

## 基于多准则特征融合的自适应跟踪算法

赵 谦<sup>1</sup>, 周 勇<sup>1</sup>, 曾召华<sup>1</sup>, 侯媛彬<sup>2\*</sup>, 刘树林<sup>2</sup>

(1. 西安科技大学 通信与信息工程学院, 西安 710054; 2. 西安科技大学 电气与控制工程学院, 西安 710054)

(\* 通信作者电子邮箱 zhouyong205@126.com)

**摘 要:**针对现有基于多特征融合的跟踪算法在复杂环境下跟踪准确度不高,且大部分采用单一判定方式来实现多特征融合的问题,提出了一种综合判定的自适应多特征融合跟踪方法。首先引入局部背景信息加强对目标的描述,然后在多特征融合过程中利用多种判定准则计算特征权值,最后在均值漂移框架下完成对目标的跟踪。在各种场景下的实验结果比较表明:该融合算法比单种判定融合有更好的稳定性和鲁棒性,有效地提高了复杂环境下跟踪准确性。

**关键词:**目标跟踪;多特征融合;多准则;局部背景;均值漂移

**中图分类号:**TP391.41 **文献标志码:**A

### Adaptive tracking algorithm based on multi-criteria feature fusion

ZHAO Qian<sup>1</sup>, ZHOU Yong<sup>1</sup>, ZENG Zhao-hua<sup>1</sup>, HOU Yuan-bin<sup>2\*</sup>, LIU Shu-lin<sup>2</sup>

(1. School of Communication and Information Engineering, Xi'an University of Science and Technology, Xi'an Shaanxi 710054, China;

2. School of Electrical and Control Engineering, Xi'an University of Science and Technology, Xi'an Shaanxi 710054, China)

**Abstract:** Multiple feature fusion based tracking is one of the most active research topic in tracking field, but the tracking accuracy needs improving in complex environment and most of them use single fusion rule. In this paper, a new adaptive fusion strategy was proposed for multi-feature fusion. First, the local background information was introduced to strengthen the description of the target, and then the feature weight was calculated by a variety of criteria in the fusion process. In addition, the framework of mean shift was considered to realize target tracking. An extensive number of comparative experimental results show that the proposed algorithm is more stable and robust than the single fusion rule.

**Key words:** target tracking; multiple feature fusion; multi-criteria; local background; mean shift

## 0 引言

实时目标跟踪在视频监控、视觉导航、人机交互、三维重建、车辆自动驾驶系统等众多领域有着广泛的应用<sup>[1]</sup>。文献[2]提出的均值迁移算法中以颜色为特征实现目标跟踪,但是在光照发生变化,有相似颜色干扰时,跟踪结果往往会出现偏差,甚至丢失目标。

近年来,研究者们提出了多种基于多特征融合的跟踪算法。文献[3]提出了一种基于 Fisher 可分性度量的权值更新机制,通过选择对当前场景区分性较高的特征模型进行跟踪;文献[4]提出了基于不确定性度量的多特征融合跟踪算法,这些方法提高了跟踪准确率但计算复杂度较高。文献[5-6]利用目标与候选目标之间的特征相似度来选择特征。文献[7]利用概率分部可分性自适应地计算特征融合权重。但是当有背景相似物干扰时,这些方法可靠性大大降低。文献[8]首先计算目标与背景直方图的 log 比率,自动选择区分性较大的特征(用三个颜色通道组合表示)。在文献[9]中,Comaniciu 试图将局部背景信息引入,提出了背景加权直方图 (Background-Weighted Histogram, BWH) 算法。但文献[10]证

明了 Comaniciu 的方法并没有引入背景信息,通过仅在目标直方图上引入背景加权的方法正确地使用局部背景信息,实现了纠正背景加权直方图 (Corrected Background-Weighted Histogram, CBWH) 算法,不仅提高了跟踪的可靠性,而且对目标的初始化相对不敏感。上述跟踪算法通过纳入背景信息和使用多特征融合方式提高了跟踪的准确性,主要不同之处是背景的使用方式和特征融合方式的差异。文献[3-4, 7-8]使用局部背景来选择合适特征;文献[9-10]利用背景加强对目标的描述。在文献[3-10]的启发下,本文期望将多特征用于目标跟踪,同时引入局部背景信息,加强目标的多特征描述能力。其特点和优势在于:

1) 目标描述中加入了局部背景信息。在目标提取时,会引入部分背景,影响跟踪可靠性。利用目标周围的局部背景与目标中部分背景的相似性,使目标点权重高而背景点权重低,减少背景对目标的干扰,因此能更精确地描述目标。

2) 多准则的多特征融合方式跟踪目标,一方面可以方便利用更多目标信息,另一方面提高特征权重的可靠性。因此,相对于单一特征和单一判定算法,本文算法的跟踪性能更稳定。

**收稿日期:**2013-04-01; **修回日期:**2013-05-17。 **基金项目:**国家自然科学基金资助项目(50977077);陕西省科技研究发展计划工业攻关项目(2012K06-49;2013K07-35);陕西省教育厅科研计划项目(12JK0508, 12JK0535);西安市科技计划项目(CX1258⑤, CX1258⑥, CXY1343(4));西安市碑林区应用技术研发项目(GX1209, GX1311, GX1310)。

**作者简介:**赵谦(1977-),男,陕西西安人,副教授,博士研究生,主要研究方向:图像智能检测、3D 虚拟现实;周勇(1986-),男,河南信阳人,硕士研究生,主要研究方向:图像处理、运动目标跟踪;曾召华(1972-),男,湖北红安人,副教授,博士,主要研究方向:B3G 移动通信、图像处理;侯媛彬(1953-),女,陕西西安人,教授,博士生导师,主要研究方向:智能控制、神经网络;刘树林(1964-),男,四川成都人,教授,博士生导师,主要研究方向:开关变压器及本质安全电路设计。

## 1 背景加权目标多特征描述与相似性度量

在目标跟踪中,如果目标中有较多背景,跟踪的精度将会降低。本文在文献[10]的基础上引入多特征,这里的特征可以是颜色、边缘、纹理、光流和尺度不变特征转换(Scale-Invariant Feature Transform, SIFT)等。计算目标背景的概率密度函数,用  $\{\delta_u\}_{u=1,2,\dots,m}$  ( $\sum_{u=1}^m \delta_u = 1$ ) 表示,背景区域的大小为目标的两倍,其中  $\delta^*$  为  $\{\delta_u\}_{u=1,2,\dots,m}$  中非零最小值。式(1)用于对目标进行更好的描述,降低目标区域中与背景相似的权重。

$$\{v_u = \min(\delta^*/\delta_u, 1)\}_{u=1,2,\dots,m} \quad (1)$$

### 1.1 目标模型描述

基于多特征的目标模型  $q = \{\{\hat{q}_{u_i}\}_{u_i=1,2,\dots,m_i}\}_{i=1,2,\dots,N}$ , 其中  $\{\hat{q}_{u_i}\}$  是第  $i$  个特征的概率密度函数,  $m_i$  为第  $i$  个特征的 bin 值,  $N$  为特征个数。假设  $\{x_j^*\}_{j=1,2,\dots,n}$  是目标区域像素位置,  $k(x)$  为核函数的轮廓函数,  $b_i(x_j^*)$  是  $x_j^*$  的第  $i$  个特征的 bin 值,  $C$  为归一化常量。

$$\hat{q}_{u_i} = C_i v_{u_i} \sum_{j=1}^n k(\|x_j^*\|^2) \delta[b_i(x_j^*) - u_i] \quad (2)$$

其中:归一化常量  $C_i = \frac{1}{\sum_{j=1}^n k(\|x_j^*\|^2) \sum_{u=1}^m v_{u_i} \delta[b_i(x_j^*) - u_i]}$ ,  $v_{u_i}$  为第  $i$  个特征的背景权重。候选目标模型可以表示为  $P(y) = \{\{\hat{p}_{u_i}(y)\}_{u_i=1,2,\dots,m_i}\}_{i=1,2,\dots,N}$ , 其中  $\{\hat{p}_{u_i}(y)\}$  是第  $i$  个特征的概率密度函数,  $y$  是候选目标的位置, 给定带宽为  $h$  的轮廓函数  $k(x)$ , 则  $\hat{p}_{u_i}(y)$  的计算方法如下:

$$\hat{p}_{u_i}(y) = C_h \sum_{j=1}^n k\left(\left\|\frac{y-x_j}{h}\right\|^2\right) \delta[b_i(x_j) - u_i] \quad (3)$$

其中:归一化常量  $C_h = \frac{1}{\sum_{j=1}^n k\left(\left\|\frac{y-x_j}{h}\right\|^2\right)}$ 。

### 1.2 相似性函数

目标模型  $q$  与候选目标模型  $P(y)$  之间某种特征相似性用 Bhattacharyya 系数  $\hat{\rho}_i(y)$  来度量分布, 其值越大表示两个模型越相似。

$$\hat{\rho}_i(y) = \rho[p(y), q] = \sum_{u_i=1}^{m_i} \sqrt{\hat{p}_{u_i}(y) \hat{q}_{u_i}} \quad (4)$$

基于式(4)可以得出目标模型和候选目标模型之间的多特征相似系数  $\hat{\rho}(y)$  如下:

$$\hat{\rho}(y) = \sum_{i=1}^N k_i \hat{\rho}_i(y) = \sum_{i=1}^N k_i \sum_{u_i=1}^{m_i} \sqrt{\hat{p}_{u_i}(y) \hat{q}_{u_i}} \quad (5)$$

其中:  $k_i$  为自适应融合权值, 且满足  $\sum_{i=1}^N k_i = 1$ 。

## 2 多准则特征融合均值漂移目标定位

### 2.1 多特征融合目标定位

假设目标在上一帧图像中的位置为  $\hat{y}_0$ , 则在当前帧中以  $\hat{y}_0$  为中心的候选目标的确定, 就是寻找使相似度  $\hat{\rho}(y)$  最大化的过程, 即使  $\sum_{i=1}^N k_i \hat{\rho}_i(y)$  中的每一项  $k_i \hat{\rho}_i(y)$  最大化。对  $k_i \hat{\rho}_i(y)$  在  $\hat{y}_0$  处进行 Taylor 展开, 并将  $\hat{p}_{u_i}(y)$  代入可得:

$$k_i \hat{\rho}_i[p(y), q] \approx \frac{1}{2} k_i \sum_{u_i=1}^{m_i} \sqrt{\hat{p}_{u_i}(\hat{y}_0) \hat{q}_{u_i}} + \frac{C_h}{2} k_i \sum_{j=1}^n w_{ij} k\left(\left\|\frac{y-x_j}{h}\right\|^2\right) \quad (6)$$

其中:  $w_{ij} = \sum_{u_i=1}^{m_i} \sqrt{\hat{q}_{u_i} / \hat{p}_{u_i}(\hat{y}_0)} \delta[b_i(x_j) - u_i]$ 。

由于式(6)中右边第一项与  $y$  无关, 通过迭代时第二项最大。在进行多特征融合时, 对所有项  $k_i \hat{\rho}_i(y)$  均进行最大化, 再根据各特征的有效性, 将它们线性加权组合起来, 得到由多特征共同决定的新的目标中心  $\hat{y}_1$ 。其中  $g(x) = -k'(x)$ , 特征融合权值  $k_i$  满足  $\sum_{i=1}^N k_i = 1$ 。

$$\hat{y}_1 = \sum_{i=1}^N k_i \frac{\sum_{j=1}^{n_R} x_j w_{ij} g\left(\left\|\frac{\hat{y}_0 - x_j}{h}\right\|^2\right)}{\sum_{j=1}^{n_R} w_{ij} g\left(\left\|\frac{\hat{y}_0 - x_j}{h}\right\|^2\right)} \quad (7)$$

式(7)即为多特征融合均值漂移目标定位公式, 其中  $N$  为特征的个数, 值得注意的是并非特征数越多越好, 考虑到一般的场景和处理的实时性, 选择 2 至 3 个特征即可。

### 2.2 自适应多准则融合权重计算

文献[8]中指出, 一种有效的特征使用简单的方法也能达到较好的跟踪效果, 但由于背景的不可预知性和场景的复杂性, 自适应融合多种特征来实现鲁棒跟踪成为众多研究者的选择, 这种方式在粒子滤波中已被广泛使用。均值漂移框架下的多特征融合都是使用单一判定方式, 文献[5-6]使用目标模板与候选目标的相似度来计算融合的权重系数。但是如果背景与目标相似度较高, 将无法区分较高相似度是来自目标还是背景, 很有可能将较小区分度的特征判定为有效的特征。针对这种问题, 文献[7-8]通过目标与背景的区分度, 来进行特征的选择, 认为目标与背景区分度较大的特征是有效特征, 易于稳定跟踪, 将赋予较高的权值。该种判定方式较好地适应了背景的变化, 但是在处理复杂背景时, 效果有所下降, 特别是当跟踪框偏离目标时, 会将目标中的部分当成背景, 从而导致错误的判断。

结合上述分析, 本文将采用多准则判定特征的有效性。当特征与背景区分能力较弱且相似度相对较小时容易造成目标丢失。由此, 本文从目标与背景的区分能力和与目标模板的相似度两个方面分析特征在跟踪中的有效性。由式(1)计算候选目标的背景权重, 式(8)计算候选目标与目标模板的相似度, 该公式是目标在不同场景下特征鉴别能力判断的依据。

$$\rho'_i(y) = \rho[p_i(y) v'_i, q_i] = \sum_{u_i=1}^{m_i} \sqrt{\hat{p}_{u_i}(y) v'_{u_i} \hat{q}_{u_i}} \quad (8)$$

其中:  $p_i(y)$  表示候选目标模型,  $v'_i$  表示候选目标背景权重,  $q_i$  为目标模型,  $i = 1, 2, \dots, N$  (其中  $N$  的取值一般不超过 4, 这里取  $N = 2$ )。

采用线性加权的方法来获得权重,  $\rho'_i(y)$  表示融合背景信息的候选目标与目标模板的相似度。令  $k_i$  表示第  $i$  个特征的权重, 则有

$$k_i = \rho'_i / \sum_{i=1}^N \rho'_i \quad (9)$$

其中:  $k_i$  满足  $\sum_{i=1}^N k_i = 1$ 。

在考虑背景信息的基础上,如果候选目标与局部背景区分度较高且与目标模型相似度则权值较大;反之较小。图1给出了视频1跟踪过程中颜色及梯度特征权值的自适应调整过程,初始跟踪时相对于梯度特征,颜色特征具有更好的可分性,当在第168帧光照发生变化时,颜色特征的权值迅速减小,而梯度特征的权值会迅速增加,从而克服了光照变化的影响。

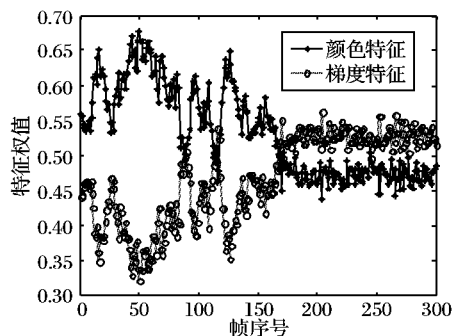


图1 特征权值的自适应过程

### 3 算法实现

实现流程如图2所示,其主要由三个模块组成:目标初始化、目标跟踪与目标更新。为了提高目标的跟踪效率,同时处理部分遮挡等情况,将前面提出的算法应用于Kalman预测框架<sup>[11]</sup>,其中的状态参量为  $[x, y, v_x, v_y]^T$ , 量测参量  $[x, y]^T$ , 其中  $x, y, v_x$  和  $v_y$  分别表示目标在  $x, y$  方向的位置和速度。在实时跟踪中,由于目标外观变化,致使长时间跟踪易出现目标漂移,导致跟踪的失败。因此,需要对目标描述进行更新,本文采用文献[3]更新方式对目标进行更新,在每次跟踪过程中对特征进行检测并更新其中相似度小的特征。

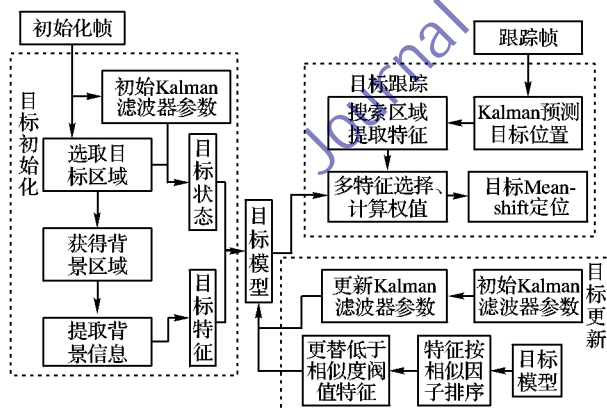


图2 算法流程图

### 4 实验结果及分析

为验证本文算法的有效性,在标准测试视频上将本文算法与文献[5,8,10]中的三种跟踪算法进行了比较,目标初始位置由手动给定。在酷睿双核2.2 GHz,2 GB内存的计算机上用Matlab与C++混合编程实现。文献中实验按照文中提供的条件进行,本文算法特征点的描述由两部分联合而成:颜色(RGB)直方图  $16 \times 16 \times 16$  和36维的梯度直方图<sup>[12]</sup>。

实验1是背景为弱光照条件下的行人跟踪。该视频序列

主要受室内灯光和室外自然光的影响,人工光照条件下,使颜色失真且部分背景区域明暗不一,引起场景杂乱,从而使目标与背景特征区分性较低。其次,目标外观与部分背景(柱子及店内的衣物)很相近。如图3所示,当目标与背景中物体相似时,文献[5,10]算法丢失跟踪(556帧),文献[8]算法经过立柱时丢失跟踪(588帧),而本文算法可以很好地持续跟踪。

实验2是自然光照下室外场景的行人跟踪。该视频序列主要与背景的颜色相近,同时目标较小,特征不明显易受外界的干扰。如图4所示,视频序列在接近附近的车辆时由于目标与局部背景难以区分导致文献[5](2360帧)算法和文献[8](2365帧)算法丢失跟踪,虽然文献[10]算法能持续跟踪,但是跟踪框跳动频繁跟踪性能不稳定,而本文算法可以很好地持续稳定跟踪。

实验所用的视频来自 <http://homepages.inf.ed.ac.uk/rbf/CAVIARDATA1/>。首先从视觉上对比以上三种算法,然后将跟踪结果与人工标注的结果进行定量分析。在前述实验参数设置下,表1给出了不同算法在跟踪过程中目标的跟踪中心坐标与手动标注中心坐标的偏差绝对值的统计量( $X_{max}$ 、 $X_{min}$ 、 $X_{mean}$ 分别是中心坐标在水平方向偏差最大值、最小值、平均值; $Y_{max}$ 、 $Y_{min}$ 、 $Y_{mean}$ 分别是中心坐标在垂直方向偏差最大值、最小值、平均值),进一步说明了本文算法在跟踪准确性方面的优势。

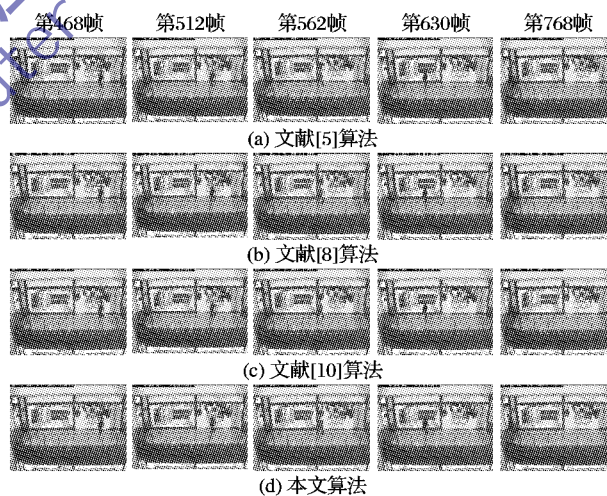


图3 实验1跟踪结果

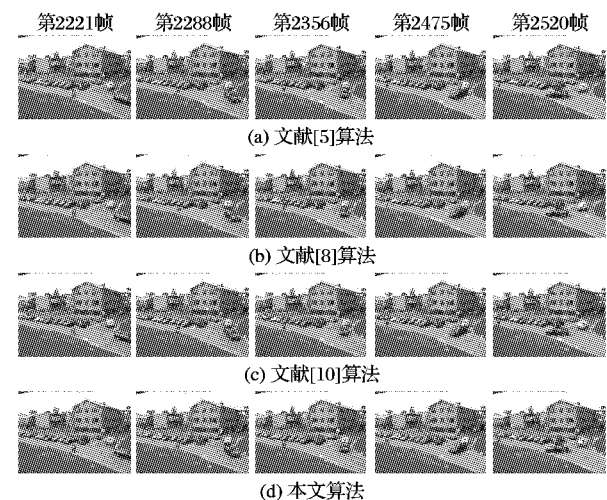


图4 实验2跟踪结果



表1 跟踪中心坐标偏差统计及平均跟踪速度对比

实验	算法	目标的跟踪中心坐标与手动标注中心坐标的偏差绝对值/像素						平均跟踪速度/ (frame · s <sup>-1</sup> )
		<i>X</i> -max	<i>X</i> -min	<i>X</i> -mean	<i>Y</i> -max	<i>Y</i> -min	<i>Y</i> -mean	
实验1	文献[5]算法	202	0	74.3	67	0	35.40	12.0
	文献[8]算法	178	0	55.6	46	0	9.60	9.6
	文献[10]算法	156	0	67.2	62	0	43.30	17.4
	本文算法	6	0	2.2	5	0	1.60	16.5
实验2	文献[5]算法	24	0	10.3	53	0	23.20	15.1
	文献[8]算法	22	0	12.1	18	0	9.30	12.3
	文献[10]算法	14	0	8.7	10	0	6.80	23.6
	本文算法	4	0	1.5	3	0	1.40	21.0

## 5 结语

通过大量分析前人的优秀研究成果,本文将目标的局部背景信息引入目标描述,减少背景干扰;其次在多特征融合跟踪过程中采用多准则的融合判定方式,有效地提高特征权值计算的准确性;最后嵌入到 Mean-shift 跟踪框架下来进行目标定位跟踪。在不同实验环境下表明:本文算法在背景复杂、光照变化和相似物干扰情况下都能较好地持续稳定跟踪,具有较好的跟踪鲁棒性。

### 参考文献:

- [1] 侯志强, 韩崇昭. 视觉跟踪技术综述[J]. 自动化学报, 2006, 32(4): 603-617.
- [2] COMANINI D, MEER P. Mean-shift: a robust approach toward feature space analysis[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(5): 603-619.
- [3] 王永忠, 梁彦, 赵春晖, 等. 基于多特征自适应融合的核跟踪方法[J]. 自动化学报, 2008, 34(4): 393-399.
- [4] 顾鑫, 王海涛, 汪凌峰, 等. 基于不确定性度量的多特征融合跟踪[J]. 自动化学报, 2011, 37(5): 551-559.
- [5] 郑玉凤, 马秀荣, 赵晓琳, 等. 基于颜色和边缘特征的均值迁移目

标跟踪算法[J]. 光电子·激光, 2011, 22(8): 1232-1235.

- [6] 刘晴, 唐林波, 赵保军. 基于自适应多特征融合的均值迁移红外目标跟踪[J]. 电子与信息学报, 2012, 34(5): 1138-1141.
- [7] 袁广林, 薛模根, 韩裕生, 等. 基于自适应多特征融合的 mean shift 目标跟踪[J]. 计算机研究与发展, 2010, 47(9): 1663-1671.
- [8] COLLINS R T, LIU Y, LEORDEANU M. Online selection of discriminative tracking features[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(10): 1631-1643.
- [9] COMANICIU D, RAMESH V, MEER P. Kernel-based object tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(5): 564-577.
- [10] NING J, ZHANG L, ZHANG D, et al. Robust mean shift tracking with corrected background-weighted histogram[J]. IET Computer Vision, 2012, 6(1): 62-69.
- [11] WENG S K, KUO C M, TU S K. Video object tracking using adaptive Kalman filter[J]. Journal of Visual Communication and Image Representation, 2006, 17(6): 1190-1208.
- [12] 刘海燕, 杨昌玉, 刘春玲, 等. 基于梯度特征和颜色特征的运动目标跟踪算法[J]. 计算机应用, 2012, 32(5): 1265-1268.

(上接第 2579 页)

### 参考文献:

- [1] DONOHO D L. Compressed sensing[J]. IEEE Transactions on Information Theory, 2006, 52(4): 1289-1306.
- [2] CANDÈS E J. Compressive sampling[C]// Proceedings of the 2006 International Congress of Mathematicians. Piscataway, NJ: IEEE Press, 2006: 1433-1452.
- [3] CHANG H, ZHAO G, SHI H, et al. Polynomial rooting based frequency estimation with sub-Nyquist sampling[J]. Journal of Astronautics, 2011, 32(1): 205-209.
- [4] 苏晓园, 练秋生. 基于迭代收缩法和三维复数小波的视频压缩传感重构[J]. 光电工程, 2010, 37(2): 108-112.
- [5] DO T T, CHEN Y, NGUYEN D T, et al. Distributed compressed video sensing[C]// CISS 2009: Proceedings of the 43rd Annual Conference on Information Sciences and Systems. Piscataway, NJ: IEEE Press, 2009: 1393-1396.
- [6] JUNG H, YE J C. Motion estimated and compensated compressed sensing dynamic magnetic resonance imaging: what we can learn from video compression techniques[J]. International Journal of Imaging Systems and Technology, 2010, 20(2): 81-98.
- [7] TROCEN M, TRAMEL E W, FOWLER J E, et al. Compressed-sensing recovery of multiview image and video sequences using signal prediction[J]. Multimedia Tools and Applications, 2013, 66(187): 1-27.
- [8] HUANG H L, ANAMITRA M. Backtracking-based matching pursuit method for sparse signal reconstruction[J]. IEEE Signal Processing Letters, 2011, 18(7): 391-394.
- [9] BARANIUK R G. Compressive sensing [lecture notes][J]. IEEE Signal Processing Magazine, 2007, 24(4): 118-121.
- [10] CANDÈS E J. The restricted isometry property and its implications for compressed sensing[J]. Comptes Rendus Mathématique, 2008, 346(9): 589-592.
- [11] 李翔, 吴国威. 一种适用于 H.264 的基于自适应搜索范围的快速运动估计算法[J]. 中国图象图形学报: A 辑, 2004, 9(4): 471-476.
- [12] MICHAEL W, JASON L, MARCO D, et al. Compressive imaging for video representation and coding[EB/OL]. [2012-10-10]. <http://inside.mines.edu/~mwakin/papers/pcs-camera.pdf>.
- [13] NEEDELL D, TROPP J A. CoSaMP: Iterative signal recovery from incomplete and inaccurate samples[J]. Applied and Computational Harmonic Analysis, 2009, 26(3): 301-321.
- [14] TROPP J A, GILBERT A C. Signal recovery from random measurements via orthogonal matching pursuit[J]. IEEE Transactions on Information Theory, 2007, 53(12): 4655-4666.
- [15] WEI D, OLGICA M. Subspace pursuit for compressive sensing signal reconstruction[J]. IEEE Transactions on Information Theory, 2009, 55(5): 2230-2249.
- [16] RICK C, WOTAO Y. Iteratively reweighted algorithms for compressive sensing[C]// ICASSP 2008: Proceedings of the 2008 IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway, NJ: IEEE Press, 2008: 3869-3872.