

行人重识别

中山大学 郑伟诗 吴岸聪

一、引言

随着现代社会的快速发展，城市人口密度越来越高。在一些公共场所，密集的人群容易导致公共安全事件的发生。为了及时预防和处理此类事件，大量的监控摄像头被安装和应用在公共场所。然而面对复杂的监控网络和海量的监控数据，依靠人工难以快速分析和处理，所以利用计算机视觉和机器学习技术辅助甚至取代视频监控中人工参与的需求应运而生。当前针对视频监控的研究受到了大量学者的关注，目的是为了减少人工的干预，实现自动分析和解读多摄像头监控系统提供的信息，并据此发展出针对不同应用需求的计算模型[1]-[3]。

行人重识别 (Person Re-identification) 作为视频监控研究领域的关键组成部分，近几年逐渐成为研究热点，受到广泛的关注[4]-[5]。其目的是对出现在监控摄像头视域内的某个目标行人，准确快速地在监控网络其他摄像头视域内的大量行人中将这个目标行人标识出来。图 1 是对行人重识别的直观展示。一个摄像头中拍摄的目标行人的图像会与其他摄像头拍摄的所有行人图像一一进行匹配，从中寻找出也属于这个目标行人的图像，从而实现目标行人的跨摄像头的搜索追踪和定位。行人重识别技术的应用可以大大减少视频监控中的人工参与，实现对监控视频中的行人及其行为的快速而准确的分析，对预防犯罪的发生和维持良好的社会治安有积极的促进作用。

二、行人重识别研究难点

经过近十年的发展，大量的行人重识别模型相继被提出，但是实际的视频监控环境往往是非可控的，导致行人重识别依然存在一些问题没有解决。具体地，影响行人重识别的主要难

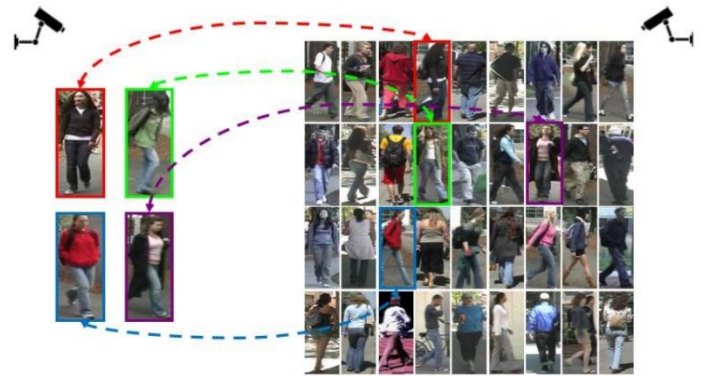


图 1 行人重识别示意图

点有以下几点：

(1) 低分辨率问题：同一行人与不同摄像头之间的距离有远有近，导致不同摄像头拍摄的行人图像的分辨率差异明显；

(2) 遮挡问题：同一行人在不同摄像头拍摄的场景中受背景和其他行人的遮挡程度不同，大量的行人遮挡问题导致了行人图像经常是不完整的；

(3) 视角和姿势问题：不同摄像头的拍摄角度差异会令同一行人在不同摄像头中的图像出现不同的姿势；

(4) 光照变化问题：不同摄像头的光照差异会使同一行人在不同摄像头中的图像表现差异明显；

(5) 视觉模糊性问题：不同行人图像之间的表现差异也可能很小，比如不同行人的衣着相似。

如图 2 所示，以上这些非可控因素的存在会导致同一行人的图像之间的差异变大而不同行人的图像之间的差异变小，所以现阶段的行人重识别依然存在较多难点，离实际应用还有不小的距离。

三、行人重识别的研究现状

行人重识别的一般技术流程是：首先根据行人特征表达方法获取行人图像的特征；然后利用相应的行人相似性判别模型对大量的行人图像进行训练，得到合适的衡量行人图像之间相似性/距



图2 非可控环境下的行人重识别存在的难点

离的判别方法；最后对一个摄像头拍摄的某个目标行人的图像，将其与其他摄像头拍摄的大量行人图像进行匹配，找到与其相似性最高或距离最小的行人图像，从而实现目标行人的再标识。

综上所述，行人重识别的关键在于行人特征表达和行人相似性判别两个步骤。下面对行人重识别领域的文献在特征表达和相似性判别这两个方面的研究进行阐述，此外，也对当前一些行人重识别中特定问题的研究进行介绍。

3.1 行人特征表达

在行人重识别中，行人特征表达的优劣是行人能否被正确再标识的关键之一。现有的行人特征表达方法研究主要是集中在行人表观信息的特征提取和表达上。

在早期的研究中，通常使用手工设计的特征来描述行人的表观信息。常用的行人表观特征包括以颜色直方图[6]为代表的颜色特征和以局部二值模式 (Local Binary Pattern, LBP) [7]和 Gabor 滤波直方图[8]为代表的纹理特征。在颜色和纹理特征的基础上，基于集成特征学习的方法[5]被提出，成为了经典的行人特征表达方法。国内外的研究者们还相继提出了空间表观信息模型[9]、多特征组合法[10]、基于费舍尔向量的组合模型[11]、跨视域不变颜色特征方法[12]、基于协方差特征的表达模型[13]、基于局部对齐的特征变换方法[14]、基于属性的表达模型[15]、自定义图形结构[59]、基于姿势先验和类判别的特征方法[16]、基于局部最大出现频率的表达模型[17]、基于协方差的分层高斯描述子[66]等行

人特征表达方法。关于中高层特征的学习和设计也被应用到行人重识别中，包括有显著性学习[18]、镜像特征表达[60]、显著颜色名称[61]、属性学习[62][63]等等。

最近，深度学习技术也被用来提取行人图像的抽象特征。其中具有代表性的有：深度配对滤波神经网络[19]、深度表观增强特征方法[20]、改进的行人深度特征学习模型[21]和基于协同训练的神经网络[64]等。近期，深度学习模型对于行人特征表达学习有更细致的设计，考虑使用人的不同身体部位信息，引导网络学习更好的特征[67][68][69]。相比起早期使用的手工特征，深度学习模型得到的特征一般比手工设计的特征具有更好的判别性，可以取得更好的行人重识别效果。

此外，特征的提取不只限于基于 RGB 图像，多模态学习的研究也逐渐开始发展。一些学者通过 Kinect 采集行人的深度信息用于完善行人的特征表达，包括基于 Kinect 骨架特征的匹配方法[22]、基于点云匹配的再标识模型[23]、基于深度信息分割增强表观信息的匹配方法[24]、基于颜色和深度信息的三维描述子[25]、基于注意力模型的深度特征学习模型[70]、基于深度信息协方差描述子的方法[71]、以及还有 RGB-D 多模态联合学习的深度学习模型[72]等方法。

3.2 行人相似性判别

行人相似性判别的主要思想是对行人训练样本进行相似性/距离判别学习，即从行人样本中学习最优的相似性/距离的计算方法来匹配相同的

行人。在最初的行人相似判别研究中，无监督学习的 L1 范数距离和 L2 范数距离被首先使用。针对简单的无监督距离鲁棒性弱的问题，有监督的行人相似性判别学习方法逐渐被提出。Prosser 等人利用排序学习思想提出了基于支持向量排序模型的匹配方法[26]，将行人相似性计算问题转化为行人相对关系排序问题，改善了行人重识别的效果。Zheng 等人发展了相对距离比较学习算法[27]，提出利用基于距离比较的优化准则弱化最大化不同行人间距的约束，避免了行人重识别模型的过学习问题，显著地提高了行人重识别的准确率。Kostinger 等人提出了基于等价约束的距离度量模型[28]，后续 Tao 等人还在其基础上进行了扩展和优化[29][30]。Pedagadi 等人还提出利用局部线性判别分析[31]来衡量行人之间的相似性，该方法通过寻找一个最优的低维子空间，使得在该空间内同一行人的距离最小而不同行人的距离最大。除了上述方法外，国内外的研究者们还提出了很多其他的行人相似性/距离判别方法，包括：基于稀疏成对约束的度量学习模型[32]、基于聚类采样的组合模板匹配模型[33]、基于自适应决策函数的行人重识别模型[34]、多任务度量学习模型[35]、基于迭代加权稀疏排序模型的行人重识别模型[36]、基于显著性的相似性判别方法[37]-[39]、基于半监督字典学习的匹配方法[40]、基于核方法的度量学习模型[41]、深度距离度量学习模型[42]、跨视角二次判别分析方法[17]、基于判别视角不变字典的学习方法[43]、基于 PSD 约束的非对称度量学习方法[44]、基于对应结构学习的行人重识别模型[45]、基于上下文判别信息分析的排序学习方法[46]、基于跨视角特征映射的非对称度量学习模型[47]和深度四元组度量学习[73]等。

以上讨论的是有监督学习的模型。实际监控中获取的大量无类标数据同样具有价值，可以加以利用。最近基于无监督学习的相似性度量模型也开始逐渐受到关注，包括有基于 L1 图的无监督学习[65]、跨数据库的无监督迁移学习[74]、非对称无监督度量学习[80]等。

此外，行人重识别也从仅仅依赖于孤立图片扩展视频，因此最近一系列的基于视频的行人重识别方法也被提出，包括 Top-push 度量学习的方法[83]、基于视频帧选择与视频排序的模型[49][50]、基于循环卷积神经网络的模型[75]、基于视频内与视频间联合度量学习的方法[86]、基于时空信息协同学习的递归神经网络模型[87]、基于双向时间信息学习的模型[88]、基于双流孪生卷积神经网络的模型[89]。

3.3 特定行人重识别问题

最近部分国内外学者针对行人重识别中的特定问题进行研究，提出了较多的解决方法。比如，部分学者提出了跨场景迁移学习模型[48]以解决行人训练样本不足的问题；将单个行人扩展到人群提出了基于群组信息的表达模型[51]。此外，针对行人确认[52][53]、行人重识别的迁移学习[54]-[57]、大规模行人重识别[58]、快速检索[76][77]、自然语言检索[78]、全图行人搜索[79]、低分辨率行人重识别[81][85][90]、残缺行人重识别[84]、可见光-红外行人重识别[82]的研究也开始出现。

四、行人重识别研究总结

行人重识别以实现于行人的跨摄像头搜索追踪和定位为目标，是视频监控研究领域的关键技术，近几年逐渐成为研究热点。然而，行人重识别是个具有较大挑战性的问题，难点主要包括低分辨率问题、遮挡问题、视角姿势变化问题、光照变化问题和视觉模糊性问题。针对以上难点，近年来开展了大量行人重识别领域的研究并有很大的进展。现有的研究可分为行人特征表达、行人相似性判别和特定行人重识别问题三类。在特征表达方面，早期使用手工设计的特征，后来特征学习技术以及深度学习技术的应用使得特征具有更好的判别性。在行人相似性判别方面，早期使用子空间和度量学习的方法，近来基于深度学习的度量学习得到发展。另外，从行人重识别的应用角度出发，一些特定问题也开始被研究，比

如数据不足问题、大规模快速检索问题、复杂环境中的非可控因素问题、多模态问题等等。目前行人重识别技术在几个大规模标准数据集上已取得阶段性成果，但仍存在一些在实际应用中的重

要问题未能解决。未来行人重识别技术将会面向实际应用的方向继续发展。

（责任编辑：任传贤）

参考文献

- [1] R. T. Collins, A. J. Lipton, and T. Kanade, Introduction to the special section on video surveillance, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22 (8), 2000, 745-746
- [2] T. Xiang and S. Gong, Video behavior profiling for anomaly detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(5), 2008, 893-908
- [3] 王亮, 胡卫明, 谭铁牛, 人运动的视觉分析综述, *计算机学报*, 25 (3), 2002, 225-237
- [4] N. Gheissari, T. Sebastian, and R. Hartle, Person reidentification using spatio temporal appearance, *CVPR 2006*
- [5] Gray and H. Tao, Viewpoint invariant pedestrian recognition with an ensemble of localized features, *ECCV 2008*
- [6] M. J. Swain, D. H. Ballard. Color indexing, *International Journal of Computer Vision*, 7(1), 1991, 11-32[7]
- [7] T. Ojala, M. Pietikäinen, T. Mäenpää, Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24 (7), 2002, 971-987
- [8] Fogel, D. Sagi. Gabor filter as texture discriminator, *Biological Cybernetics*, 61 (2), 1989, 102-113
- [9] X. Wang, G. Doretto, T. Sebastian, J. Rittscher, and P. Tu, Shape and appearance context modeling, *CVPR 2007*
- [10] M. Farenzena, L. Bazzani, A. Perina, M. Cristani, and V. Murino, Person re-identification by symmetry-driven accumulation of local features, *CVPR 2010*
- [11] B. Ma, Y. Su, F. Jurie, Local descriptors encoded by fisher vectors for person re-identification, *ECCV Workshops 2012*
- [12] Kviatkovsky, A. Adam, and E. Rivlin, Color Invariants for Person Reidentification, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35 (7), 2013, 1622-1634
- [13] B. Ma, Y. Su, and F. Jurie, Covariance Descriptor based on Bio-inspired Features for Person Re-identification and Face Verification, *Image and Vision Computing*, 32 (6-7), 2014, 379-390
- [14] W. Li and X. Wang, Locally Aligned Feature Transforms across Views, *CVPR 2013*
- [15] R. Layne, T. M. Hospedales, S. Gong, Towards person identification and re-identification with attributes, *ECCV Workshops 2012*
- [16] Z. Wu, Y. Li, and R. Radke, Viewpoint invariant human re-identification in camera networks using pose priors and subject-discriminative features, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37 (5), 2014, 1095-1108
- [17] S. Liao, Y. Hu, X. Zhu, and S. Z. Li, Person re-identification by local maximal occurrence representation and metric learning, *CVPR 2015*
- [18] R. Zhao, W. Ouyang, and X. Wang, Learning mid-level filters for person re-identification, *CVPR 2014*
- [19] W. Li, R. Zhao, T. Xiao, and X. Wang, Deepreid: Deep filter pairing neural network for person re-identification, *CVPR 2014*
- [20] Wu, Y. Y. Chen, X. Li, A. Wu, J. You, and W.-S. Zheng, An enhanced deep feature representation for person re-identification, *WACV 2016*
- [21] E. Ahmed, M. Jones, and T. K. Marks, An improved deep learning architecture for person re-identification, *CVPR 2015*
- [22] B. I. Barbosa, M. Cristani, A. Del Bue, L. Bazzani, and V. Murino, Re-identification with rgb-d sensors, *ECCV Workshops 2012*
- [23] M. Munaro, A. Basso, A. Fossati, L. V. Gool, and E. Menegatti, 3D reconstruction of freely moving persons for re-identification with a depth sensor, *ICRA 2014*
- [24] B. Takac, A. Catala, M. Rauterberg and W. Chen, People identification for domestic non-overlapping RGB-D camera networks, In *Proceedings of International Multi-Conference on Systems, Signals & Devices*, 2014

- [25] J. Oliver, A. Albiol, A. Albiol, 3D descriptor for people re-identification, CVPR 2012
- [26] B. Prosser, W.-S. Zheng, S. Gong, and T. Xiang, Person re-identification by support vector ranking, BMVC 2010
- [27] W.-S. Zheng, S. Gong, and T. Xiang, Re-identification by relative distance comparison, IEEE Transactions on Pattern Analysis and Machine Intelligence, 35 (3), 2013, 653-668
- [28] M. Kostinger, M. Hirzer, P. Wohlhart, P. M. Roth, and H. Bischof, Large scale metric learning from equivalence constraints, CVPR 2012
- [29] D. Tao, L. Jin, Y. Wang, Y. Yuan, and X. Li, Person re-identification by regularized smoothing kiss metric learning, IEEE Transactions on Circuits and Systems for Video Technology, 23(10), 2013, 1675-1685
- [30] D. Tao, L. Jin, Y. Wang, and X. Li, Person re-identification by minimum classification error-based kiss metric learning, IEEE Transactions on Cybernetics, 45 (2), 2014, 242-252
- [31] S. Pedagadi, J. Orwell, S. Velastin, and B. Boghossian, Local fisher discriminant analysis for pedestrian re-identification, CVPR 2013
- [32] Mignon and F. Jurie, Pcca: A new approach for distance learning from sparse pairwise constraints, CVPR 2012
- [33] Y. Xu, L. Liang, W.-S. Zheng, and X. Liu, Human re-identification by matching compositional template with cluster sampling, ICCV 2013
- [34] Z. Li, S. Chang, F. Liang, T. S. Huang, L. Cao, and J. R. Smith, Learning locally-adaptive decision functions for person verification, CVPR 2013
- [35] L. Ma, X. Yang, and D. Tao, Person re-identification over camera networks using multi-task distance metric learning, IEEE Transactions on Image Processing, 23 (8), 2014, 3656-3670
- [36] G. Lisanti, I. Masi, A. Bagdanov, and A. Del Bimbo, Person re-identification by iterative re-weighted sparse ranking, IEEE Transactions on Pattern Analysis and Machine Intelligence, 37 (8), 2014, 1629-1642
- [37] R. Zhao, W. Ouyang, and X. Wang, Unsupervised salience learning for person re-identification, CVPR 2013
- [38] R. Zhao, W. Ouyang, and X. Wang, Person re-identification by salience matching, ICCV 2013
- [39] R. Zhao, W. Ouyang, and X. Wang, Person re-identification by saliency learning, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016
- [40] X. Liu, M. Song, D. Tao, X. Zhou, C. Chen, and J. Bu, Semi-Supervised Coupled Dictionary Learning for Person Re-identification, CVPR 2014
- [41] F. Xiong, M. Gou, O. Camps, M. Sznajder, Person re-identification using kernel based metric learning methods, ECCV 2014
- [42] D. Yi, Z. Lei, S. Liao, and S. Z. Li, Deep metric learning for person re-identification, ICPR 2012
- [43] S. Karanam, Y. Li, and R. J. Radke, Person re-identification with discriminatively trained viewpoint invariant dictionaries, ICCV 2015
- [44] S. Liao and S. Z. Li, Efficient PSD constrained asymmetric metric learning for person re-identification, ICCV 2015
- [45] Y. Shen, W. Lin, J. Yan, M. Xu, J. Wu, and J. Wang, Person re-identification with correspondence structure learning, ICCV 2015
- [46] J. Garcia, N. Martinel, C. Micheloni, and A. Gardel, Person re-identification ranking optimisation by discriminant context information analysis, ICCV 2015
- [47] Y. C. Chen, W.-S. Zheng, J. H. Lai, and P. Yuen, An asymmetric distance model for cross-view feature mapping in person re-identification, IEEE Transactions on Circuits and Systems for Video Technology, 2016
- [48] X. Wang, W.-S. Zheng, X. Li, and J. Zhang, Cross-scenario transfer person re-identification, IEEE Transactions on Circuits and Systems for Video Technology, PP(99), 2015
- [49] T. Wang, S. Gong, X. Zhu, and S. Wang, Person re-identification by video ranking, ECCV 2014
- [50] T. Wang, S. Gong, X. Zhu, and S. Wang, Person re-identification by discriminative selection in video ranking, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016

- [51] W.-S. Zheng, S. Gong, and T. Xiang, Associating Groups of People, BMVC 2009
- [52] W.-S. Zheng, S. Gong, and T. Xiang, Transfer re-identification: From person to set-based verification, CVPR 2012
- [53] W.-S. Zheng, S. Gong, and T. Xiang, Towards open-world person re-identification by one-shot group-based verification, IEEE Transactions on Pattern Analysis and Machine Intelligence, 38 (3), 2015, 591-606
- [54] R. Layne, T. M. Hospedales, and S. Gong, Domain transfer for person re-identification, In Proceedings of the 4th ACM/IEEE international workshop on Analysis and retrieval of tracked events and motion in imagery stream, 2013
- [55] W. Li, R. Zhao, and X. Wang, Human reidentification with transferred metric learning, ACCV 2012
- [56] T. Avraham, I. Gurvich, M. Lindenbaum, and S. Markovitch, Learning implicit transfer for person re-identification, ECCV Workshops 2012
- [57] J. Ma, P. C. Yuen, and J. Li, Domain transfer support vector ranking for person re-identification without target camera label information, ICCV 2013
- [58] L. Zheng, L. Shen, L. Tian, S. Wang, J. Wang, and Q. Tian, Scalable person re-identification: a benchmark, ICCV 2015
- [59] D. S. Cheng, M. Cristani, M. Stoppa, L. Bazzani, and V. Murino, Custom pictorial structures for re-identification, BMVC 2011
- [60] Y.-C. Chen, W.-S. Zheng, and J. Lai, Mirror representation for modeling view-specific transform in person re-identification, IJCAI 2015
- [61] Y. Yang, J. Yang, J. Yan, S. Liao, D. Yi, and S. Z. Li, Salient color names for person re-identification, ECCV 2014
- [62] C. Su, F. Yang, S. Zhang, Q. Tian, L. S. Davis, and W. Gao, Multi-task learning with low rank attribute embedding for person re-identification, ICCV 2015
- [63] C. Su, S. Zhang, J. Xing, W. Gao, and Q. Tian, Deep attributes driven multi-camera person re-identification, ECCV 2016
- [64] T. Xiao, H. Li, W. Ouyang, and X. Wang, Learning deep feature representations with domain guided dropout for person re-identification, CVPR 2016.
- [65] E. Kodirov, T. Xiang, Z. Fu, and S. Gong, Person re-identification by unsupervised ℓ_1 graph learning, ECCV 2016
- [66] T. Matsukawa, T. Okabe, E. Suzuki and Y. Sato, Hierarchical Gaussian Descriptor for Person Re-identification, CVPR 2016
- [67] Liming Zhao, Xi Li, Yueting Zhuang, and Jingdong Wang. Deeply-Learned Part-Aligned Representations for Person Re-Identification, ICCV 2017.
- [68] Dangwei Li, Xiaotang Chen, Zhang Zhang, Kaiqi Huang. Learning Deep Context-aware Features over Body and Latent Parts for Person Re-identification, CVPR 2017
- [69] Haiyu Zhao, Maoqing Tian, Shuyang Sun, Jing Shao, Junjie Yan, Shuai Yi, Xiaogang Wang, Xiaoou Tang, Spindle Net Person Re-identification with Human Body Region Guided Feature Decomposition and Fusion, CVPR 2017
- [70] Haque, A. Alahi and L. Fei-Fei, Recurrent Attention Models for Depth-Based Person Identification, CVPR 2016
- [71] Wu, W.-S. Zheng and J. H. Lai, Robust Depth-Based Person Re-Identification, IEEE Transactions on Image Processing, vol. 26, no. 6, pp. 2588-2603, June 2017.
- [72] Liangliang Ren, Jiwen Lu, Jianjiang Feng, Jie Zhou, Multi-modal uniform deep learning for RGB-D person re-identification, Pattern Recognition, vol. 72, pp. 446-457, Dec. 2017.
- [73] Weihua Chen, Xiaotang Chen, Jianguo Zhang, Kaiqi Huang, Beyond triplet loss a deep quadruplet network for person re-identification, CVPR 2017
- [74] P. Peng et al., Unsupervised Cross-Dataset Transfer Learning for Person Re-identification, CVPR 2016
- [75] N. McLaughlin, J. M. d. Rincon and P. Miller, Recurrent Convolutional Network for Video-Based Person Re-identification, CVPR 2016
- [76] Jiaxin Chen, Yunhong Wang, Jie Qin, Li Liu, Ling Shao, Fast Person Re-identification via Cross-camera Semantic Binary Transformation, CVPR 2017
- [77] Xiatian Zhu, Botong Wu, Dongcheng Huang, Wei-Shi Zheng, Fast Open-World Person Re-Identification, IEEE Transactions on Image Processing, in press, 2017 (DOI: 10.1109/TIP.2017.2740564).

- [78] Shuang Li, Tong Xiao, Hongsheng Li, Bolei Zhou, Dayu Yue, Xiaogang Wang, Person Search with Natural Language Description, CVPR 2017
- [79] Tong Xiao, Shuang Li, Bochao Wang, Liang Lin, Xiaogang Wang, Joint Detection and Identification Feature Learning for Person Search, CVPR 2017
- [80] Hongxing Yu, Ancong Wu, Wei-Shi Zheng. Cross-view Asymmetric Metric Learning for Unsupervised Person Re-identification. ICCV 2017
- [81] Jiening Jiao, Wei-Shi Zheng, Ancong Wu, Xiatian Zhu, and Shaogang Gong. Deep Low-resolution Person Re-identification. AAAI 2018
- [82] Ancong Wu, Wei-Shi Zheng, Hongxing Yu, Shaogang Gong, Jianhuang Lai. RGB-Infrared Cross-Modality Person Re-Identification. ICCV 2017
- [83] Jinjie You, Ancong Wu, Xiang Li, Wei-Shi Zheng. Top-push Video-based Person Re-identification. CVPR 2016
- [84] Wei-Shi Zheng, Xiang Li, Tao Xiang, Shengcai Liao, JianHuang Lai, Shaogang Gong. Partial Person Re-identification. ICCV 2015
- [85] Xiang Li, Wei-Shi Zheng, Xiaojuan Wang, Tao Xiang, Shaogang Gong. Multi-scale Learning for Low-resolution Person Re-identification. ICCV 2015.
- [86] X. Zhu, X.-Y. Jing, F. Wu, and H. Feng, Video-based person reidentification by simultaneously learning intra-video and inter-video distance metrics. IJCAI 2016.
- [87] Z. Zhou, Y. Huang, W. Wang, L. Wang, and T. Tan, See the forest for the trees: Joint spatial and temporal recurrent neural networks for video-based person re-identification. CVPR 2017.
- [88] W. Zhang, X. Yu, and X. He, Learning bidirectional temporal cues for video-based person re-identification, IEEE Transactions on Circuits and Systems for Video Technology, 2017.
- [89] D. Chung, K. Tahboub, and E. J. Delp, A two stream siamese convolutional neural network for person re-identification, CVPR 2017
- [90] Xiao-Yuan Jing, Xiaoke Zhu, Fei Wu, Ruimin Hu, Xinge You, Yunhong Wang, Hui Feng, and Jing-Yu Yang, Super-Resolution Person Re-Identification With Semi-Coupled Low-Rank Discriminant Dictionary Learning, in IEEE Transactions on Image Processing, vol. 26, no. 3, pp. 1363-1378, March 2017.



郑伟诗

中山大学数据科学与计算机学院教授，机器智能与先进计算教育部重点实验室副主任，国家优秀青年科学基金获得者、广东省自然科学基金杰出青年基金获得者。主要研究方向是视频监控下的行人身份识别与行为信息理解。在 IEEE TPAMI、IEEE TIP、IEEE TNN 等国际期刊和 ICCV、CVPR、IJCAI 等国际会议发表学术论文 80 余篇。曾担任 IEEE AVSS 的 Area Chair 和 Publication Chair，担任 2012 年、2015 年和 2016 年全国生物特征识别学术会议的联合程序委员会主席。

Email: wszheng@ieee.org



吴岸聪

中山大学电子与信息工程学院博士生。主要研究方向是计算机视觉、行人重识别。

Email: wuancong@mail2.sysu.edu.cn

群目标行为理解：从运动表观到内在驱动

清华大学 苏航 董胤蓬 邹昊晟

随着我国社会和经济的快速发展，群体活动的规模和频率都呈现大幅度增加趋势。这就迫切需要发挥视频分析等技术的优势，及时发现异常运动目标或者行为区域，进而采取措施以降低危险事件发生概率，提高我国社会的安全管理水平。

从本质上来说，目标在群体中的运动过程实际上是其内在决策过程一种外化表现，如果能够表观运动来解析其个体的内在决策要素，并利用这些决策要素来深入理解运动模式，具有重要的意义。根据社会力模型的研究显示，一个典型的个体在群体当中除了考虑自身的运动之外，同时需要考虑环境约束以及目标时间的相互作用。因此，我们提出基于模仿学习的群目标运动决策要素解析算法，利用 virtual agent 对运动决策过程的模仿，分析人类在群体中运动中的决策要素，并利用这些决策要素特征实现行为预测和组群状态分析。同运动轨迹等表观特征相比，基于决策要素的运动目标分析可以更好的反映其运动的本质因素，受到遮挡、扰动等干扰较小。

另一方面，运动群目标可以从宏观上被当作一个包含大量运动粒子的流体场，因而可以融合流体力学和计算机视觉的相关理论，来实现群目标运动的模式发现和分析。通常来说，规则运动的群目标可以利用稳定的层流场进行建模，而人群中潜在高危区域则可以利用紊流场来解释。群



图 1 群目标行为的分析和理解

目标行为的分析重要目标则是及时发现潜在的高分析群体运动的相关结构，对人群中潜在高危区域(紊流区)进行及时检测，为人群管控提供依据。在这方面，我们提出利用粘性流体对运动高密度群体进行建模，利用流体场的散度和旋度等分析工具对群体内目标之间的交互关系进行理解，并进一步应用深度递归神经网络等相关技术，提取群目标运动轨迹的隐含特征，有效地实现了群目标的运动模式发现和运动趋势预测，为群体管理提供了重要手段。

近 5 年来，课题组在群目标行为的理解和预测等领域包括 IJCV、IJCAI、AAAI、CVPR 等相关论文 20 余篇，完成国家自然科学基金等多个科研项目，提出的基础粘性流体的群体运动建模方法获得 AVSS 最佳论文奖。

(责任编辑：申抒含)



苏航

CCF 计算机视觉专委会委员，CCAI 机器学习专委会通讯委员，清华大学计算机系助理研究员。在 CVPR、IJCAI 等国际会议和期刊发表论文 50 余篇，曾获得 MICCAI “青年学者”、AVSS “最佳论文”等奖项。

Email: suhangss@mail.tsinghua.edu.cn

基于深度学习的高分遥感图像理解

西北工业大学 程堃 韩军伟

我国于 2010 年启动的高分辨率对地观测技术显著提高了我国自主获取高分遥感影像的技术水平。高分辨率对地观测技术的终极目标是对遥感影像中隐藏的知识进行挖掘，实现遥感观测数据向知识的转化。而高分遥感图像理解，是遥感数据信息提取的关键技术，具有重要的应用价值。

目前深度学习已经成为高分遥感图像理解的主流方法，但也存在着一些难点问题：1) 受限于遥感图像标注的难度，已有的遥感数据集规模较小；2) 高分遥感图像中的地物目标具有各种各样不同的朝向，如何实现旋转不变特征提取是一个难点问题；3) 高分遥感图像包含大量的场景语义信息，且类内差异性和类间相似性问题突出。

针对上述问题，我们围绕图 1 所示的整体研究思路，取得了一些代表性的研究工作：1) 构建并发布了遥感图像场景分类数据库“NWPU-RESISC45”，该数据库共有 31500 幅图像，涵盖了 45 个场景类别，是目前遥感图像场景分类最大规模的数据集；2) 提出了基于弱监督学习的高分遥

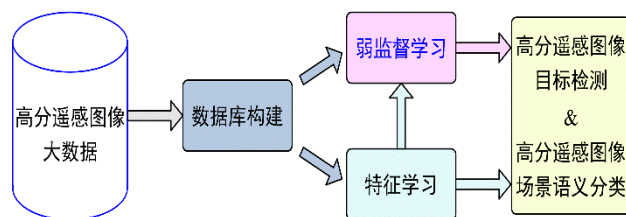


图 1 高分遥感图像理解整体框架

感图像理解方法，实现了在只有图像级标签的情况下可以准确地进行目标检测和场景语义标注；3) 提出了基于旋转不变卷积神经网络的高分遥感图像目标检测方法；4) 为了解决场景语义的类内多样性和类间相似性问题，提出了基于判别地物目标属性特征和基于深度度量学习的遥感图像场景理解方法。

以上工作获得了包括国家自然科学基金在内的多个科研项目支持。这些工作发表在 PIEEE、IEEE T-GRS、CVPR 等国际期刊和国际会议，并获得教育部自然科学二等奖一项。

(责任编辑：邓成)



程堃

CCF 计算机视觉专委会委员，西北工业大学副研究员。主要研究领域为计算机视觉、模式识别、高分遥感图像理解等。

E-mail: gcheng@nwpu.edu.cn



韩军伟

CCF 计算机视觉专委会委员，IEEE 高级会员，西北工业大学教授，自动化学院副院长，信息融合技术教育部重点实验室副主任。主要研究领域为计算机视觉和多媒体处理等。

E-mail: jhan@nwpu.edu.cn