本文算法 |
|-------|--------|------|
| 光照变化  | 31     | 100  |
| 旋转遮挡  | 86     | 96   |
| 平均成功率 | 58.5   | 98   |

#### 4 结语

本文基于传统的压缩感知算法,针对其在彩色序列中跟踪的缺陷提出了改进,形成了一种面向平均曲率图的压缩感知目标跟踪算法. 该算法的特征的构造仅仅与目标与摄像头之间的空间相对位置有关,可以

有效的避免光照变化和旋转遮挡因素的影响,基于压缩降维,快速而精准地实现人脸跟踪.

#### 参考文献

- Gordon G, Chen XR, Buck R. Person and gesture tracking with smart stereo cameras. Proc. of the 2008 SPIE. 2008.
- 王莹. 基于 Kinect 的 Tri-tracking 视频跟踪算法研究[硕士学位论文]. 大连: 大连理工大学, 2012.
- Tang F, Harville M, Hai T, et al. Fusion of local appearance with stereo depth for object tracking. IEEE Conference on Computer Society. 2008, 6. 23–28.
- Donoho DL. Compressed sensing. IEEE Trans. on Information Theory, 2006, 52(4): 1289–1306.
- Hong J, Sung K, Nayak KS, et al. k-t FOCUSS: a general compressed sensing framework for high resolution dynamic MRI. Magnetic Resonance in Medicine 2009, 61(1):103–116.
- Ji SH, Xue Y, Carin L. Bayesian Compressive Sensing. IEEE Trans. on Signal Processing, 2008, 56(6): 2346–2356.
- Do TT, Chen Y, Nguyen DT, et al. Distributed compressed video sensing. Proc. of the 16th IEEE International Conference on Image Processing. 2009:1393–1396.
- Li HX, Shen CH, Shi QF. Real-time visual tracking using compressed sensing. 2011 IEEE Conference Computer Vision and Pattern Recognition. 2011. 1305–1312.
- Zhang KH, Zhang L, Yang MH. Real-Time Compressive Tracking. Proc. of the 12th European Conference on Computer Vision. 2012. 864–877.
- Mei X, Ling HB. Robust visual tracking using L1 minimization. Proc. IEEE Int. Conf. Comp. Vis., Kyoto, Japan. 2009. 1436–1443.
- Zhang QS, Kamata S, Zhang J. Face detection and tracking in color images using color centroids segmentation. IEEE International Conference on Robotics and Biomimetics(ROBIO 2008). 2009. 1008–1013.
- Houhou N, Thiran JP, Bresson X. Fast texture segmentation model based on the shape operator and active contour. IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2008). 2008, 6. 23–28.