分类号	TP39	
UDC _		

密级	
编号	

中国科学院研究生院

硕士学位论文

图像序列中的光流估计与运动分割

马东民

指导教师		普林特 副研究	充员
申请学位级别_	工学硕士	_学科专业名称	模式识别与智能系统
论文提交日期_	2010年5月	_论文答辩日期_	2010年5月
培养单位	中	国科学院自动化	研究所
学位授予单位_		中国科学院研究	生院

答辩委员会主席_刘成林 研究员_

Segmentation and Motion Estimation in Image sequences

Dissertation Submitted to

Institute of Automation, Chinese Academic of Sciences

in partial fulfillment of the requirements

for the degree of

Master of Engineering

by

Ma Dongmin

Pattern Recognition and Intelligent System

Dissertation Supervisor: Veronique Prinet

独创性声明

本人声明所递交的论文是我个人在导师指导下进行的研究工作及取得的研究成 果。尽我所知,除了文中特别加以标注和致谢的地方外,论文中不包含其他人已经发 表或撰写过的研究成果。与我一同工作的同志对本研究所做的任何贡献均已在论文中 作了明确地说明并表示了谢意。

签 名:_____日 期: _____

关于论文使用授权的说明

本人完全了解中国科学院自动化研究所有关保留、使用学位论文的规定,即:中 国科学院自动化研究所有权保留送交论文的复印件,允许论文被查阅和借阅;可以公 布论文的全部或部分内容,可以采用影印、缩印或其他复制手段保存论文。

(保密的论文在解密后应遵守此规定)

签 名:________导师签名:_____日 期: _____

摘要

随着计算机技术、网络技术以及各种摄像设备制造技术的发展,人们可以很容易地获取到海量的视频。这些视频所包含的运动内容丰富多样,它为需要对视频中所含运动进行研究的学科提供了丰富的数据资料,如:运动检测、视频跟踪、流分析等。在海洋学、大气学等需要对视频运动进行研究的领域,研究人员也进行了大量的理论与应用研究。

作为视频分析的一个重要方面,光流估计由于其能直观的表述运动模式和不易受运动 物体外表的干扰等优点受到众多研究者的青睐。现在光流估计方法已经成为运动检测、流 分析等领域的一个重要方法。尽管光流估计方法已经有了极大的改进,能成功应用于许多 方面,但是对于处理一些含形变运动如漩涡、扩散等运动仍然需要进一步改进。

本硕士论文工作就是在这样的背景下展开的,围绕图像序列中的运动估计这一主题进行了研究。本文主要关注于如下两个问题: 1. 建立更具描述性的光流估计模型; 2. 建立含 多运动的图像序列中的光流估计与分割模型。本文在对以往光流估计方法的分析与总结的 基础上,对上述两个问题进行了研究,并提出了两种新的光流估计方法。本文的主要贡献 总结如下:

① 提出了基于统计学习的光流估计方法(第三章)

本方法在Sun的基于统计学习的光流估计方法的基础上,通过引入Radial Basis Function (RBF) 神经网络模型,分别对于关于亮度变化的统计与速度场空间梯度的 统计进行学习,以建立基于统计的光流估计模型。这个方法的优势在于,通过选用 不同的训练集,可以被应用于求解不同类型运动的运动场。

② 提出了图像序列中的光流估计与运动分割方法(第四章)

为了克服已有的方法只能针对特定类型运动进行运动估计,而对超出此运动类型范围的运动不鲁棒的缺点,我们提出了平滑项约束式选择的方法。它是一个适用范围更广的方法。通过建立基本运动表达式集合,并将这些基本运动表达式进行组合,可以对含这些基本运动的图像序列进行运动估计。这种方法的优势在于:1.对视频中的每种运动,通过选择一个最佳的约束来求解更准确的光流场;2.该方法在求解光流场的同时可以对运动进行分割,这对于需要进行运动分割的领域很有用;3.通过选择不同的基本运动集,可以将其拓展到求解更多类型的运动的光流场,如大气图像视频中,可以选择旋度与散度等表达式作为基本运动。

关键词: 统计学习模型,马尔可夫随机场,RBF模型,约束函数选择,运动分割,图像 序列



- ii -

Segmentation and motion Estimation in Image Sequences

Author: Dongmin Ma

Supervisor: Veronique Prinet

Abstract

People can easily acquire huge videos as a result of the development of computer technology, Internet Technology and Manufacturing technology of all kinds of cameras. For it contain abundant information in videos, video analysis technology is given more and more attention, such as video tracking, fluid analysis, 3D reconstruction, etc. In many other fields of science where need image sequences to conduct the researches, such as oceanography, atmospheric science, climatology, medicine, people also have done a lot of theory and application works.

As an important aspect of video analysis, optical flow estimation is becoming popular among researchers, because of its advantages such as the direct description about motions and not easily affected by the appearances of motion objects. Now, optical flow estimation is one important method in fields of motion detection, edge extraction and others. Though great achievements have been done about motion estimation and it can be applied to many aspects, there is still much work to do.

In this master thesis, the theme is to retrieve more accurate motion field in videos. The specific objectives of this thesis are to 1. build more descriptive optical flow estimation model, 2. build model that can retrieve more accurate motion field for multi-motion and segment the motion simultaneously. Based on a general review and analysis of existing methods on motion estimation (Chapter 2) and after studying the above two objectives, we propose two optical flow estimate methods. The main contributions are,

(1) Propose Statistical Modeling of Optical Flow method.

In this method, by introducing Radial Basis Function Neural Network model, the statistical for brightness variant and spatial derivative of velocity field are learnt, and thus establish the statistical model of optical flow. The advantage of this method is it can be applied to more motions by choosing different training set.

(2) Propose Optical Flow Estimation and Segmentation in Image Sequences.

In order to conquer the problem in existing methods which can only retrieve motion field for one specific type of motion, and thus not robust for multimotion, we propose motion regularization selection method. It is a more general method. By propose basis motions and combining them efficiently, the method can choose the best basis motion combination for different motions. And thus can we retrieve more accurate motion field. The contributions of this method are: 1. For different motions, by choosing a most appropriate regularization, can we retrieve more accurate motion field. 2. It can retrieve motion field and segmentation simultaneously. 3. By choosing different basis motion set, it can be extended to detect more types of motions.

Key Words: Statistical modeling, Radial Basis Function Neural Network(RBF), motion selection, motion segmentation, Markov Random Field(MRF), image sequence.

E	录

第一章	绪论	1
1.1	研究意义与背景	1
1.2	问题描述	1
1.3	文章组织结构	2
第二章	研究现状概述	5
2.1	引言	5
2.2	光流法建模分类	5
	2.2.1 差分法建模	6
	2.2.1.1 主要方法介绍	6
	2.2.1.2 刚体运动的光流估计	9
	2.2.1.3 非刚体运动的光流估计	10
	2.2.1.4 基于统计学习的光流估计	12
	2.2.2 其它光流估计建模方法	12
2.3	多尺度方法	13
2.4	Markov Random Field (MRF)模型回顾	13
2.5	优化求解方法	16
-	2.5.1 置信传递方法Belief Propagation (BP)	17
	2.5.2 迭代条件模型Iterated Conditional Modes (ICM)	18
2.6	小结	20
第二音	基于统计学习的光流估计模型	21
31	引言	21
3.2	其于统计学习的光流估计模型	21
0.2	391 数据项建立与学习	21
	3.2.1	24
33	光流估计的名尺度模型	24
3.4	参数学习	25
3.5	立影结果与分析	28
0.0	351	31
	3.5.2 结果分析	31
3.6	小结	33
		00
第四章	图像序列中的光流估计与运动分割	37
4.1	引言	37
4.2	模型思想与描述	37
4.3	数据项	38

4.4	先验知识项	39	
	4.4.1 运动分类	40	
	4.4.1.1 基于角度与幅度的分类	40	
	4.4.2 基于运动分类的平滑项描述	41	
	4.4.2.1 函数向量	41	
	4.4.2.2 选择子函数	41	
	4.4.2.3 参数矩阵	42	
4.5	运动类间的平滑约束	42	
4.6	实现问题	43	
	4.6.1 运动模拟与评估	43	
	4.6.2 BP算法实现	44	
4.7	实验结果与分析	49	
	4.7.1 实验设定	49	
	4.7.2 结果分析	49	
4.8	小结	53	
第五章	结束语	55	
5.1	本文主要工作	55	
5.2	未来工作展望	56	
参考文献 57			
附录A	光流评估方法介绍	i	
附录B	图像数据介绍	iii	
个人简历及攻读硕士学位期间发表的论文 vi			
致谢		ix	

插图目录

1-1	Yosemite图像序列及其矢量光流场	2
1-2	Hydrangea 图像序列及其彩图光流场, 1.2(c)位彩色值到速度值的 映射图。	3
2-1	Lucas-Kanade解的分情况讨论	9
2-2	规则格子(lattice)的邻域(neighborhood)与团块(clique)示意图(图 片来源 [1]),图(a)为四邻域系统,图(b)(c)为团块	14
2-3	消息更新方式。n为结点, M为消息, x为随机变量	19
2-4	BP算法:消息更新算法与执行向量求解视图(图片来源Aggeliki Tsoli)	19
3-1	校正图像与第一帧图像的差值结果,由于实际中差值结果无法 阅读,我们在这里做了简单处理,将结果以4为阈值进行了二值 化,图像中黑色区域为差值大约4,白的区域为差值小于等于4。	25
3-2	差值图像的直方图统计	26
3-3	差值图像的处理后的直方图统计	27
3-4	学习结果。其中点划线为原始数据,实线为学习曲线	27
3-5	速度分量u水平向导数的统计	28
3-6	速度分量v水平向导数的统计	29
3-7	约束项模型学习结果	29
3-8	采用Graph-Cut与TRW-S算法对Yosemite序列求得的光流估计结果以及角度误差。	34
3-9	采用Graph-Cut与TRW-S算法对真实流运动图像序列求解得到的 光流估计。	35
4-1	运动类型1. 所有的4.1(c), 4.1(d), 4.1(e), 4.1(f) 的值均为0。因此, 这个运动场满足运动类型1的约束条件。	45
4-2	运动类型2. 在 4.2(c)中 $ \Delta \rho = 0.05$ 其它结果4.2(d), 4.2(e), 4.2(f)的值均为0。因此,这个运动场满足运动类型2的约束条件。	46
4-3	运动类型3。在 4.3(c)与 4.3(d)中它们的值的绝对值 $\leq 10^{-15}$,可以 视作0。在 4.3(e)中 $\Delta \alpha \parallel = 0.01131$,不可以视作0。因此,这个 运动场满足运动类型3的约束条件。	47
4-4	运动类型4。 $\Delta \rho$ = 0.01414, $\Delta \alpha$ = 0.01131 。其它结果为0。因此,这个运动场满足运动类型4的约束条件。	48
4-5	合成图像序列光流估计与分割结果。在图 4.5(d)中绿色表示垂直 向运动,蓝色表示水平向运动	50
4-6	Hydrangea图像序列光流估计与分割结果。在图 4.6(d)中蓝色代表匀速运动,红色代表单向加速运动,绿色代表角度加速其幅值不变运动,黑色代表角度幅值均加速的运动	51

4-7	Urban图像序列光流估计与分割结果。在图 4.6(d)中蓝色代表匀 速运动,红色代表单向加速运动,绿色代表角度加速其幅值不变 运动,黑色代表角度幅值均加速的运动	52
5-1	Yosemite图像序列及其矢量光流场(源于 [2])	iv
5-2	Hydrangea 图像序列及其彩图光流场(源于middlebury网站)。 图 5.2(b)的速度可以通过 5-3查得。	iv
5-3	彩色速度矢量表。图中不同颜色的坐标即为速度矢量,这样可以使得从彩色图域到速度矢量域的映射转换	v
5-4	流运动图像序列(源于 [3])	V

表格目录

3-1	Yosemite图像序列光流计算结果的量化评估。AAE=Average An-	
	gle Error, STD= Standard Deviation. \ldots \ldots	32
4-1	均角度误差量化评估。	53
4-2	角度误差的标准方差量化评估。	53

第一章 绪 论

1.1 研究意义与背景

随着计算机技术、网络技术以及各种摄像设备制造技术的发展,人们可以 很容易地获取到海量的视频。这些视频所包含的运动内容丰富多样,它为需要 对视频中所含运动进行研究的学科提供了丰富的数据资料如:运动检测、视频 跟踪、流分析等。在海洋学、大气学等需要对视频运动进行研究的领域,研究 人员也进行了大量的理论与应用研究。

作为视频分析的一个重要方面,光流估计由于其能直观的表述运动模式和 不易受运动物体外表的干扰等优点受到众多研究者的青睐。现在光流估计方法 已经成为运动检测、流分析等领域的一个重要方法。尽管光流估计方法已经有 了极大的改进,能成功应用于许多方面,但是对于处理一些含形变运动如漩 涡、扩散等运动仍然需要进一步改进。

1.2 问题描述

在开始这个研究之前,我们首先要定义什么是"光流场"或"运动场"。 光流场(optical flow field)指的是图像灰度模式的表观运动。它是一个二维矢 量场,它包含的信息即是各像点的瞬时运动速度矢量信息。我们的目的就是对 给定图像序列,求取其所含运动的运动场。

为了更清晰的解释什么是光流场,我们给出两个示图如图 1-1与图 1-2。这两组图均给出了光流场的示例,并用不同的方式对光流场进行表示。图 1.1(a)为tsukuba图像序列中的一帧。这个图像序列中的运动分为两部分:从左向右运动的云层与垂直图像向外运动的山体。在图 1.1(b)中给出了这个图像序列的光流场真值。此图中的光流场用速度矢量图表示。从图中可以看出两种不同的运动。图 1.2(a)取自Hydrangea图像序列。其内容也包含两种运动:一种为背景的从左向右的平移运动;另一种为物体的旋转运动。在图 1.2(b)中我们给出了它的彩色光流场图。这是表示光流场的另一种方法。对于图中的每一点的速度,可以通过其彩色值从图 1.2(c)中查找速度值,即与彩色值-速度值映射图中对应的彩色值的坐标即为速度两个分量。

在本论文中,我们给出了一些当前研究方法的优势与缺点,并提出我们自





国 1-1 105emme国家门列及来八重元

己的解决方法。

通过Horn-Schunck以及Lucas-Kanade的工作表明,光流估计是一个病态 (ill-posed)的问题。人们只能通过定义成本函数(cost-function)来求取它的最 优解。在我们的方法中,我们将光流估计问题描述为建立在马尔可夫框架上的 能量最小化问题。这个能量由数据项与平滑项组成。在本文中,我们的主要贡 献在于1.利用学习方法,对马尔可夫模型设置最佳的参数;2.通过采用约束函 数选择的方法,使得我们的方法能对多运动进行更好的约束。

1.3 **文章组织结构**

本论文共分为五章。本章主要介绍本研究的意义与背景、研究问题的描述 以及论文的组织结构与试验平台等。

第二章对当前光流法研究的进展进行综述,并对一些相关的优化求解方法 记性介绍。首先,我们回顾了MRF模型,对其的建立进行了介绍;其次,我们 介绍了光流法的一些已有的方法,并分析了他们的一些优缺点;最后,我们对 两种常用的优化求解方法进行了介绍,尤其是置信传递算法。

第三章主要介绍我们提出的基于统计学习的光流估计模型。在这一章中, 分别对模型的数据项与平滑项的建立方法进行了详细说明和分析,在最后给出 了实验结果分析并做章节总结。

第四部分主要介绍我们提出的图像序列中的光流估计与运动分割模型。我



(a) Hydrangea图像序列





们给出了我们建立此模型的目的与意义。在这个模型中,我们重点介绍了通过 平滑项选择方法对不同运动进行不同的约束的方法。我们对常见运动进行了分 析,并对典型运动进行了提取。根据这些提取的典型运动我们建立了函数向 量。基于这个函数向量我们生成了一些典型运动的合成图像序列以用于测试。 由于可以采用平滑函数选择的方法,在这个模型中我们还能对图像序列中的运 动进行分割。在最后我们给出了初步的实验结果并进行分析。在总结中,我们 对模型需要改进的地方进行了讨论。

第五部分小结。对本文主要工作和贡献进行了总结,并对未来此方向的工 作进行了展望。

第二章 研究现状概述

2.1 引言

本章主要对本研究主题内国内外已有的主要方法进行综述,并对后面章节 将要用到的一些基本的概念与方法进行介绍。

光流估计方法主要包括光流模型建模与光流求解两个步骤。由于基于马尔 可夫随机场与光流求解方法本文中要经常使用,在本章中也将对尔可夫随机场 模型进行简要回顾。

本章的组织如下:2.1节简要概述本章内容;2.2节对已有的光流求解方 法进行综述;2.3对多尺度方法进行简要介绍;2.4节对尔可夫随机场模型进 行回顾;2.4节对两种主要的优化求解方法:置信传递(Belief Propagation) 与ICM(Iterated Conditional Mode)进行介绍。最后在2.5节中给出本章小结。

2.2 光流法建模分类

光流场(optical flow field)指的是图像灰度模式的表观运动,是空间运动物体在观测成像面上的像素运动的瞬时速度。它是一个二维矢量场,它包含的信息即是各像点的瞬时运动速度矢量信息。我们的目的就是对给定图像序列, 求取其所含运动的运动场。研究光流场的目的就是为了从序列图像中近似计算运动场。在光流场估计中,人们假设图像中运动点的亮度不随时间变化即,

$$\frac{dI(x(t), y(t), t)}{dt} = 0$$
(2-1)

在实际研究中,针对视频或者图像序列中含有的不同运动物体的物理属 性(如刚体、非刚体),人们往往采用不同的方法进行研究,这样就可以满 足不同物体所特有的运动情况。而这种"不同的方法"指的是研究者主要根据 所研究图像序列中的运动物体的物理属性和运动属性,通过数学方法或者物 理方法进行建模,以达到准确描述这些运动的目的。根据 [2]光流法主要分 为四类:差分法(differential)、区域匹配法(matching)、基于能量的方法 (energy-based)或者根据 [4]频域法(frequency-based method)、基于相的方 法(phase-based)。由于在实际研究中,我们主要采用的是差分法建模,下面 将对差分法建模做主要回顾,其余方法进行简单介绍。

2.2.1 差分法建模

基于差分的光流估计方法主要通过对图像序列进行时间、空间求导来求解 出运动物体的速度估计 [2],

 $I(x + \Delta x, y + \Delta y, t + \Delta t) - I(x, y, t) = \frac{\partial I}{\partial x} \Delta x + \frac{\partial I}{\partial y} \Delta y + \frac{\partial I}{\partial t} \Delta t = 0 \quad (2-2)$ 其中, I为图像数据, x, y, t分别为水平坐标, 垂直坐标以及时间。它只能用 于检测小位移运动($\Delta x, \Delta y$)。

2.2.1.1 主要方法介绍

在这一节我们主要介绍光流估计的两个先创性的工作。一个是Horn和Schunck光流方法 [5];另一个是Lucas-Kanade光流方法 [6]。

• Horn和Schunck光流方法 [5]

1981年,Horn和Schunck创造性地将二维速度场与图像灰度相联系,引入 了光流约束方程,并通过为光流约束方程添加约束项最终将问题解决,得 到光流估计的基本方法。此后许多的研究者都按照这种方法进行他们的研 究。

基于Horn [5]提出的三个假设条件(即亮度不变假设):

- 假设被拍摄的物体的面为平面

- 假设物体的光照均匀
- 假设物体表面光反射是平滑的且空间连续

得到,

$$\nabla I \cdot \mathbf{w} + I_t = 0 \tag{2-3}$$

其中**w** = (u, v)为速度矢量, I为图像像素值, $\nabla I = \left(\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}\right)$, $I_t = \frac{\partial I}{\partial t}$ 。 很显然,这是一个开孔问题 [4],由光流约束方程:该方程含有两个未 知u, v,只用使用附加的约束条件才能对其进行求解。

为了解决这个问题,Horn提出了一阶平滑约束:物体邻近点有相似的速度,来为光流约束方程添加约束项即,最小化光流速度梯度的幅度的平方,其表达式为:

$$\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2 \qquad \forall x, \forall y$$
(2-4)

其中u = u(x, y), v = v(x, y)综合光流约束方程可得出:

$$\min_{(\mathbf{u},\mathbf{v})} \int_{\Omega} (\nabla I \cdot \mathbf{w} + I_t)^2 + \lambda^2 (\|\nabla u\|_2^2 + \|\nabla v\|_2^2)^2 d\mathbf{x}$$
(2-5)
其中: $\mathbf{x} = (x, y)$ 为图像坐标, λ 为权重系数, Ω 为整个图像域。

这样光流场的计算问题就转换为一个离散的变分问题 [7]。因此通过下面 迭代方法求解得:

$$u^{n+1} = \overline{u}^n - I_x [I_x \overline{u}^n + I_y \overline{v}^n + I_t] / (\alpha^2 + I_x^2 + I_y^2)$$
(2-6)

$$v^{n+1} = \overline{v}^n - I_y [I_x \overline{u}^n + I_y \overline{v}^n + I_t] / (\alpha^2 + I_x^2 + I_y^2)$$

$$(2-7)$$

其中: ū, v为速度u, v的局部平均, n为迭代次数。

Horn基于其提出的三个条件假设对光流估计进行了很好的求解,但由于在现实世界中,许多物体的运动并不满足这三个条件,比如物体遮挡、物体形变等,另外,由于此方法只能对位移不超过一个像素的运动进行光流求解,需要通过寻找别的方法比如多分辨率分级的方法对其进行扩展,以适用于大位移运动。针对这些问题出现了一系列的文章,人们提出了很多很好的解决方法,部分方法将会在文中介绍(见节2.2.1-2.2.2)。

• Lucas-Kanade光流方法 [6]

Lucas-Kanade光流法是20世纪80年代提出的方法。它是经常被用到的另一种进行光流求解的方法。为了解决差分光流法的问题,

$$\nabla I \cdot \mathbf{w} + I_t = 0 \tag{2-8}$$

在Lucas-Kanade方法中,以帧与帧之间的时间很短、物体的运动很不明显 为前提,通过假设在图像上的局部小范围内(如小窗口)邻近像素点的光流 速度相等,他们通过这些小范围的像素点得到一个方程组,

$$I_{x1}u + I_{y1}v = -I_{t1}$$

$$I_{x2}u + I_{y2}v = -I_{t2}$$

$$I_{x3}u + I_{y3}v = -I_{t3}$$

$$\cdots$$

$$I_{xn}u + I_{yn}v = -I_{tn}$$

其中n是小范围窗口的像素点个数, *I_{xi}、I_{yi}和I_{ti}为小范围窗口内第i个像素* 点的水平方向、垂直方向的导数和时间导数。 对上面所述式子进行整理可以得到,

$$\begin{bmatrix} I_{x1} & I_{y1} \\ I_{x2} & I_{y2} \\ \cdots \\ I_{xn} & I_{yn} \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} I_{t1} \\ I_{t2} \\ \cdots \\ I_{tn} \end{bmatrix}$$
(2-9)

令
$$A = \begin{bmatrix} I_{x1} & I_{y1} \\ I_{x2} & I_{y2} \\ \dots \\ I_{xn} & I_{yn} \end{bmatrix}$$
, $b = -\begin{bmatrix} I_{t1} \\ I_{t2} \\ \dots \\ I_{tn} \end{bmatrix}$, $\overrightarrow{w} = \begin{bmatrix} u \\ v \end{bmatrix}$, 上式可以写做,

$$A\overrightarrow{w} = b \tag{2-10}$$

这个方程组自然是个超定方程,也就是说方程组内有冗余,首先对上式进 行改写,得到要求解得目标方程,

$$\min_{\overrightarrow{w}} \|A\overrightarrow{w} - b\|^2 \tag{2-11}$$

其中 $\vec{w} = \vec{w}(x,y)$ 。

可以通过最小二乘方法对上式进行求解:

$$\overrightarrow{w} = (AA^T)^{-1}A^Tb \tag{2-12}$$

其中, $AA^{T} = \begin{bmatrix} \sum I_{x}^{2} & \sum I_{x}I_{y} \\ \sum I_{x}I_{y} & \sum I_{y}^{2} \end{bmatrix}$

对这个方程的解我们讨论如下,

- 处于亮度相同位置时如图 2.1(b),此时有 $AA^T = 0$,也是不可逆。

- 处于纹理位置时如图 2.1(c),此时AA^T可逆。



(b) 亮度相同位置 图 2-1 Lucas-Kanade解的分情况讨论

(c) 纹理位置

从上面的讨论可以得知这个算法的不足在于它不能产生一个密度很高的流 向量,比如在运动的边缘和亮度相同区域中的微小移动方面流信息会很快 的褪去。它的优点在于有噪声存在的鲁棒性还是可以的。

刚体运动的光流估计 2.2.1.2

由于光流约束方程是一个开孔问题,无法对光流进行求取,于是研究者寻 求通过对光流速度矢量添加约束并综合光流约束方程来求解光流。这种约束方 法主要分为两类 [2], 一类是全局约束, 另一类是局部约束。

• 全局约束 [4,5,7,8,9]

Nagel [8,9]通过采用二阶导数方法来求取光流。在此方法中, Nagel提出了 面向平滑的约束(oriented-smoothness constraint),在这种约束中没有对梯 度陡峭的边缘进行平滑约束以解决遮挡问题。但是如果开孔问题在局部邻 域中严重,由于数字差分的敏感性(比如噪声)原因,二阶导数方法通常不 能准确的确定光流速度v的切向部分 [2]。因此二阶导数方法对速度估计的 准确性不如一阶导数方法好。

• 局部约束 [4,10,11,6,12,13,14,15]

[11]用一个模型通过最小二乘或Hough变换来拟合每个像素m的邻域的 光流值 [4]。根据 [11,6,12,13,14,15], Barron [2]实现了对窗口领域进行 一阶平滑约束的加权最小二乘拟合光流估计方法。由于此方法需要连 续15帧才能进行运算, [10]使用IIR(Infinite impulse response)迭代滤波器

替代了FIR(finite impulse response)滤波器,这种方法只需要连续三帧图像,因此计算只延迟两到三帧。

2.2.1.3 非刚体运动的光流估计

对于另一类物体运动——流运动,如大气云图、灰尘、沙尘暴等运动中带 有严重形变如扩散、聚集、形状拉伸、涡流等,另外对于大气云图,云层在图 像中的亮度和云层的高度有关,不同时间同一云区由于高度不同在图像中的亮 度也不同。这样Horn提出的两个假设——亮度不变性和一阶平滑性——已经不 再适用。为了解决这些问题,必须对Horn的假设进行修改或者采用新的方法以 解决问题。在这一节我们我们对关于非刚体运动光流估计的部分方法进行介 绍。

Corpetti [16]提出了新的基于最小化的方法。该方法的提出是为了解决流运动中呈现出的大量的时空失真问题。在此方法中其数据项是基于连续方程导出的,相对于亮度不变假设这个约束项能更好的与真实的情况相符合,

$$I(\mathbf{x} + \mathbf{d}(\mathbf{x}), t + \Delta t) = I(\mathbf{x}, t)exp(-div\mathbf{d}(\mathbf{x}))$$
(2-13)

根据上式,亮度信息可以由*exp*(-*div*d(x))进行调节。当散度为正时,亮度 下降;反之则相反。

对于约束项,他们基于div-curl公式,引入了带附加常量场的鲁棒性二阶约 束,这样就能更好地捕捉未知流中的扩散运动和旋转运动,约束项为:

 $\int_{\Omega} |div\mathbf{d}(\mathbf{x}) - \xi|^2 + \lambda f_2(|\nabla\xi|) + \int_{\Omega} |curl\mathbf{d}(\mathbf{x}) - \zeta|^2 + \lambda f_2(|\nabla\zeta|) d\mathbf{x} \qquad (2-14)$ $| \xi \mathbf{\psi} : div\mathbf{d} = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}, curl\mathbf{d} = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}$

这样通过用以上数据项和约束项来替代Horn的两个假设,这样就能更好的 适用于流运动的特殊运动性质。对于非刚体运动,由于该方法采用的模型更接 近真实的物体运动情况,因此能很好的解决Horn中遇到的困难,能更有效的解 决非刚体运动中的形变问题尤其是含有漩涡类的运动。由于本方法是针对非刚 体运动,需要整合其他物理模型、先验知识以适应更多的实际里情况。

Bereziat [17]采用了体积不变假设的方法,此方法为了克服灰度不变假设在 非刚体运动中的不适用性,指出解决非刚体运动光流求解的一个更好的办法是 对运动物体整体进行考虑而不是只针对单像素进行光流求解。因此,通过假设 物体的总体灰度或者灰度总和是暂时不变的,他们建立了体积约束方程。它能 更好的适用于非刚体物体运动。其体积约束方程为:

$$\nabla I \cdot \mathbf{v} + I_t + I \cdot div(\mathbf{v}) = 0 \tag{2-15}$$

同时作者还指出,在存在多个运动物体的图像序列中,由于各个物体运动 方式不一样,即有的物体运动方式满足体积约束,有的物体运动满足灰度值不 变约束,可以通过创建一种混合模型,即先对图像进行预分割,在对不同区域 采用不同方法进行处理,此种处理模型即为混合模型。根据作者给出的实验结 果,这个方法能很好的应用于卫星云图光流的求解,但是在这类图像中由于云 层的高低随时间变化(影响图像的亮度)、云层的扩散聚集运动为三维运动, 很难保持图像的亮度体积不变假设,因此这种方法仍需要进一步改进。

Bereziat [18]采用将物体整体视为研究对象的方法,通过约束物体运动中的整体亮度来求解光流。在此方法中,他们采用了一种双模型方法(bi-model)即混 合模型方法,同时包含了亮度不变约束与总体亮度不变约束。对于约束项他们 通过采用仿射先验模型来对运动方程进行过约束。这种模型适用于多种情况尤 其是红外气象数据图像。对于双模型或者混合模型,他们通过将亮度不变约束 与总体亮度不变约束两种约束结合起来得到一个混合模型。此模型可以为运动 物体和静止背景提供不同的约束。此模型可以写为,

 $\forall (x,y) \in \Omega, \nabla I \cdot \mathbf{v} + I_t + I \cdot div(\mathbf{v})\Psi_O = 0$ (2-16) 其中, Ω为图像空间域, O为云区, 当 $(x,y) \in O$ 时 $\Psi_O = 1$ 否则 $\Psi_O = 0$ 。

此时可以看出当 $(x, y) \in O$ 时即 $\Psi_O = 1$,混合模型变为体积不变方程,可用 于求解云层区域运动;当 $(x, y) \notin O$ 时即 $\Psi_O = 0$,混合模型变为亮度不变约束方 程,可用于求解云层区域外的运动,这样通过对不同运动区域使用不同的模型 可以更好的求解出光流。

基于差分光流方程 [19]提出了一种鲁棒性多尺度技术进行图像序列的速度 场估计。在此方法中作者提出了一种新的混合多尺度多网格方法,即将高斯金 字塔下采样分解与多网格方法结合起来。这种方法使得计算大位移运动成为可 能,同时这种方法也能有效的改善计算时间。

[20]提出了一种新的方法用于求解多结构物体的全局运动和局部运动。虽 然最近关于非刚体物体运动光流求解已经出现了很多有效的方法,但是对于 带有亚结构运动物体运动,这些方法却不能给出有效的求解,其原因在于物 体的形变运动包含了全局性的和局部性的。文章 [20]找出了一种有效的公式与 计算框架以对非紧致性非刚体物体运动进行全局和局部的光流求解。作者采用的方法主要有两点:一个是亮度不变约束一般化公式(Generalized Constant intensity Constraint),另一个是多尺度分析框架的下采样优化方法。其亮度不变约束一般化公式为:

$$\nabla I \cdot (\mathbf{v}_g + \mathbf{v}_l) + \frac{\delta I}{\delta t} = 0, \mathbf{v}_l \ll \mathbf{v}_g \tag{2-17}$$

其中: v_g 为全局运动速度, v_l 为局部运动速度。通过本文给出的结果,可以看出本方法对含有全局运动与局部运动的非刚体运动能进行很好的光流求解。

2.2.1.4 基于统计学习的光流估计

通过学习方法进行光流建模 [21,22,23]也是一个重要的方向。与采用标准启 发式公式的方法不同,Sun [22]建立了一个运动的概率模型,并通过利用已有 的含有光流真值的图像序列对数据项与约束项进行独立的训练学习。通过利用 光流真值与双线性差值法,作者将每组中第二帧图像校正(warp)到第一帧并与 第一帧图像求差,以此作者可以获知自然图像序列中一致性假设是如何被违反 的,并且可以得到关于一个亮度不一致的概率模型。

2.2.2 其它光流估计建模方法

他光流估计方法主要包含区域匹配法(matching)、基于频域法 (frequency-based method)、基于相的方法(phase-based)等。这里以模 板匹配方法做一下简述。

在差分光流估计中,由于图像序列或者视频存在噪声,因此准确的求取数 字差分变得不现实。因此人们很自然的找到了另一种方法即基于区域的匹配方 法。这种方法通过找到两个时刻图像的最佳匹配区域来求取光流 [4]。为了找到 最佳匹配,可以采用最大化相似度如归一化的相互关系数,进行最大化;光流 强度场的平方和,进行最小化等。

Anandan [24]采用了基于拉普拉斯金字塔和从粗到细的SSD(sum-of-squared difference)匹配策略。其框架有三部分构成: 空域分解,它可以依据尺度对图像的强度变化分开;匹配策略;控制策略,用来控制不同尺度的度量处理并合并他们的运算结果。Brad [25]采用将联合快匹配算法与和最佳候选快搜寻与向量中位数约束项合并的方法,以解决非刚体运动。

虽然现在人们已经提出了很多新的理论方法,其中包括对彩色图像的光流

求解方法 [26,27,28,29,30]的研究,但由于时空序列图像自身的一些失真 [16] (如变形)原因,对真实的图像序列进行准确的运动估计仍然是一个比较困难 的问题。

2.3 多尺度方法

在光流估计中,由于图像序列中经常含有大位移运动(位移大于1个像素的运动),这使得建立的光流模型很难准确的检测到这些大位移运动,如在求解含大位移运动的光流场时,很容易陷入一个局部能量最小化,而不能达到全局最优[31],这样求取的光流场显然不准确。可以通过采用多尺度方法对这个问题加以解决[31],即采用从粗到细(coarse-fine)的方法[31,32,33,34]。在这个方法中,人们通过一些多尺度方法如高斯金字塔[35,36]将图像序列从低尺度到高尺度分级。为了达到全局最优,首先对低尺度的图像序列进行光流求解。这是由于它尺度小,所含运动位移小,更易于达到全局最优。其次,将低尺度求取的光流场插值到其相邻的更高一级图像大小,并用插值后的光流场对这一级的像进行校正(warped)。由于进行了校正,校正后的图像序列的运动位移会变成小的相对位移。这样就可以很容易的对这一级图像序列求取使得全局最优的光流场。由于这一级中,所求的光流场为相对位移,需要与插值所得的光流场进行迭加才能得到这一级图像序列的真正位移。通过这种从低尺度到高尺度、逐级校正的方法,最终解决了光流法求取含大位移运动的图像序列的光流场的问题。

2.4 Markov Random Field (MRF)模型回顾

MRF随机场理论是分析物理现象依赖关系的概率理论的一个分支 [1]。由于 它能刻画随机变量间的相互影响,MRF理论为时间、空间相关的过程建模提供 了一个方便有效的理论 [37]。因此,它被广泛应用于计算机视觉与图像处理问 题 [38]如图像修复、图像分割、纹理分析、光流估计、边缘检测等领域。本节 将对MRF模型进行回顾。

为了介绍MRF,我们首先对一些需要用到的符号做一些约定,然后以图像问题为例简单介绍MRF理论。很多视觉问题都可以被视为标签问题。在这些问题中,它们的解便是分配给图像像素或者特征的一系列标签 [1]。标签问题由基元集与标签集组成。我们的研究对象由很多基元组成,(如每一个像素可视为

一个基元)。假设有m个基元,对基元集我们用如下方式表示,

$$S = \{1, 2, 3, \dots, m\}$$
(2-18)

基元通过邻域系统与其它基元联系起来,对基元的邻域系统表示如下,

$$\mathcal{N} = \{\mathcal{N}_i | \forall i \in S\} \tag{2-19}$$

其中, N_i为基元i的邻域。这些邻域关系有如下属性,

- 一个基元不属于它自己的邻域: $i \notin N_i$;
- 邻域关系可以互换: $i \in \mathcal{N}_{i'} \iff i' \in \mathcal{N}_{i}$ 。

对于规则的格子组成的集合*S*,*i*的邻域可以定义为在半径为*r*的圆内的邻近基元,

$$\mathcal{N}_{i} = \{i' \in S | \|x_{i'} - x_{i}\|^{2} \leqslant r, i' \neq i\}$$
(2-20)

其中,*x_i*为基元*i*的空间坐标; || ||表示两个坐标的欧拉距离; *r*取整数。通过上面方式,所有点的邻域构成了邻域系统。





 (S, \mathcal{N}) 构成了我们常见的关系图。其中,*S*包含结点,*N*根据邻域关系确定 结点之间的连接。在关系图上,我们再进一步定义团块(clique),它是*S*的一 个子集。它可以由单个基元组成如 $c = \{i\}$ 2-2(b),也可以由邻域内一对基元组 成如 $c = \{i, i'\}$ 2-2(c),或者由邻域内三个不同基元组成,等等。我们用如下方 式对它们进行表示,

$$C_1 = \{i | i \in S\} \tag{2-21}$$

 $C_2 = \{\{i, i'\} | i' \in \mathcal{N}_i, i \in S\}$ 且它们互相为其它结点的邻近结点} (2-22) 图(*S*, *N*)上的所有团块可以通过*C* = *C*₁ \cup *C*₁ \cup *C*₁ \cdots 来表示。邻域与团 块的示意图可以参见图2-2。在图中,2-2(a)为一个一阶邻域系统也叫4邻 域系统。2-2(b)与2-2(c)分别为为一阶邻域系统的单个基元团块、水平与 垂直成对基元团块。对于结点*i* = 5, $\mathcal{N}_i = \{2, 4, 6, 8\}, C_1 = \{5\}, C_2 = \{(5, 2), (5, 4), (5, 6), (5, 8)\}$ 。

在MRF理论中, 图像被视为定义在*S*集合上的随机变量集*X* = $\{X_1, X_2, \ldots, X_m\}$ 。在这个集合中,每个随机变量 X_i 取值 $x_i \Box x_i \in L$, *L*为给定标签集合。我们称集合*X*即为随机场。我们用 $X_i = x_i$ 表示随机变量 X_i 取 x_i 的事件,用 $(X_1 = x_1, X_2 = x_2, \ldots, X_m = x_m)$ 表示联合事件。为了简化起见,联合事件可以简写为X = x,其中 $x = x_1, x_2, \ldots, x_m$ 对应随机场的一个实现。对于给定标签集*L*,随机变量 X_i 取值 x_i 的概率用 $P(X_i = x_i)$ 表示,简写为 $P(f_i)$ 。联合概率用 $P(X_1 = x_1, X_2 = x_2, \ldots, X_m = x_m)$ 表示,简写为P(x)。那么,当X当且仅当满足下面两个条件时,它才成为定义在 (S, \mathcal{N}) 上的马尔可夫随机场,

$$P(x) > 0, \forall x \in L^m \tag{2-23}$$

$$P(x_i|x_{S-\{i\}}) = P(x_i|x_{\mathcal{N}_i})$$
(2-24)

其中, L^m 为所有可能的f组成的空间, $S - \{i\}$ 为集合之差, $x_{S-\{i\}}$ 表示集 合 $S - \{i\}$ 上的标签集,并且 $x_{N_i} = \{x_{i'} | i' \in N_i\}$ 。第一个公式保证了标签集内 的任何标签的概率都为正。第二个公式保证了马尔可夫性,即描述了集合F的 局部特性。任意位置的标签只依赖于其邻域基元,或者说只有邻域标签之间才 能进行直接的相互作用,但是在实际中很难确定满足这两个条件的概率 [39]。 通过Hammersley等人的工作,他们证明了吉布斯随机场(Gibbs Random Field, GRF)与马尔可夫随机场的等价性。从而可以使用Gibbs分布来求解马尔可夫 随机场的概率分布 [1,38,39,40,41]。下面对吉布斯随机场进行简单介绍。

对于 (S, \mathcal{N}) 上随机变量集F,当其满足吉布斯分布时,被称为吉布斯随机场。吉布斯分布的表达式如下,

$$P(x) = \frac{1}{Z} e^{-\frac{1}{T}U(x)}$$
(2-25)

其中, $Z = \sum_{x \in L^m} e^{-\frac{1}{T}U(x)}$ 为归一化因子, T为温度常数,通常取值为1, U(x)为能量函数,它是所有团块的能量和即 $U(x) = \sum_{c \in C} V_c(x)$, C为邻域系统所包含的所有团块的集合, $V_c(\cdot)$ 是定义在团块上的势函数,它只依赖于团块c的局部取

值。

为了表述方便, 吉布斯分布有时使用如下方式进行表示,

$$U(x) = \sum_{\{i\}\in C_1} V_1(x_i) + \sum_{\{i,i'\}\in C_2} V_2(x_i, x_{i'}) + \cdots$$
(2-26)

对于两阶团块系统已经足够,在这种情况下能量可以表述如下,

$$U(x) = \sum_{\{i\} \in C_1} V_1(x_i) + \sum_{\{i,i'\} \in C_2} V_2(x_i, x_{i'})$$
(2-27)

根据Hammersley-Clifford定理 [41],一个随机场是关于邻域系统的MRF, 当且仅当这个随机场是关于邻域系统的Gibbs分布。关于邻域系统N的马尔可夫 随机场与吉布斯分布的等价形式可以表示为 [37]。

$$P(x) = \frac{1}{Z} \exp\left(-\frac{1}{T} \left(\sum_{\{i\}\in C_1} V_1(x_i) + \sum_{\{i,i'\}\in C_2} V_2(x_i, x_{i'})\right)\right)$$
(2-28)

通过上式解决了求MRF中概率分布的难题,使对MRF的研究转化为对势函数 $V_c(\cdot)$ 的研究 [39],即最大化P(x)等价于最小化总能量 $U = \sum_{\{i\}\in C_1} V_1(x_i) + \sum_{\{i,i'\}\in C_2} V_2(x_i, x_{i'})$ 。为了反映成对基元之间的相互作用的关系,在能量中加入了 控制平滑强度的常数系数 λ ,这样得到了在MAP框架下的最优估计值 [37],

$$\widehat{x} = \arg\min_{x} \left(\sum_{\{i\} \in C_1} V_1(x_i) + \lambda \sum_{\{i,i'\} \in C_2} V_2(x_i, x_{i'}) \right)$$
(2-29)

2.5 优化求解方法

基于MRF框架的光流模型最后被转化为求解在MAP框架下的最优估计值 [37],

$$\widehat{x} = \arg\min_{x \in L^m} (\sum_{(i,j)} V(x_i, x_j) + \sum_i D(x_i, y_i))$$
(2-30)

其中, $V(x_i, x_j)$ 与 $D(x_i, y_i)$ 分别为平滑项能量函数与数据项能量函数。在这里,为了表述更具描述性,我们对上节最后式子2-28做了改写。 $V(x_i, x_j)$ 与式2-28中 $V_2(x_i, x_{i'})$)相同, $D(x_i, y_i)$ 与2-28中 $V_1(x_i)$ 相同,添加 y_i 是为了表示这一项与给定数据的关系。

通过优化方法如Belief Propagation (BP) [42, 43, 44, 45, 46, 47]、Iterated Conditional Modes (ICM) [48]、simulated annealing [49]、Graph Cuts [50, 51]、Tree-Reweighted Max-Product Message Passing (TRW) [52]、sequential tree-reweighted algorithm (TRW-S) [53]等可以对模型进行求解,并且在 [54,55,56]中给出了这些方法的性能比较。本节将对后续章节对Belief Propagation (BP)和Iterated Conditional Modes (ICM)进行简单介绍。

2.5.1 置信传递方法Belief Propagation (BP)

对于任意的贝叶斯网络,计算其结点的边缘分布是一个NP hard问题。为 了解决这个问题,Pearl于1988年提出了置信传递方法Belief Propagation (在下 文中用缩写BP表示)。它是一种基于局部消息传递来解决推理问题的有效的方 法,即一个结点通过收集其邻域结点传向它的消息(不包括它要传向的结点) 并考虑它与要传向的结点之间的约束关系确定要传递的消息值。下面对其进行 简单介绍。

MRF的求解是一个NP问题,应用BP算法可以对其进行近似求解。之所以 是近似求解是因为,MRF为有环图。BP算法对无环图可以通过有限次迭代求取 精确解。但是对于有环图,由于可能会出现震荡、发散等不收敛的情况,只能 通过对收敛条件进行设定或者通过指定迭代次数来获取近似解。

本节对应用BP算法求取MRF最大似然估计的算法进行例举介绍。

在求解上述能量的过程中,BP算法包含两部分:消息更新算法与最终结 果获取算法。在这里,为了更清楚的表述结果获取算法,我们将其分解为两部 分:一部分为最终的置信向量的求解;另一部分为最终结果的求解或者为标签 的求解。下面给出消息更新算法与结果获取算法:

• 消息更新算法。

$$m_{ij}^{t}(x_j) = \min_{x_i} (D(x_i, y_i) + V(x_i, x_j) + \sum_{k \in \mathcal{N}(i)/j} m_{ki}^{t-1}(x_i))$$
(2-31)

*t*为迭代次数,通过上式并参考图2.4(a),当更新任意一个结点*i*传递到其领域结点*j*的消息时,*i*结点首先收集除*j*点外的领域结点传向它的消息,并考虑到两个能量函数*D*(·),*V*(·)来进行消息更新。

• 置信向量估计算法

$$b(x_j) = D(x_j, y_j) + \sum_{q \in \mathcal{N}(j)} m_{q \to j}(x_j)$$
(2-32)

当消息更新迭代停止,需要求取各未知变量结点的置信向量以最后求取 各节点标签。对于任意结点j求取置信向量的方法是2.4(b):首先收集其领 域结点传递给它的消息并考虑数据项能量函数*D*(·),来求取最终的置信向量。

• 最终结果求解算法

$$\widehat{x}_j = \arg\min_{x \in I^m} b(x_j) \tag{2-33}$$

求取置信向量后,对每一个结点求使得*b*(*x_j*)最大的标签*x_j*。这一步为BP算法的最终结果。

BP算法是一个迭代算法,在迭代过程中需要对每一个结点到其邻域结点的 消息进行更新,因此需要采用某种消息更新机制。消息更新机制的不同,直接 影响到消息传递效率的不同,下面对两种常用的消息更新机制进行介绍。

- 同步消息更新机制,即对每一个节点依次更新其4个领域节点,再更新下 一节点的消息。
- 单方向消息更新机制,即每次消息单方向从一个节点传递到相应方向的下 一个领域节点。

最后我们对BP算法做如下归纳:

- BP为迭代算法。其收敛条件有全局能量收敛与固定迭代次数两种。
- 在迭代运算中,消息从一个节点传递到其领域节点,以此对消息进行更新,如图2-3所示。
- 对于无环图, BP可以通过有限次迭代获得准确解; 对于有环图, BP可以通过设定消息更新次数或者设定收敛阈值获取近似解。

2.5.2 迭代条件模型Iterated Conditional Modes (ICM)

ICM(Iterated Conditional Mode)是另一种用于求解MRF的方法。由于ICM算法计算量小,没有复杂的操作,很适合通过计算机实现,因而使用较多,下面对这种算法进行介绍。

ICM算法是由Besag [48]提出,用于计算最大后验分布。ICM算法为迭代算法,采用对逐个像素进行扫描的方式,根据条件分布逐点对节点进行估计。



图 2-3 消息更新方式。n为结点, M为消息, x为随机变量





假定二维随机场的大小为*M* × *N*, ICM算法步骤 [37]如Algorithm 1所示。 它具有两个特点1.为迭代算法; 2.用于求取最优值。它通过对每一个结点求取不 同状态值时的局部能量,并以最小能量对应的状态值作为此结点此时的状态。 通过对每一结点依次进行如此运算达到对结点状态的不断更新。

标准的ICM算法规定每一个结点的状态值不再改变时,才是达到收敛。但 是实验结果表明这一规定太苛刻,往往很难达到此要求,因此,可以选用其它 标准作为收敛的条件:

① 通过设定固定迭代次数k。当整个图像迭代次数达到k时,终止迭代。

Algorithm 1 ICM算法

1: 对要随机场的每一结点进行初始化,初始化最大迭代次数k。

2: 从节点0到节点MN,计算每一个结点在取不同状态值时的局部能量。

3: 对每个结点,选择能够使局部能量最小的状态作为该点的当前状态值。

4: 重复步骤2 和步骤3, 直到收敛或者达到最大迭代次数。

② 使用全局能量的相对变化量作为收敛依据。计算每次迭代结果的能量并与 上次迭代计算结果的能量比较,如果在设定阈值内,则停止迭代。

事实上,上述两种标准可以同时使用,即同时给算法设定能量收敛阈值与最大 迭代次数。当在最大迭代次数内能量收敛时,迭代停止;当超过最大迭代次 数,能量仍未收敛时,同样也停止迭代运算。这样可以防止由于能量出现震动 或发散而引起的死循环。

2.6 小结

通过对研究课题领域内的发展与研究状况的总结,可以得出:由于图像序 列中所含运动物体的物理属性与运动属性具有多样性,很难找到一种通用的算 法去解决问题。同时由于同一图像序列中含有几种运动,很难找到一种通用的 约束去惩罚这些不同的运动。因此人们普遍采用的研究方法是,对特定图像序 列集建立专门的光流模型,以期解决问题。

另外,由于光流估计要求求解图像中每一点的速度,即要求解的未知变量 为图像像素大小个。这就要求不但要建立合适的光流模型,还要有快速有效的 优化求解方法去求解模型。

针对我们研究的问题,我们将在下面章节研究并提出我们的解决方法。
第三章 基于统计学习的光流估计模型

3.1 引言

自1981年Horn [5]提出光流的三个假设(1. 假设被拍摄的物体的面为平面; 2. 假设物体的光照均匀;3. 假设物体表面光反射是平滑的且空间连续)以及两个平 滑约束假设(1. 物体邻近点有相似的速度;2. 亮度模式的速度场几乎处处平滑), 并在此基础上进一步提出了光流估计模型。此后,很多研究者在这个领域上的 工作是基于Horn的这一理论进行的。

目前,采用基于统计学习的方法的研究工作还不是太多。受 [22,23,21]的 启发,在本章我们提出基于MRF框架的统计学习方法。这个方法使得我们一 方面可以利用统计学习方法对已有训练集进行学习,这样可以使建立的模型对 图像序列中的运动更具描述性,如通过统计学习可以更好的反映遮挡,边缘等 自然图像序列中固有特性是如何违反Horn假设的 [22];另一方面,由于将统基 于MRF模型,我们可以利用已有且有效的优化方法对模型进行求解。

在本方法中,我们引入了基于RBF(Radial Basis Function)神经网络 [57, 58]的学习模型。我们采用了高斯金字塔下采样方法用以检测图像序列中的大位移运动。

本章的内容组织如下: 3.1节引言,列出章节组织; 3.2节提出基于统计学 习的光流估计模型,在这一节中,将分别详细阐述数据项与约束项的建立与学 习。3.3节给出光流估计的多尺度模型,这个模型使得大位移运动可以被求解出 来。3.4节给出数据项与约束项学习结果; 3.5节给出了模型实现的一些参数设定 与实验结果,并与已有的一些方法的结果进行了比较分析。最后3.6节给出了小 结,我们在这里讨论了本章所提出方法的优缺点,并以此为基础分析了下一步 工作如何展开。

3.2 基于统计学习的光流估计模型

在本模型中,我们将光流估计问题归纳为一阶MRF概率推理问题。在给定 图像序列*I*₁,*I*₂的条件下,我们求取速度场的后验概率*P*(**w**|*I*₁,*I*₂)。具体描述如下:

$$P(\mathbf{w}|I_1, I_2; \Omega) \propto p(I_1|\mathbf{w}, I_2; \Omega_d) \cdot p(\mathbf{w}; \Omega_r)$$
(3-1)

其中, $\Omega = \{\Omega_d, \Omega_r\}$ 分别为数据项模型参数与平滑项模型参数, $\mathbf{w} = (u, v)$ 为速度矢量, I_1 , I_2 为连续的两帧图像。

对式子 3-1两边取-log得到基于MRF的能量表达式,

$$E(\mathbf{w}, I_1, I_2; \Omega) = \sum_{\mathbf{s} \in C_1} V_d(\mathbf{w}_{\mathbf{s}}, I_1; I_2, \Omega_d) + \lambda \sum_{\mathbf{s}, \mathbf{s}' \in C_2} V_r(\mathbf{w}_{\mathbf{s}}, \mathbf{w}_{\mathbf{s}'}; \Omega_r)$$
(3-2)

其中, $V_d(\mathbf{w_s}, I_{1s}; I_{2s}, \Omega_d) = -\log(\rho(V'_d(I_{1s}, I_{2s}, \mathbf{w_s}); \Omega_d))$ 为数据项势函数, $V_r(\mathbf{w_s}, \mathbf{w_{s'}}; \Omega_r) = -\log(\rho(u_s - u_{s'}; \Omega_h) \cdot \rho(v_s - v_{s'}; \Omega_v))$ 为约束项势函数, $\Omega_r = \{\Omega_h, \Omega_v\}$, $\mathbf{w} = (u, v)$ 为速度矢量, s为图像位置坐标, C_1 为图像单位置团块 (single-site clique), C_2 是二邻域点位置团块 (pair-site clique), λ 为平滑项权重, I_1 , I_2 为连续两帧图像。

在这个方法中,通过采用RBF神经网络模型来对数据项与约束项进行学 习。在建立数据项模型*p*(*I*₁|**w**, *I*₂; Ω_{*d*})的时候,我们首先通过使用光流真值对第 二帧图像*I*₂进行校正,这样得到的图像除了遮掩、边界的地方外都是与第一帧 图像相应位置相同的。其次,在得到第二帧图像的校正后,将结果图像与第一 帧图像求差得到亮度不变误差。最后,求取亮度不变误差的统计并通过RBF模 型进行学习。这样便可以得到数据项。对于平滑项,我们通过计算光流真值速 度场两个分量导数的统计,并同样采用RBF模型进行学习以建立平滑项。下节 我们对数据项与平滑项的具体实现进行详细介绍。

3.2.1 数据项建立与学习

我们利用亮度不变假设将光流模型的数据项建模为关于亮度变化的概率模型。模型描述如下:

$$p(I_1|\mathbf{w}, I_2; \Omega_d) = \prod_{\mathbf{s} \in C_1} \rho(V'_d(I_{\mathbf{s}}, \mathbf{w}_{\mathbf{s}}); \Omega_d)$$
(3-3)

其中, w为速度矢量w = (u, v), $\rho(\cdot)$ 为数据项函数, s为图像位置矢量, C_1 为图像单位置团块(single-site clique), Ω_d 为数据项模型参数, $V'_d(I_s, w_s)$ 为亮度不变假设,在本模型中,我们根据需要使用了它的两种不同表达方式,其具体描述如下式:

• 亮度不变假设的定义的模型: 用于数据项模型的学习,

$$V'_d(I_{\mathbf{s}}, \mathbf{w}_{\mathbf{s}}) = I_2(\mathbf{s} + \mathbf{w}_{\mathbf{s}}) - I_1(\mathbf{s})$$
(3-4)

• 亮度不变假设的定义的模型的变形:用于计算光流场,

$$V'_d(I_{\mathbf{s}}, \mathbf{w}_{\mathbf{s}}) = \nabla I_2(\mathbf{s}) \cdot \mathbf{w}_{\mathbf{s}} + I_t(\mathbf{s})$$
(3-5)

其中, I_1 , I_2 为连续的两帧图像, I_t 为图像序列的时间导数, ∇I_2 为第二帧图 像 I_2 的空间导数。

我们之所以要采用亮度不变假设的两种表达式,是由我们建立的数据项的 在进行参数学习与光流计算时的特点决定的。对于训练部分中,虽然可以通过 上面式子 3-5进行计算,但是我们更容易通过利用光流真值直接对第二帧图像进 行校正,然后与第一帧图像求差。这种方法使得计算更直观,更简洁。因此, 在训练中我们直接采用第一种表达方式。对于在光流场计算过程中,根据计算 算法的特点,采用第二种表达方式更简单,而第一种表达式会使得计算很麻 烦、耗时。

为了建立数据项,我们首先通过利用已有的光流真值对第二帧图像进行 校正,并求取第一帧图像与校正后图像插值的直方图统计。通过对统计结 果进行学习,就可以建立我们的数据项。有很多方法可以用于学习此模型, 如GSM(Gaussian Scale Mixture)模型。当然,也有一些人通过选择一些已有的 与上统计直方图结果曲线相似的函数模型作为数据项模型。但这种模型由于只 是在某些部分与直方图结果比较相似,总体上还是不准确,即不能准确的反应 计算得到的直方图的信息。

在我们的模型中,我们采用了RBF(Radial Basis Function)神经网络模型来 学习数据项,

$$f(x;\Omega) = \sum_{i=1}^{N} w_i \cdot \exp(-(x-\mu_i)^2/\sigma^2) + C$$
(3-6)

其中, w_i是模型中每个子表达式的权重, μ_i是模型中每个子表达式的均 值, C常量值, σ为模型方差, 其值通过经验手动设定。

RBF模型有两个主要的优点:

- •参数可以通过简单有效的训练方法进行学习。
- •规模不大,并且不会随着输入变量个数的增加而增加。

需要指出的是,RBF神经网络学习的结果是数据项: $-\log(x;\Omega) = f(x;\Omega)$ 。

3.2.2 约束项建立与学习

我们的约束项描述如下:

$$p(\mathbf{w};\Omega_r) \propto \prod_{(\mathbf{s},\mathbf{s}')\in C_2} \rho(u_{\mathbf{s}} - u_{\mathbf{s}'};\Omega_h) \cdot \rho(v_{\mathbf{s}} - v_{\mathbf{s}'};\Omega_v)$$
(3-7)

其中u,v为速度w的分量, s, s[']是图像位置坐标, C_2 是二邻域点对集合。 $\rho(\cdot)$ 为约束项函数, $\Omega_r = (\Omega_h, \Omega_v)$ 分别为水平向势函数与垂直向势函数的的参数。

我们利用己有的训练数据集分别计算两个速度分量的梯度。并采用与计算 数据项直方图相似的方式分别计算速度梯度的统计。我们仍然采用式 3-6来学习 约束项。这样,我们建立了最终的约束项模型。这个模型描述了邻近像素点速 度之间的关系。

需要指出的是,RBF神经网络学习的结果是约束项: $V_r(\mathbf{w}_s, \mathbf{w}_{s'}; \Omega_r) = -\log(\rho(u_s - u_{s'}; \Omega_h) \cdot \rho(v_s - v_{s'}; \Omega_v))$ 。

3.3 光流估计的多尺度模型

到上节为止,我们的光流模型已经建立起来。但是它只能对小位移运动(大小为1个像素的运动)进行求解。为了使得上述所建模型能应用于大位移运动, 我们对其进行改进得到光流估计的多尺度模型,使得模型可以求解图像序列或 者视频中的大位移运动。

我们采用的多尺度的方法改进模型。在模型中,多尺度金字塔的总层数 用N表示, I^k代表第k层的图像, I^N为最底一层或者说是尺度最小的一层。在这 个多尺度金字塔中,最底层两帧图像I^N和I^N直接使用式3-2进行计算。然后计 算所得的结果会被插值到邻近的上一层,在那里它将用于对第二帧图像进行校 正。在本方法中,我们采用双线性差值法对速度进行插值。插值后的速度用于 两个目的:

- ① 校正第二帧图像,以用于相应多尺度金字塔相应层的光流求解。
- ② 与相应多尺度金字塔层计算的速度结果叠加,以用于邻近的上一层的光流运算。

我们用**w**^{*k*}表示从低一层插值的来的速度值,这样我们可以通过下式对第二帧图像进行校正:

$$\overline{I}_{2}^{k}(\mathbf{s}) = I_{2}^{k}(\mathbf{s} + \overline{\mathbf{w}}^{k}) \tag{3-8}$$

式中, \overline{I}_2^k 为在k层校正后的结果,s为位置坐标。

当*I*^{*k*}^{*k*}计算出来后,便可以与*I*^{*k*}^{*k*}进行的*k*层的光流求解运算。在我们的方法中,我们采用了4层的多尺度金字塔。

3.4 参数学习

下面我们对本数据项与约束项的建立方法进行简单介绍:

• 数据项学习

(1) 图 像 数 据 我 们 釆 用 的 : 是middlebury(http://vision.middlebury.edu/flow/data/)的 冬 像 序列库,它有一系列带有光流真值的图像序列可以提供给研究者使 用。在本方法中,我们使用的是Hydrangea图像序列以及其光流真值 (参见附录B)。

② 图像差值及其统计结果

利用光流真值将第二帧图像校正,获取它与第一帧图像的差值3-1, 并求取差值结果的直方图统计。



图 3-1校正图像与第一帧图像的差值结果,由于实际中差值结果无法阅读,我们在这 里做了简单处理,将结果以4为阈值进行了二值化,图像中黑色区域为差值大约4,白 的区域为差值小于等于4。



图 3-2 差值图像的直方图统计

从图3-2中可以看出,图像差值几乎集中在[-20,20]之间,与图3-1相 比较,可以看出,差值比较大的地方(差值大于4的地方)位于图像 的边缘与遮挡部分。这样我们可以建立关于图像差值的统计模型。当 求取式(3-1)的最大后验估计时,此统计模型用于对像素差值的约束 (像素差值越小,模型获得的概率越大,在后续介绍内容中,由于对 模型做了处理,其等效理解为,像素差值越小,惩罚越小,像素差值 越大则相反)

- ③ 对统计结果进行处理,使其与式3-2的表达式相符合。 由于我们对数据项表达式做了理。这一步,我们对上一步骤获取的统 计直方图做相应处理结果如图3-3,并对结果进行学习。
- ④ 学习结果

使用RBF模型对上节计算得到的结果进行学习,我们可以得到学习结果如图3-4。此结果即为数据项。

• 约束项学习

我们利用训练数据集,分别计算速度两个分量水平向梯度得到结果如 图3-5,3-6。在这里我们假设速度场是各向同质的,因此只对速度场的一个



图 3-3 差值图像的处理后的直方图统计



图 3-4 学习结果。其中点划线为原始数据,实线为学习曲线

方向进行求导。根据计算的结果3-5,3-6:速度两个分量在水平向的统计结果 很相似,因此为了简化模型我们对这两个势函数取同样的参数。

图3-7为RBF模型学习的结果。通过图3-7可以看出,邻近两点的速度差异 越大,对其的惩罚也越大。通过这种惩罚方式可以达到约束邻近点之间速 度差异过大的目的。



图 3-5 速度分量u水平向导数的统计

3.5 **实验结果与分析**

我们采用离散的优化求解方法对建立的模型进行求解。对于本模型,我们 使用Sequential Tree-Rewighted message passing(TRW-S)算法与Graph-Cut算法 对光流场进行求解,下面先简单介绍这两种方法:

• Sequential Tree-Reweighted message passing(TRW-S)算法

TRW-S也叫序列化的Tree-Reweighted max-product message passing(TRW)算法由Kolmogorov [53]提出,它是TRW [52]的改进型版 本。TRW提出的目的是为了解决在max-product求解中,使得能量下 界最大化的问题。但是,它并不能保证能量一直上升,反而在有些时候计 算的能量会下降,另外TRW能量并不总是收敛 [53]。基于此Kolmogorov 对TRW进行了修改,以使得获得的能量不会下降。与BP算法和Graph-Cut等算法相比,TRW-S有很多优点,如Graph-Cut只能解决 [50,51]有限 的一类问题。如果带解决问题超过这个限制,就只能寻找其它的算法进行 求解如选用BP,TRW等。

• Graph-Cut算法



图 3-7 约束项模型学习结果

我们也使用了Graph-Cut算法求解我们的模型。对两个状态α与β,交换算 法通过找出可变为状态β的状态α,或者可变为状态α的状态β 来达到局部 最小化 [54]。交换状态一直到无α与β可交换为止,此时的标签使得最终能 量达到最小。为了实现交换算法,通常采用最小割/最大流算法。

由于离散优化求解方法用于求解MRF中的单标签值,如在图像去噪音处理中,可以使用像素值直接作为标签,最后求解的结果即为图像像素值。换句话

说,在这种应用中,像素值很容易离散化(其本身就是离散值)。但是在光流 求解中,速度矢量为连续量,并且不是单值量。因此,我们需要进行一些预处 理,使得这些优化方法可以求解我们的模型。

通过设计一个标签域-速度域映射函数,我们解决了这个问题。由于我们所 采用的优化求解方法为离散方法,我们首先需要将速度场离散化。考虑到在每 一级金字塔中图像所含运动不会大约2个像素,我们将速度单向分量[0,2]等分 为N份,这样得到的速度域为含有N²个速度矢量的矩阵如下式,

$$V = \begin{bmatrix} u_1 & v_1 \\ u_2 & v_2 \\ u_3 & v_3 \\ \vdots & \vdots \\ u_{N^2} & v_{N^2} \end{bmatrix}$$
(3-9)

我们以速度域每个行的索引i作为其标签得到如下映射函数,

$$V(i) = [u_i, v_i]$$
(3-10)

通过上述方法,我们得到了标签域-速度域映射函数。在运算过程中,对于 任意结点所用到的状态,我们都通过使用这个映射函数做标签-速度映射。这 样,算法运行时虽然是在对标签进行选择,实际上是在对速度向量进行选择。 通过这种方法,我们克服了所用到的离散优化求解方法不能求解我们的模型的 问题。

在进入下节前,需要对将要用到的光流评估方法做一下介绍。常用到的 量化的光流评估方法有平均角度误差AAE(Average Angle Error)、角误差的标 准差STDAE(Standard Deviation of Angle Error)、平均幅值误差AEP(average end-point)、幅值误差的标准差STDEP(standard deviation of end-point),另外 还有角度误差AE(Angle Error),它通常用于直观的给出所求得光流场与光流场 真值相应位置的角度差别(详见附录A)。

有了上面的一些光流评估方法,我们在下节介绍关于实验的一些参数设定 与实验结果分析。

3.5.1 实验设定

在实验中,我们对两种图像序列进行了测试。一种是包含云层运动的Yosemite图像序列,另一种是真实的流运动图像序列。对这两种图像序列 我们分别用Sequential Tree-Reweighted message passing(TRW-S)算法与Graph-Cut算法对我们建立的模型进行光流求解。并分别进行了光流估计结果评估。 由于图像本身的边界,会使得边缘中的运动物体(局部)消失或者增加,因此,我们在进行光流评估时对图像边缘的10个像素进行了忽略。

在光流估计模型中,我们设定金字塔模型为4层,光流算法如Algorithm 2。

3.5.2 结果分析

对于我们的模型,我们采用TRW-S和Graph-cuts两种优化求解方法最两种不同的图像序列进行了光流求解,下面分别对所求结果进行分析。

• Yosemite图像序列它包含云层的平移运动与山体的垂直向外运动。我们采用TRW-S和Graph-cuts算法,用我们建立的模型对此图像序列的光流进行

Algorithm 2 光流求解算法

- 初始化金字塔模型层数N = 4,初始化最底层累加速度矩阵W = 0(累加速 度矩阵是一个变化量,会随着金字塔模型层的变化而变化,并且其大小与 相应金字塔模型层中图像的大小一致)。
- 2: 以图像 I_1^N , \overline{I}_2^N (\overline{I}_2^N 为校正后的图像, 对于最底层的图像 $\overline{I}_2^4 = I_2^4$) 为输入, 采用TRW-S算法或者Graph-Cut算法计算此层光流 w_N 。
- 3: 将计算结果w_N与累加速度矩阵W相叠加,所得结果通过双线性插值法插值 到邻近的上一层,并将W的大小调整到与上一层金字塔中图像相同大小。 插值结果重新存储到大小调整后的累加速度矩阵W。
- 4: N = N 1.
- 5: 如果N = 0, 进入下面步骤, 否则回到第2步。
- 6: 通过状态-速度映射函数, 求取最终求解的光流场
- 7: 对求解得到的光流场进行AAE、STDAE评估。

求解,结果如图3-8。从图 3.8(d)与图 3.8(f)中,可以看出角度误差比较大的地方位于山体与云层的相交处,另外,还有在云层处的小块区域角度误差也比较大。对于前一个问题,原因是由于物体遮挡,使得第二帧新暴露出的像素在第一帧图像中没有对应的像素,这样就不能求解出其正确的速度;对于第二个问题,原因是在云层运动中,云层的亮度会发生变化,这样云层中同一点,在前后两帧中由于亮度发生变化,使得对应点无法被对应起来。这样会影响这个点所求速度的精度。

在表3-1中,我们给出了用不同优化求解方法对Yosemite图像序列求解结 果的量化评估。在进行量化评估时,我们对图像边缘的10个像素进行了忽 略,这样消除了运动物体在边缘由于进入或者超出图像边界,因为的像 素无法对应而导致的光流结果不确定。从图中,这应用两种优化求解方 法求取的结果大致相同。用TRW-S求解的结果相对好一些。原因是,应 用TRW-S求解的结果不会出现块状效果。

• 流图像序列运动

流图像序列是在实验室中拍摄的到的,其所含运动为流体的对流运动。图像序列平均位移为4个像素,最大位移为15个像素 [19]。在图3-9中给出了

光流估计方法	AAE	STD
Anandan(Barron 1994)	13.36	15.64
Liu et al.(2002)	13.18	13.24
Roy et al.(2000)	13.00	
Singh(Barron 1994)	10.44	13.94
Our Model+Graph-Cut	10.42	12.31
Our Model+TRW-S	10.23	12.27
Cassisa et al.(2008)	7.02	7.96

表 3-1 Yosemite图像序列光流计算结果的量化评估。AAE=Average Angle Error, STD= Standard Deviation。

图像序列的的第一帧图像与采用TRW-S算法与Graph-Cut算法对模型求得的结果。在结果图中,计算出的光流场能比较好的符合实际的流运动。从图中可以看出,采用TRW-S算法求解的光流场在左侧无运动的空白区域误求解要好于Graph-Cut算法。但是在光流场从左上到右下对角向运动区域,TRW-S算法求解的光流场不如Graph-Cut算法求解的好。但是总体上这两种优化方法对我们的模型求解的结果还是比较鼓舞人的。

3.6 小结

在这一章,我们详细的介绍了我们提出的基于统计学习的光流估计模型, 并采用两种不同的优化方法TRW-S算法和Graph-Cut算法对模型进行求解,并 给出了应用于两种图像序列的实验结果。这个模型的创新在于:在本模型中, 我们引入了RBF(Radial Basis Function Neural Network)模型对图像序列已有光 流真值数据进行学习,以得到更准确的光流估计模型。我们还提出了建立了将 统计学习与MRF混合的光流估计模型,并在数据项建立与进行计算过程中,采 用了不同的亮度不变表达式。通过采用统计学习与MRF模型混合建模,我们可 以使用离散的优化求解方法统计学习如TRW-S算法和Graph-Cut算法对模型进 行求解。

通过实验结果,我们的方法虽然有了一些初步的结果,但是还有很多不 足,有很多地方需要进一步改进:

① 训练数据。

我们对光流真值图像序列进行筛选,选用了含有典型运动(匀速运动,旋 转等)的数据作为训练数据。如果将大量真值数据进行整合,一起用于统 计模型的学习,学习的模型应用范围会更具普遍性。

② 模型。

模型仍然对遮挡、边界等区域无法求解正确的光流场。如果将模型进行改进,使得其可以对运动进行分类,并将遮挡、边界等区域通过分类方法归类到其领域运动区域。这样可以利用被归类到的区域的运动信息对这些"盲"区域进行光流估计。



(a) Yosemite图像





(c) Graph-Cut算法计算的光流



(d) Graph-Cut算法计算的光流角度误差



(e) TRW-S算法计算的光流









第四章 图像序列中的光流估计与运动分割

4.1 引言

从第二章的文献综述可以了解到,目前已有的方法几乎都是针对图像序列 中某种运动提出的。它们可以很好的用于求解含有这类运动图像序列。但是, 当图像序列中的运动类型发生了变化,或者图像序列中含有多种运动方式时, 它们求解的光流场不再如只针对它们特定的运动求解的结果好。原因是,他们 针对特定类型运动提出了专门的数据项与约束项,这些一旦确定就只能适用于 此类型运动。在本章,我们承认图像序列中含有多种不同的运动,并提出一种 基于运动分类的光流估计模型对图像序列中含有的多种运动进行光流求解,同 时对这些运动进行分割。

本章组织如下: 4.2 节给出我们提出的模型的思想,并对模型进行总体介绍; 4.3 节详细介绍模型的数据项; 4.4 节介绍模型的约束项,在本节将详细给出不同运动约束式,基于约束式的运动模拟以及不同约束式的选择; 4.5 节介绍不同邻接运动的平滑约束;在4.6中,给出典型运动的模拟结果以及在方法实现时要处理的一些问题; 4.7 节中我们给出实验的实现与结果分析;最后在4.8 节进行本章总结,讨论本章所提出的方法的优点与不足,并提出下一步的改进。

4.2 模型思想与描述

通过研究已有光流估计方法面临的问题我们发现,已有方法是针对特定类 型运动提出的。如果将其应用于含有超出这种运动的图像序列,他们获取的光 流场将变得不如以前好。因是,他们针对特定类型运动提出了专门的数据项与 约束项,这些一旦确定就只能适用于此类型运动。我们发现它们一般都含有多 种运动,并且我们进一步发现它们可以有几个基本的运动组合实现。我们提取 出了4种基本的运动,并在此基础上提出了光流估计与运动分割模型。

在我们的模型中,我们承认图像序列含有多种运动。我们的目的是通过采 用一系列典型运动表达式建立一种光流估计模型,使其可以对含有常见的运动 图像序列进行光流场估计。在本文中,我们提出了4种典型的运动:匀速运动、 单向加速运动、速度幅值不变角度加速的运动以及速度幅值与角度均加速的运 动。由于这四类运动以幅值与角度进行描述,它们在极坐标中更易于表达。因 此,我们建立了基于极坐标系的光流估计模型。

由于采用了运动分类的思想,在求解过程中,我们不但可以求取图像序列 的光流场,还可以获取不同运动的分割。这对于需要进行运动物体分割的领域 如目标跟踪等很有用。

在第二章中,我们简单介绍了MRF模型,它由数据项与约束项构成。为了 能建立更一般化的光流估计模型,我们将约束项进行了改写,即由典型运动表 达式构成,过平滑约束式选择子函数可以对不同的运动采用不同的平滑约束。 同时我们添加了运动类型间的的平滑约束,模型表达式如下,

 $E(I, \rho, \alpha, \theta, \ell) = E_d(I, \rho, \alpha) + \lambda E_r(\rho, \alpha, \theta, \ell) + E_c(\ell)$ (4-1) 其中, I 为图像, $\vec{w} = (\rho, \alpha)$ 为极坐标系下的速度表达式, θ 为约束不同运动的 参数矩阵, ℓ 为运动类型标签, λ 为约束项权重, E_d , E_r 与 E_c 分别为数据项, 平 滑项,运动类间的平滑约束项。

在这个模型中,我们采用的是极坐标。原因是:1. 这样可以使得我们更 能方便的描述4中典型运动;2. 通常的光流方法认为笛卡尔坐标系系的速度分 量*u*与*v*是相互独立的,我们认为在极坐标下ρ与α是独立的;这样可以使得3.我 们通过假设ρ与α 是各向同质的,能更方便的通过程序实现本方法。

4.3 数据项

我们基于亮度不变假设进行光流模型的数据项建模。其表达式描述如下,

$$E_d(I,\rho,\alpha) = \sum_{\mathbf{s}\in C_1} V_d(I_{\mathbf{s}},\rho_{\mathbf{s}},\alpha_{\mathbf{s}})$$
(4-2)

式中, I 为图像, s 为图像位置坐标, C_1 为图像单位置团块 (single-site clique), V_d 为数据项势函数, (ρ, α) 为极坐标系下的速度表达式。

我们对数据项势函数V_d进行进一步展开,

 $V_d(I_{\mathbf{s}}, \rho_{\mathbf{s}}, \alpha_{\mathbf{s}}) = (I_x(\mathbf{s}, t+1)\rho_{\mathbf{s}} \cos \alpha_{\mathbf{s}} + I_y(\mathbf{s}, t+1)\rho_{\mathbf{s}} \sin \alpha_{\mathbf{s}} + I_t(\mathbf{s}, t))^2$ (4-3) 其中, $I_x(\mathbf{s}, t+1) = I_y(\mathbf{s}, t+1)$ 分别为第二帧图像的水平向与垂直向导数,第二 帧图像的空间梯度, $I_t(\mathbf{s}, t) = I_2(\mathbf{s}, t+1) - I_1(\mathbf{s}, t)$ 为图像的时间导数。

在上式中,

$$\rho_{\mathbf{s}}\cos(\alpha_{\mathbf{s}}) = u_{\mathbf{s}} \tag{4-4}$$

$$\rho_{\mathbf{s}}\sin(\alpha_{\mathbf{s}}) = v_{\mathbf{s}} \tag{4-5}$$

其中, (u, v)为笛卡尔坐标系下的速度表示。

我们采用了在极坐标系上进行建模,但是这里数据项的势函数仍将极坐标 系下的速度转换到了笛卡尔坐标系下的速度,原因为:我们的方法建立在以像 素为基元的固定的拓扑图上。对于图上的每一个基元,必须通过笛卡尔坐标系 下的坐标(*x*,*y*)进行索引。如果直接采用接坐标进行索引,将变得很困难。所以 对于图上的任意基元的索引,最终要通过将极坐标转换到笛卡尔坐标系下来实 现。

4.4 先验知识项

为了对不同的运动进行更好的平滑约束,我们提出了对典型的运动建立专 门的描述式的方法。他们一起被用于对图像序列中的运动进行约束。为了实现 这种约束方式,我们使用了一个选择子函数。它可用于对不同的运动选择不同 约束式。我们的先验知识项描述如下,

$$E_r(\rho, \alpha, \theta, \ell) = \sum_{\mathbf{s}, \mathbf{s}' \in C_2} V_r(\rho_{\mathbf{s}}, \alpha_{\mathbf{s}}, \rho_{\mathbf{s}'}, \alpha_{\mathbf{s}'}, \theta, \ell_{\mathbf{s}})$$
(4-6)

式中, s 为图像单位置坐标, C_2 为图像二领域点位置团块, V_r 为平滑项势函数, (ρ, α) 为极坐标系下的速度表达式, θ 为约束不同运动的参数矩阵, ℓ 为运动类型标签。

约束项势函数通过对不同约束函数进行选择,以实现对不同运动的约束, 其展开式如下,

$$V_r(\rho_{\mathbf{s}}, \alpha_{\mathbf{s}}, \rho_{\mathbf{s}'}, \alpha_{\mathbf{s}'}, \theta, \ell_{\mathbf{s}}) = \sum_{k \in L} (\hbar_k(\ell_{\mathbf{s}}) \ \theta_k \cdot \overrightarrow{f}(\rho_{\mathbf{s}}, \alpha_{\mathbf{s}}, \rho_{\mathbf{s}'}, \alpha_{\mathbf{s}'}))$$
(4-7)

其中, $\hbar(\cdot)$ 为选择子函数, ℓ 为运动类型标签。 θ 为参数矩阵, θ_j 为 θ 为的第j行向量, \vec{f} 为函数向量,用于描述不同的运动,L为运动类型标签集,其表达式如下,

$$\hbar_k(\ell_{f s}) = egin{cases} 1 & { t if} & \ell_{f s} = k \in L \ 0 & { t elsewise} \end{cases}$$

其中,s为位置坐标,L为运动类型标签集,k为运动类型标签。

在这个式子中,有三个关键部分:1. θ ; 2. \overrightarrow{f} ; 3. ℓ 。参数 θ 决定了约束 式的线性组合方式,使得相应的运动类型可以被算法检测到。函数向量 \overrightarrow{f} 是一 个典型运动表达式集,由典型的运动描述式组成。所提取的基本运动是否具有 代表性,是决定模型准确的一个重要方面。另一方面,运动类型标签直接决定 着选取何种运动描述式的线性组合,因此,运动类型标签的准确性很重要。

4.4.1 运动分类

这一节,我们对运动分类方法进行详细介绍。根据速度幅值与角度的变化,我们提取出了4类典型的运动。我们对这些运动建立了表达式,他们被组合在一起用于对图像序列中不同的运动进行不同的平滑约束。为了使得模型更具描述性,在这个模型中,我们采用了极坐标系进行建模。

4.4.1.1 基于角度与幅度的分类

基于速度幅值与角度的变化,我们提取了4类典型的运动,并对这些运动建立了相应的表达式。在给出具体的表达式之前,我们给出导数的定义, $\Delta y = y_s - y_{s'}$,其中y为任意要求导的变量,s = s'为位置坐标。具体的描述式如下,

① 匀速运动,

 $\parallel \Delta \rho \parallel = 0, \ \parallel \Delta \alpha \parallel = 0.$

② 速度角度不变,幅值变化的加速运动,

 $\parallel \Delta(\Delta \rho) \parallel = 0, \parallel \Delta \alpha \parallel = 0.$

③ 速度幅值不变,角度变化的加速运动,

 $\parallel \Delta \rho \parallel = 0, \parallel \Delta (\Delta \alpha) \parallel = 0.$

④ 速度幅值与角度均变化的加速运动,

 $\| \Delta(\Delta \rho) \| = 0, \| \Delta(\Delta \alpha) \| = 0.$

式中, ρ为速度幅值, α为速度角度。将这4种描述式组合在一起, 就得到了我 们的函数向量。对于以上式子, 它不仅能够对所描述的4类典型运动进行约束, 对于同时具有以上运动组合的运动, 通过选择不同的参数它们也可以被很好的 约束。

为了使得模型更易于程序编写,我们对上述表达式进行了进一步改进使 得: $\| \Delta(\Delta \rho) \| = \| \Delta \rho - \overline{\Delta \rho} \|, \| \Delta(\Delta \alpha) \| = \| \Delta \alpha - \overline{\Delta \alpha} \|.$ 这里 $\overline{\Delta \rho}$ 与 $\overline{\Delta \alpha}$ 分别 为 $\Delta \rho$ 与 $\Delta \alpha$ 的均值。通过这种方式,我们可以近似获得 ρ 与 α 的二阶导数近似。

4.4.2 基于运动分类的平滑项描述

在介绍完典型运动模型已经基于这些模型生成的图像序列后,本节将对平 滑项进行进一步的展开。为了实现模型,我们首先需要构建函数向量,这是模 型的基础。其次,我们需要定义选择子函数。选择子函数用于给不同像素点选 定合适的平滑约束,以期求取更准确的光流场。最后,我们需要一个参数矩 阵,用于确定各种平滑约束表达式的组合方式,它也相应的确定了可以准确检 测到的运动的类型。下面我们对这三个部分分别予以详细介绍。

4.4.2.1 函数向量

函数向量是本模型的基础,它的准确性直接决定着可以检测到的运动的全 面性。通过在小节"基于角度与幅度的分类"中,我们可以定义如下函数向 量,

$$\vec{f} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{bmatrix} = \begin{bmatrix} \| \Delta \alpha \|^2 \\ \| \Delta \rho \|^2 \\ \| \Delta \alpha - \overline{\Delta \alpha} \|^2 \\ \| \Delta \rho - \overline{\Delta \rho} \|^2 \end{bmatrix}$$
(4-8)

式中, $\overline{\Delta \alpha}$ 与 $\overline{\Delta \rho}$ 是 $\Delta \alpha$ 与 $\Delta \rho$ 的均值。

需要指出的是,本函数向量不是唯一的。它由我们提取出的典型运动决定,人们可以根据不同的需要,选择不同的基运动典型运动描述式比如,在云 层运动中,其主要运动为平移、加速、漩涡已经扩散运动,这些运动用于建立 一个函数向量以检测这些运动。另外,由于我们提取的是典型运动,或者说是 常见运动,目前它还不能涵盖所有的运动,但是相对于已有方法,它的适用性 更强,尤其是含多种运动的图像序列。

4.4.2.2 **选择子函数**

本模型的一个关键点是约束式的选择。我们首先需要对像素点打标签,然 后根据这个标签进行不同的运动约束。因此,标签的正确性直接决定着所求解 光流场的准确性。只有在像素被打上正确的标签时,才能够为其选择合适的约 束式以进行光流求解。另一方面,基运动典型运动表达式集的选择也影响着标 签的正确性。合理的基运动典型运动表达式集有助于为不同的运动打上正确的 标签。 在本模型中,我们采用局部能量比较的方法进行打标签。对于一个像素 点,我们按顺序依次为其分配一个标签,并计算采用这个标签时的局部能量。 我们选择使得局部能量最小的标签作为这一步骤的结果。这样打标签的问题可 以被解决。在完成打标签之后,我们就可以通过选择子进行运动约束。

为了实现选择子描述如下,

$$\hbar_k(\ell_{\mathbf{s}}) = \begin{cases} 1 & \text{if} \quad \ell_{\mathbf{s}} = k \in L \\ \\ 0 & \text{elsewise} \end{cases}$$

其中,s为位置坐标,L为运动类型标签集,k为运动类型标签。

通过采用选择子,可以对不同的运动类型标签选择相应的运动约束。这样 就实现了对不同运动的约束。

4.4.2.3 参数矩阵

参数矩阵式模型的另一个重要方面。参数的设定直接影响着约束式的线性 组合方式。我们将模型参数设计为一个参数矩阵,通过对这个矩阵进行设定, 可以将约束式有效的组合起来。对参数矩阵的设定可以通过两种方法实现:一 种是通过利用己有的图像序列与其光流真值对模型进行训练来设定参数矩阵; 另一种是通过手工方式对参数进行设定。下面给出我们的参数矩阵,

$$\theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \end{bmatrix} = \begin{bmatrix} \theta_{11} & \theta_{12} & \theta_{13} & \theta_{14} \\ \theta_{21} & \theta_{22} & \theta_{23} & \theta_{24} \\ \theta_{31} & \theta_{32} & \theta_{33} & \theta_{34} \\ \theta_{41} & \theta_{42} & \theta_{43} & \theta_{44} \end{bmatrix}$$
(4-9)
b) δ δ ϕ $h \equiv H \sum \theta_{ii} = 1$.

式中, θ_{ℓ} 是像素标签为 ℓ 的参数向量且 $\sum_{i} \theta_{ij} = 1$

4.5 运动类间的平滑约束

在光流求解中,有一些局部的小块会被误分类,因此会使得计算的分割结 果出现小块儿。为了消除这种问题,我们引入了运动类间的平滑约束。它能够 对被设定不同标签的邻接点进行"强惩罚"。如果某些点被设定错误标签,它 将会受到"强惩罚",因此会引入更多的能量,是总能量增加。这样,其错分 的标签会被修正过来,以降低能量。通过这种方法,这些误分的区域会被修正 过来。下面给出运动类间的平滑约束表达式,

$$E_c(\ell) = \sum_{(\mathbf{s},\mathbf{s}')} \delta(\ell_{\mathbf{s}},\ell_{\mathbf{s}'})$$
(4-10)

式中, $\delta(\ell_{\mathbf{s}}, \ell_{\mathbf{s}'})$ 为knechel function模型,

$$\delta(\ell_{\mathbf{s}}, \ell_{\mathbf{s}'}) = \begin{cases} 0 & \text{if } \ell_{\mathbf{s}} = \ell_{\mathbf{s}'} \\ 1 & \text{otherwise} \end{cases}$$
(4-11)

这个模型课对分配有不同标签的邻接像素进行"惩罚",以修正那些错分的标签。

4.6 实现问题

4.6.1 运动模拟与评估

本节我们模拟并生成了一些符合这4类约束的图像序列。

通过对4类运动表达式进行推导,我们得到了这些方程的解。在这些解得基础上,我们通过程序生成了符合这类约束的两帧图像序列。下面我们给出生成的图像序列的方法。

① 匀速运动,

约束式:

$$\begin{cases} \parallel \Delta \rho \parallel = 0 \\ \parallel \Delta \alpha \parallel = 0 \end{cases}$$
(4-12)

约束式的解:

$$\begin{cases} u = C_1 \\ v = C_2 \end{cases}$$
(4-13)

② 速度角度不变,幅值变化的加速运动,

约束式:

$$\begin{cases} \parallel \Delta \rho - \overline{\Delta \rho} \parallel = 0 \\ \parallel \Delta \alpha \parallel = 0 \end{cases}$$
(4-14)

约束式的解:

$$\begin{cases} f_1(x,y) = -\frac{1}{C}(Ax + By) \\ f_2(x,y) = -\frac{1}{C}(Ax + B(-y + width)) \end{cases}$$
(4-15)

③ 速度幅值不变,角度变化的加速运动,

约束式:

$$\begin{cases} \parallel \Delta \rho \parallel = 0 \\ \parallel \Delta \alpha - \overline{\Delta \alpha} \parallel = 0 \end{cases}$$
(4-16)

约束式的解:

$$\begin{cases} u = C_1 \cos(C_2 x + C_3 y) \\ v = C_1 \sin(C_2 x + C_3 y) \end{cases}$$
(4-17)

④ 速度幅值与角度均变化的加速运动,

约束式:

$$\begin{cases} \parallel \Delta \rho - \overline{\Delta \rho} \parallel = 0 \\ \parallel \Delta \alpha - \overline{\Delta \alpha} \parallel = 0 \end{cases}$$
(4-18)

约束式的解:

$$\begin{cases} u = (C_1 x + C_2 y) \cos(C_3 x + C_4 y) \\ v = (C_1 x + C_2 y) \sin(C_3 x + C_4 y) \end{cases}$$
(4-19)

我们利用上面的解,通过程序模拟并生成了含相应运动的图像序列,并对 这些结果进行了评估验证,如图4-1,4-2,4-3,4-4。

本节生成的图像序列,可用于对模型的验证。

4.6.2 BP**算法实现**

我们通过BP算法对模型进行求解。在实现过程中,由于我们的模型与其它可以采用BP求解的模型有一下区别,

- 在我们的模型中对任意不同的两组邻接点,他们的约束平滑矩阵是不同 的,这是由于我们的约束矩阵的值与局部区域的均值有关。而在常见的可 用BP求解的模型中,这个平滑矩阵式固定的,可以提前被初始化。
- ② 我们采用了对平滑约束函数进行选择的方案,需要对局部能量就行求解, 使得在进行BP消息传递时要对BP算法进行一些修改。



图 4-1 运动类型1. 所有的4.1(c), 4.1(d), 4.1(e), 4.1(f) 的值均为0。因此,这个运动场 满足运动类型1的约束条件。



图 4-2 运动类型2. 在 4.2(c)中 || Δρ ||= 0.05 其它结果4.2(d), 4.2(e), 4.2(f)的值均为0。 因此,这个运动场满足运动类型2的约束条件。



图 4-3 运动类型3。在 4.3(c)与 4.3(d)中它们的值的绝对值≤ 10⁻¹⁵,可以视作0。
在 4.3(e)中 || Δα ||= 0.01131,不可以视作0。因此,这个运动场满足运动类型3的约束
条件。



图 4-4 运动类型4。 $\| \Delta \rho \| = 0.01414$, $\| \Delta \alpha \| = 0.01131$ 。其它结果为0。因此,这个运动场满足运动类型4的约束条件。

基于以上原因,我们需要对BP进行一些修改,使得其在进行消息更新时要:1.重新计算平滑约束矩阵;2.计算局部能量,以对平滑约束函数进行选择。由于以上修改,使得我们的模型在求解过程中会比采用常用BP算法耗时。因为在采用BP求解一般的模型时,它们的平滑矩阵由于处处一样,可以在算法开始时通过初始化完成求解,相对于在消息传递时每次重新计算平滑矩阵,它们会节省很多时间。另外一个原因是,它们没有我们模型的平滑约束函数选择这一步,这也会使得计算时间下降。

4.7 实验结果与分析

4.7.1 实验设定

通过分析,我们手工设定参数矩阵,并约束矩阵每行之和为1。

$$\theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \end{bmatrix} = \begin{bmatrix} 0.25 \ 0.25 \ 0.25 \ 0.25 \\ 0.33 \ 0 \ 0.33 \ 0.33 \\ 0 \ 0.33 \ 0.33 \\ 0 \ 0 \ 0.5 \ 0.5 \end{bmatrix}$$
(4-20)

幅值标签个数: 5,幅值步长: 0.3;

角度标签个数: 4 (整个范围为[0,2π]);

运动类型标签数: 4, 平滑系数λ: 1.0, 多尺度分级数: 3;

采用BP(Belief Propagation) [45]算法求解,

$$(\widehat{\rho}, \widehat{\alpha}, \widehat{\ell}) = \arg\min_{(\rho, \alpha, \ell)} lb(\rho, \alpha, \ell)$$
(4-21)

其中, (ρ, α) 为速度, ℓ 为运动类型标签, $lb(\cdot)$ 为置信向量。

4.7.2 结果分析

我们对三组数据进行了测试:

• 合成图像序列

这两帧图像由我们自己合成生成。它包含两类运动: 左部分向下的运动与 有部分水平向左的运动。在这个测试实验中,我们对我们的基本运动集进 行了改写。图 reffig4:results5为光流估计与分割结果。从图 4.5(c)中(速 度量参见附录B),所求解的速度场的方向是正确的,但是速度的幅值不



太平滑。在图 4.5(d)中,我们成功的分割出了两类运动: 左边的垂直向运动与右边的水平向运动。

图 4-5 合成图像序列光流估计与分割结果。在图 4.5(d)中绿色表示垂直向运动,蓝色表示水平向运动

• Hydrangea 图像序列

Hydrangea 图像序列(源于http://vision.middlebury.edu/flow/data/) 4.7.2包含两种运动:背景的水平运动与物体的旋转运动。图 4.6(c)为我们求解

出的光流结果,我们对比速度映射表(参见附录B),可以看出,所求解 出光流场的背景运动方向正确,但是幅值存在不平滑。图 4.6(d)为分割出 的结果。在本实验中,我们采用了4中基本运动,成功的分割出了两类运 动: 红色的单向加速运动与黑色的幅值角度均加速的运动。



图 4-6 Hydrangea图像序列光流估计与分割结果。在图 4.6(d)中蓝色代表匀速运动,红 色代表单向加速运动,绿色代表角度加速其幅值不变运动,黑色代表角度幅值均加速 的运动

• Urban 图像序列

Urban 图像序列源于http://vision.middlebury.edu/flow/data/) 4.7.2包含 旋转运动,其左上角速度幅值很小,右下角速度幅值很大,这给光流估计 造成了困难。图 4.7(c)为计算的光流结果。我们求解出了左上部分的运动 场,但是对于剩余部分没有成功求解出:一个原因这一部分的纹理性太 强;另一原因在于这一部分的运动幅值比较大。图 4.7(d)为分割结果。我

们同样采用了4中基本运动。对比光流真值,我们成功的分割出了两个运动区域。剩下的区域没有分割出,尤其是右下部分。其原因在于,由于光流没有能正确的求解出,相应的分割结果也是错误的。



图 4-7 Urban图像序列光流估计与分割结果。在图 4.6(d)中蓝色代表匀速运动,红色代表单向加速运动,绿色代表角度加速其幅值不变运动,黑色代表角度幅值均加速的运动

另外,我们将约束式选择的方法与约束式固定的方法通过两组数据进行了比较。在表4-1中为角度误差量化评估结果。表4-2中为角度误差的标准方差量化估计。从结果可以看出,采用约束式选择的方法所求取的结果要好于采用约束式固定的方法求解的结果。

• 平均角度误差量化估计

数据	约束式选择的方法	约束式固定的方法
Hydrangea	16.92	22.05
Urban	44.48	87.89

表 4-1 均角度误差量化评估。

数据	约束式选择的方法	约束式固定的方法
Hydrangea	22.55	36.14
Urban	43.25	55.01

表 4-2 角度误差的标准方差量化评估。

• 角度误差的标准方差量化估计

4.8 小结

这一章,我们详细介绍了我们提出的图像序列中的光流估计与运动分割模型,并展示了初步的实验结果。这个方法的创新性在于:1.与以往的针对某种运动进行建模的方法不同,我们提出了对多运动建模的思想。通过提取4种典型运动,模型通过对不同运动设定运动类型标签,并选取相应的平滑约束函数。2.由于采用了对运动分类的方法,我们不仅能求取视频的光流场,还能获取这些运动的分割。

虽然我们目前只取得了初步的结果,但是从实验结果还是可以感觉到这个 模型本身蕴含的巨大潜力。一方面是这个模型能够克服以前模型的不鲁棒性, 因为他们只是针对某种运动进行建模,不易应用于其它类型的运动。而我们的 模型由于对多种典型的运动进行了建模,能适应于更多的运动类型;另一方 面,我们的模型可以直接对视频中的物体进行分割,这对于需要进行目标分割 的领域很有帮助。

从实验结果中,我们还有如下需要进一步改进的地方;

 将图像边缘部分与遮挡等无像素匹配对点的区域进行分类,利用所分到的 类的信息对其进行光流估计。 目前,图像边缘、遮挡等引起的光流无法估计仍然是一个挑战性问题。原 因在于,在连续两帧图像中,由于遮挡等使得一帧图像中的像素点无法在 另一帧图像中找到匹配点。由于我们采用的仍然是差分模型,这个无法匹 配的像素点会被在另一帧中匹配与其像素值近的任何有可能的像素点,于 是造成了误匹配,当然这点的速度值也是不正确的。如果能充分利用运动 分类信息,通过有效的分类方法,将这些"盲区"进行分类,这样可以通 过类域的信息对其进行光流估计。使用这种方法,可以使得求解得到的光 流场更加准确。

② 参数学习

在我们的模型中,参数是通过经验手动设定的。由于我们已经人工生成了 所提取的典型运动的图像序列,如果能利用学习方法,对这些人工图像序 列进行充分利用,我们会得到更合理的参数设置。这样也会使得求解的光 流场更加准确。

第五章 结束语

本章我们将对本文的工作以及主要的贡献进行总结,并对我们下一步的工 作进行展望。

5.1 本文主要工作

本文在已有光流场求解方法的分析与总结的基础上,围绕光流场估计这一 主题进行了研究。本文的主要工作及贡献如下:

① 基于统计学习的光流估计

受 [22]的启发,我们提出了基于统计学习的光流估计模型。在本模型中, 我们建立了基于马尔科夫随机场的统计学习光流估计模型,并引入了入 了RBF(Radial Basis Function)神经网络模型对图像序列已有光流真值数 据进行学习,以得到更准确的光流估计模型。在模型学习与光流场求解两 个部分,我们分别采用了两种不同的亮度不变约束式。以使得模型的学习 训练与光流求解更有效。由于采用了基于MRF框架的方法进行建模,这 使得我们的统计模型可以通过一些有效的优化求解方法如TRW-S [53]算法 和Graph-Cutciteboykov2001fast,kolmogorov2002energy算法对光流场进行 求解。

② 图像序列中的光流估计与分割模型

我们提出了图像序列中的光流估计与运动分割模型。在本模型中,我们 基于MRF框架建立了可以对含多运动的图像序列进行求解。在这个模型 中,我们的目的就是求解含多运动的图像序列的光流场,并对这些运动进 行分割。在本模型中,我们对常见运动进行了分析并提取出4种典型的运 动。我们对这些典型运动分别进行了建模,并在这些约束的条件下我们创 建了相应的图像序列。通过对这些典型运动进行线性组合,我们建立了我 们的模型。这个模型可以对不同的运动选取不同的平滑约束,以使得这些 运动获得最佳约束,这样就能求取更准确的光流场。我们还进行了实验, 并获得了初步的结果。实验结果不但给出了求解的光流场还给出了初步的 运动分割结果。实验表明我们的模型蕴含有极大的潜力和优势,为今后的 工作奠定了基础。

5.2 未来工作展望

本文提出的图像序列中的光流估计与运动分割模型具有巨大的潜力,也为 今后的工作打下了良好的基础。未来的工作可以从以下方面进行展开:

- ① 建立基于图像序列中的光流估计与运动分割、分类的模型
 - 目前,图像边缘、遮挡等引起的光流无法估计仍然是一个难点问题。在我 们的图像序列中的光流估计与运动分割模型模型中,可以对不同的运动进 行分类。这是一个非常有用的信息,尤其是对前面提到的遮挡等区域。这 些区域一般要么位于不同物体或背景之间的边界要么位于图像边缘,他们 事实上是这些物体或者背景的一部分。之所以无法对其进行正确求解是因 为它们不能在其相邻的另一帧图像中找到对应的像素点,或者说无法匹 配。如果加入有效的分类算法,对这些区域进行分类,根据其所属类的运 动物体的运动特点可以估算出其速度值。使用这种方法对我们已有模型进 行改进,可以获取更准确的光流场。
- ② 建立基于学习的图像序列中的光流估计与运动分割模型
 - 在我们的图像序列中的光流估计与运动分割模型中,参数是通过手工设定 的。虽然具有一定的合理性,但是仍不是太准确。如果通过加入学习算 法,利用已有的光流真值对模型进行训练,获取的参数会更具合理性。在 今后的工作中,我们可以加入学习算法对模型进行改进。
参考文献

- S.Z.Li. Markov Random Fileld Modeling in Computer Vision. Springer-Verlag Press, 1995.
- [2] J.L. Barron, D.J. Fleet, and S.S. Beauchemin. Performance of optical flow techniques. International Journal of Computer Vision, 12(1):43–77, 1994.
- [3] G.M. Quenot, J. Pakleza, and T.A. Kowalewski. Piv with optical flow. Exp. in Fluids, pages 177–189, 1998.
- [4] 马颂德, 张正友. 计算机视觉--计算理论与算法基础. 北京: 科学出版社, 1998.
- [5] B.K.P. Horn and B.G. Schunck. Determining Optical Flow. Artificial Intelligence, 17(1-3):185–203, 1981.
- [6] B.D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In International joint conference on artificial intelligence, volume 3, page 3. Citeseer, 1981.
- [7] C. Zhen, G. Man-Tun, and S. Yun-Wen. The current situation and trends of optical flow estimation. Journal of Image and Graphics, 7(5):434, 2002.
- [8] H.H. Nagel. Displacement vectors derived from second-order intensity variations in image sequences. Computer Vision, Graphics and Image Processing, 21(1):85–117, 1983.
- [9] H.H. Nagel. On the estimation of optical flow: relations between different approaches and some new results. Artificial Intelligence, 33(3):298–324, 1987.
- [10] D.J. Fleet. Measurement of Image Velocity. 1992.
- BD Lucas. Generalized image matching by the method of differences. Dissertation Abstracts International Part B: Science and Engineering, 46(11), 1986.

- [12] E.H. Adelson and J.R. Bergen. The extraction of spatiotemporal energy in human and machine vision. In Proc. IEEE Workshop on Visual Motion, Charleston, pages 151–156, 1986.
- [13] J.K. Kearney, W.B. Thompson, and D.L. Boley. Optical flow estimation: An error analysis of gradient-based methods with local optimization. IEEE Transactions on Pattern Analysis and Machine Intelligence, 9(2):229–244, 1987.
- [14] E.P. Simoncelli. Distributed representation and analysis of visual motion. Massachusetts Institute of Technology, Cambridge, MA, 1993.
- [15] EP Simoncelli, EH Adelson, and DJ Heeger. Probability distributions of optical flow. In IEEE Conference on Computer Vision and Pattern Recognition, pages 310–315, 1991.
- [16] T. Corpetti, É. Mémin, and P. Pérez. Dense Estimation of Fluid Flows. IEEE Transaction on Pattern Analysis and Machine Intelligence, pages 365– 380, 2002.
- [17] D. Bereziat, I. Herlin, and L. Younes. Motion estimation using a volume conservation hypothesis. In IEEE Conference on Acoustics, Speech, and Signal Processing, volume 6, 1999.
- [18] D. Bereziat and I. Herlin. Object based Optical Flow Estimation with an Affine Prior Model. In Conference on pattern recognition, volume 15, pages 1048–1051, 2000.
- [19] C. Cassisa, V. Prinet, L. Shao, S. Simoens, and CL Liu. Optical flow robust estimation in a hybrid multi-resolution MRF framework. In IEEE Conference on Acoustics, Speech and Signal Processing, pages 793–796, 2008.
- [20] V. Prinet, C. Cassisa, and F.F. Tang. MRF Modeling for Optical Flow Computation From Multistructure Objects. In IEEE Conference on Image Processing, pages 1093–1096, 2006.

- [21] W. T. Freeman, E. C. Pasztor, and O. T. Carmichael. Learning low-level vision. International Journal of Computer Vision, 40(1):25–47, 2000.
- [22] D. Sun, S. Roth, J. P. Lewis, and M. J. Black. Learning optical flow. European Conference on Computer Vision: Part III, pages 83–97, 2008.
- [23] S. Roth and M. J. Black. On the spatial statistics of optical flow. International Journal of Computer Vision, 74(1):33–50, 2007.
- [24] P. Anandan. A computational framework and an algorithm for the measurement of visual motion. International Journal of Computer Vision, 2(3):283– 310, 1989.
- [25] R. BRAD, I.A. LETIA, and I. Alfred. Cloud Motion Detection from Infrared Satellite Images. In Second International Conference on Image and Graphics, pages 408–412.
- [26] N. OHTA. Optical flow detection by color images. NEC research & development, (97):78–84, 1990.
- [27] 项学智,赵春晖,李康.一种彩色光流场估计算法.哈尔滨工程大学学报, 29(006):604-609, 2008.
- [28] 陈震,高满屯. 机器人视觉中彩色时变图像光流场计算的综述. 机器人, 23(006):559-562, 2001.
- [29] 项学智,赵春晖. 色彩梯度恒常性的光流场估计算法. 哈尔滨工程大学学报, page 04, 2008.
- [30] L. Xu, J. Chen, and J. Jia. A segmentation based variational model for accurate optical flow estimation. European Conference on Computer Vision: Part I, pages 671–684, 2008.
- [31] N. Papenberg, A. Bruhn, T. Brox, and J. Weickert. Numerical Justification for Multiresolution Optical Flow Computation, 2003.

- [32] P. Anandan. A computational framework and an algorithm for the measurement of visual motion. International Journal of Computer Vision, 2(3):283– 310, 1989.
- [33] M.J. Black and P. Anandan. The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields. Computer Vision and Image Understanding, 63(1):75–104, 1996.
- [34] E. Mémin and P. Pérez. Hierarchical estimation and segmentation of dense motion fields. International Journal of Computer Vision, 46(2):129–155, 2002.
- [35] P. Burt and E. Adelson. The Laplacian pyramid as a compact image code. Communications, IEEE Transactions, 31(4):532–540, 1983.
- [36] C. Cassisa, S. Simoens, and V. Prinet. Two-Frame Optical Flow Formulation in an Unwarping Multiresolution Scheme. Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications, pages 790–797, 2009.
- [37] 汤峰峰. 高分辨率遥感图像的结构变化检测. [硕士学位论文]. 北京: 中国科 学院自动化研究所, 2007.
- [38] S. Li. Markov random field models in computer vision. European Conference on Computer Vision, pages 361–370, 1994.
- [39] 李旭超,朱善安.图像分割中的马尔可夫随机场方法综述.中国图象图形学报,12(005):789-798,2007.
- [40] S. Geman, D. Geman, K. Abend, TJ Harley, and LN Kanal. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images*. Journal of Applied Statistics, 20(5):25–62, 1993.
- [41] J.M. Hammersley and P. Clifford. Markov fields on finite graphs and lattices. Unpublished manuscript, 3, 1971.
- [42] J. Pearl. Probabilistic reasoning in intelligent systems: networks of plausible inference. Morgan Kaufmann, 1988.

- [43] Y. Weiss. Belief propagation and revision in networks with loops. CBCL-155, 1997.
- [44] K. Murphy, Y. Weiss, and M.I. Jordan. Loopy belief propagation for approximate inference: An empirical study. In Proceedings of Uncertainty in Artifical Intelligence, pages 467–475, 1999.
- [45] Pedro F. Felzenszwalb and Daniel P. Huttenlocher. Efficient belief propagation for early vision. International Journal of Computer Vision, 70(1):41–54, 2006.
- [46] J.S. Yedidia, W.T. Freeman, and Y. Weiss. Generalized belief propagation. Neural Information Processing Systems, 2001.
- [47] J.S. Yedidia, W.T. Freeman, and Y. Weiss. Understanding belief propagation and its generalizations. Exploring artificial intelligence in the new millennium, 8:236–239, 2003.
- [48] Besag J. On the statistical analysis of dirty pictures. Journal of the Royal Statistical Society, B 48(3):259–302, 1986.
- [49] S.T. Barnard. Stochastic stereo matching over scale. International Journal of Computer Vision, 3(1):17–32, 1989.
- [50] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(11):1222–1239, 2001.
- [51] V. Kolmogorov and R. Zabih. What energy functions can be minimized via graph cuts? European Conference on Computer Vision, pages 185–208, 2002.
- [52] M. J. Wainwright, T. S. Jaakkola, and A. S. Willsky. Map estimation via agreement on trees: Message-passing and linear-programming approaches. IEEE Transactions on Information Theory, 51(11):3697–3717, 2005.

- [53] V. Kolmogorov. Convergent tree-reweighted message passing for energy minimization. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 28(10), 2006.
- [54] M. F. Tappen and W. T. Freeman. Comparison of graph cuts with belief propagation for stereo, using identical mrf parameters. Ninth IEEE International Conference on Computer Vision, pages 900–906, 2003.
- [55] R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother. A comparative study of energy minimization methods for markov random fields with smoothness-based priors. IEEE Transactions on Pattern Analysis and Machine Intelligence, 30(6):1068–1080, 2008.
- [56] R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother. A comparative study of energy minimization methods for markov random fields. European Conference on Computer Vision, pages 16–29, 2006.
- [57] S. A. Billings, H. L. Wei, and M. A. Balikhin. Generalized multiscale radial basis function networks. Neural Networks, 2007.
- [58] Y. J. Qu and B.G. Hu. Rbf networks for nonlinear models subject to linear constraints. IEEE conference for granular computing, 2009.

附录A 光流评估方法介绍

在这里我们给出常用的光流评估方法,并对这些方法做简单的介绍。常用 到的量化的光流评估方法有平均角度误差AAE(Average Angle Error)、角误差的 标准差STDAE(Standard Deviation of Angle Error)、平均幅值误差AEP(average end-point)、幅值误差的标准差STDEP(standard deviation of end-point),另外 还有角度误差AE(Angle Error),它通常用于直观的给出所求得光流场与光流场 真值相应位置的角度差别。

• 角度误差AE(Angle Error)

角度误差常用于直观的给出光流估计结果的好或者差。其计算方法为,对 图像每一点计算其估计的速度与光流真值速度的夹角,夹角越小表示角度 结果越准确。通常会将计算好的结果用图像表示。角度误差的计算公式如 下,

$$AE = \arccos(\frac{\mathbf{w}_e(\mathbf{s}) \cdot \mathbf{w}_{gt}(\mathbf{s})}{\|\mathbf{w}_e(\mathbf{s})\| \|\mathbf{w}_{gt}(\mathbf{s})\|})$$
(5-1)

这里, s为图像位置坐标, $\mathbf{w} = (u, v, 1)$, $\mathbf{w}_e(\mathbf{s}) = \mathbf{w}_{gt}(\mathbf{s})$ 分别表示估计的速度与光流速度真值, AE为估计速度 \mathbf{w}_e 与光流真值速度 \mathbf{w}_{gt} 之间的夹角。

 平均角度误差AAE(Average Angle Error)
平均角度误差是一种重要的量化评估光流估计结果的方法。设图像任意 点s的角度误差为AE(s),

$$AAE = \frac{\sum_{\mathbf{s}\in C_1} AE(\mathbf{s})}{N} \tag{5-2}$$

式中,s为图像位置坐标, C_1 为图像单点位置集合,N为 C_1 大小,AAE为整个 C_1 上的平均角度误差。

• 角误差的标准差STDAE(Standard Deviation of Angle Error)

角误差的标准差也是一种重要的光流估计结果评估方法,它和AAE将一起 被用于我们的结果评估。设图像任意点s的角度误差为AE(s),

$$STDAE = \sqrt{\frac{\sum_{\mathbf{s}\in C_1} (AE(\mathbf{s}) - AE)^2}{N}}$$
(5-3)

平均幅值误差AEP(average end-point)
平均幅值误差用于评估估计速度幅值与光流真值幅值的差别,其表达式如下,

$$AEP = \frac{\sum_{\mathbf{s} \in C_1} \|\mathbf{w}_{gt}(\mathbf{s}) - \mathbf{w}_e(\mathbf{s})\|}{N}$$
(5-4)

式中, s为图像位置坐标, C_1 为图像单点位置集合, $N \Rightarrow C_1$ 大小, $\mathbf{w}_e = \mathbf{w}_{gt}$ 分别表示估计的速度与光流速度真值, AEP为平均幅值误差。

• 幅值误差的标准差STDEP(standard deviation of end-point)

$$STDEP = \frac{\sum_{\mathbf{s}\in C_1} (\|\mathbf{w}_{gt}(\mathbf{s}) - \mathbf{w}_e(\mathbf{s})\| - AEP)^2}{N}$$
(5-5)

式中, s为图像位置坐标, C₁为图像单点位置集合, N为C₁大小, w_e与w_{gt}分别表示估计的速度与光流速度真值, AEP为平均幅值误差, STDEP为幅值误差的标准差。

附录B 图像数据介绍

本 文 实 验 使 用 的 数 据 来 源 于middlebury的 网 站(http://vision.middlebury.edu/flow/data/)、 真 实 图 像 序 列 等 。 这 些 图 像数据以及相应的光流真值如图5-1,5-2,5-3,5-4。在图 5-1中,云层的运动为从 左向右,山体的运动为向外运动,表现为发散运动。在图 5-2中,他包含两种运 动:匀速向右的背景运动与物体的旋转运动。在图 5-4中,他包含的内容为真实 的流运动。它是从实验室获取 [3]:首先,对充满水的盒子加热,并且使得盒子 的不同区域受热不同,以产生对流;其次,在盒子里放入小的颗粒,这样在水 的对流中,小的颗粒会产生运动。这样就可以对这种运动进行拍摄,获取流运 动的图像序列了。在图 5-3中给出了彩色-速度的映射图。在这个图中,不同颜 色的坐标即为其对应的速度的两个分量。



(a) Yosemite图像
(b) 光流真值
图 5-1 Yosemite图像序列及其矢量光流场(源于 [2])



(a) Hydrangea图像序列

(b) 光流真值

图 5-2 Hydrangea 图像序列及其彩图光流场(源于middlebury网站)。图 5.2(b)的速度可以通过 5-3查得。



图 5-3 彩色速度矢量表。图中不同颜色的坐标即为速度矢量,这样可以使得从彩色图 域到速度矢量域的映射转换



图 5-4 流运动图像序列(源于 [3])

个人简历及攻读硕士期间发表的论文

马东民,男,汉族,中共党员,1983年3月29日出生于河北新河,2003年 毕业于河北新河振堂中学,同年9月进入河北工业大学自动化系学习,并 于2007年7月毕业,获工学学士学位。2007年9月至2008年7月,在中国科学院研 究生院学习研究生基础课程。2008年7月至2010年6月在中国科学院自动化研究 所模式识别国家重点实验室攻读硕士学位。主要研究方向为光流估计。

攻读硕士期间发表的论文:

 Dongmin Ma, Veronique Prinet and Cyril Cassisa, "Statistical Modeling of Optical Flow", ICIG2009, Xi'an, September. 2009.

致 谢

在中科院自动化所攻读硕士学位的三年里,自动化研究所与模式识别国家 重点实验室、中法联合实验室为我的学习和科研工作提供了良好的环境和资 助。在导师和同学的帮助下,我的研究工作得以顺利的展开。感谢他们,是他 们共同创造了一个良好的学习氛围,能和这些志同道合的朋友们愉快共事,对 我而言是一段非常宝贵而且难忘的人生经历,更何况得到了他们大量热情无私 的帮助。

在完成论文之际,首先感谢我的导师普林特。感谢她在我三年研究生的学 习和工作中,对我的指导、关怀和帮助。她对学术研究的严谨执着,对工作的 认真负责,已经对学术研究的敏锐眼光和对工作的热情都深深的影响了我。从 她那里我不仅获得了科学知识、学习研究方法,还领悟到了生活和工作的意 义。这些都令我终生受益。

感谢刘成林研究员、赵思雄研究员、Thomas Corpetti副研究员对我论文的 指导与帮助。

感谢我的父母,为了我能按时完成学业,他们省吃俭用,拖着病老的身体 终日在田地里劳作。还要感谢我的姐姐,感谢她和我的父母一直在背后对我的 支持和鼓励。

感谢我们实验室的何亮亮、喜力、瞿亚军、袁飞、李尔,尤其感谢何亮 亮、喜力、瞿亚军,谢谢他们对我的程序的贡献与帮助。

感谢模式识别国家重点实验室和中法联合实验室的各位老师和同学,是他 们创造了一个良好的科研学习氛围。

感谢自动化所的各位领导,感谢研究生部的各位老师,谢谢他们为我们提 供的学习和生活上的帮助

2010年5月