

文章编号:1001-9081(2009)12-3310-04

## 改进的序列图像超分辨率重建方法

袁 琦,荆树旭

(长安大学 信息工程学院, 西安 710064)

(angeleeyuan@gmail.com)

**摘要:**为解决局部运动的视频序列在超分辨率重建过程中,由于采用传统的图像间全图一致变换模型可能导致的运动估计误差增大,影响重建效果的问题,提出了基于三角网不规则分块运动估计思想和基于 DTN-POCS 的重建算法。从边缘点中提取特征点,利用配准获得的同名点集,在主/从图像对间构建同名不规则三角网,并以此分割图像,默认每个三角块中像素运动一致。在此基础上,把所有低分辨率帧分块投影到高分辨率坐标,以凸集投影(POCS)迭代优化。试验结果表明:重建的高分辨率图对图像中的局部运动鲁棒性更强,能有效改善重建精度。

**关键词:**超分辨率重建; 视频序列; 不规则三角网; 运动估计; 凸集投影

中图分类号: TP391.4 文献标志码:A

## Improved super resolution reconstruction method for video sequence

YUAN Qi, JING Shu-xu

(School of Information Engineering, Chang'an University, Xi'an Shaanxi 710064, China)

**Abstract:** To decrease the negative effect caused by local shift motions, this paper proposed an approach based on the triangulated irregular block motion estimation and DTN-POCS super resolution reconstruction. First of all, select feature points from edges, obtain matching points set between image pairs from the motion estimation process; after that, for each image pair, creates the Delaunay triangulation net for slave image based on the matching feature points, and also creates its corresponding triangular net for the master image; and then, suppose the homologous triangulations between images have the relationship of affine transformation, project all the low resolution images by sub-block onto the high resolution coordinate, generate the initial image by interpolation methods; in the end, iteratively optimize the super resolution image based on the improved method. In the experiment, the method with image uniform affine transformation parameters has less accuracy than that of the improved one. The simulation results testify that the proposed method can improve the reconstruction performance.

**Key words:** Super Resolution Reconstruction (SRR); video sequence; triangulated irregular network; motion estimation; Projection Onto Convex Sets (POCS)

### 0 引言

超分辨率图像重建(Super Resolution Reconstruction, SRR)技术的目的,就是利用多幅图像间的信息互补,消除由成像系统导致的图像退化,同时恢复出光学衍射极限分辨率所决定的截止频率外的信息,形成空间分辨率更高的清晰图像<sup>[1]</sup>。这一技术广泛应用于由于技术限制或经济原因而无法获得理想高分辨率(High Resolution, HR)图像的应用,比如交通视频监控、遥感等。SRR方法通常分为两类:频域法和空域法,其中频域法只能局限于全局平移运动和线性空间不变退化模型,不再成为研究的主流。空域法中常用的有插值法、迭代背投法、凸集投影法、最大似然(Maximum Likelihood, ML)估计、混合ML/MAP/POCS<sup>[1]</sup>,以及自适应滤波、维纳滤波、Kalman<sup>[2]</sup>滤波法等,以Milanfer<sup>[3]</sup>、Elad<sup>[2]</sup>等人为代表的多个机构和个人多年来一直在对SRR方法进行系统的总结和研究,并针对不同应用领域提出改进。此外,针对特定领域如遥感图像等的SRR技术一直是研究热点<sup>[4]</sup>。

亚像素级的运动估计和配准是图像SRR的关键。早期超分辨率算法的研究中,图像配准问题主要针对由镜头运动

引起的全局平移或旋转,不能对视频中存在的局部运动进行有效估计。基于块匹配的运动估计可以提高搜索效率和精度<sup>[5]</sup>,但是块尺寸的选择是算法的关键。大的块尺寸难以对细节运动实现准确的估计,而尺寸太小不仅带来运算量的无谓增加,还会造成平坦区域估计效果的严重下降。为此出现了不少改进策略,如文献[6]中引入了多尺度思想,但如何根据图像内容实现块尺度的灵活切换则不易解决,而且运算量也较大。

近年来,针对SRR的分块配准策略逐渐引起关注。如基于分块快速仿射变换算法<sup>[7]</sup>以及分级或多分辨率自适应块的方法<sup>[8]</sup>等,大大提高了重建精度。总的来看,这些方法普遍存在计算量大、占用内存高、系统效率低等问题,同时也没有将配准与后期应用综合考虑,适用性不高。不规则三角网(Triangulated Irregular Network, TIN)技术在对三维纹理的近似上具有优势,因此被广泛应用于图像配准、三维表面重建<sup>[9]</sup>以及动态地形表现等领域<sup>[10]</sup>。本文中引入TIN技术,把TIN的分割与表现思想贯穿整个重建过程,提出了一种基于TIN的序列图像配准及SRR过程。首先,选择边缘点作为特征点,在主/从帧之间配准获得同名点集,每帧图像分别以同

收稿日期:2009-06-24;修回日期:2009-08-18。 基金项目:陕西省自然科学基金资助项目(SJ08F23)。

作者简介:袁琪(1970-),女,云南蒙自人,讲师,博士,主要研究方向:交通信息控制、图像处理、模式识别、数据挖掘; 荆树旭(1978-),男,陕西西安人,讲师,博士,主要研究方向:CAD、CSCW、模式识别。

名点构造 TIN, 在主、从图之间建立同名三角形的仿射关系, 按照此关系将所有图像投影到统一的坐标系中, 采用插值技术估计初始 HR 图, 最后, 基于 Delaunay 三角网 (Delaunay Triangulation Net, DTN) 三角分块建立的仿射关系, 通过 POCS 重建方法迭代优化, 估计出 HR 图像。仿真实验证明, 本文提出的算法在提高图像分辨率的同时能够有效降低前景不同运动轨迹对重建结果的负面影响。

## 1 基于 TIN 的 SRR 方法

### 1.1 SRR 过程和关键

视频图像的 SRR 过程一般分三步<sup>[1]</sup>: 运动估计(配准)可以获得图像间的变换关系, 初始化过程由非均匀像素分布经过插值生成一幅初始 HR 估计图像, 并通过迭代过程(依赖于观测模型)不断优化重建效果。在整个 SRR 过程中有三个影响复原效果的关键因素: 可靠的子像素级运动信息估计最为至关重要; 准确的观测模型能正确描述图像系统及其降质过程; 合理的重建方法能更充分地利用先验知识以恢复出效果更好的高分辨率图像。

下面按照超分辨率重建执行过程, 详细说明本算法对三个关键因素所提出的解决方案。

### 1.2 运动估计

LR 视频序列中, 各帧之间存在不同的偏移。如果偏移量估计能够达到子像素精度, 就能够综合这些 LR 帧获得的图像信息, 并由此重构出 HR 图像。因此, 帧间位移关系的子像素级运动估计(配准)成为能否充分利用 LR 帧信息的关键。

配准过程是这样的: 首先, 在所有参与 SR 图重构的 LR 帧中选择一帧作为参考帧(主帧), 一般选择第一帧; 其他帧则作为从帧, 从帧分别与主帧配准以获得各从帧向主帧坐标系的变换关系。

配准算法中, 基于特征的方法最大的优点是能够将对整个图像进行的分析转化为对图像特征的分析, 从而大大减小图像处理过程的运算量, 降低图像存储, 提高传输效率。其中基于点特征的方法特征数量少, 有利于提高匹配、存储及传输效率, 能较好的适应图像形变和遮挡。每帧中特征点的选择要求能够充分利用帧间差异。考虑到帧内物体的非均匀分布特点, 物体间存在分界线的特点, 本文从边缘点中采用一定的策略非均匀选择特征点。从帧与主帧的配准采用局部累计系数差最小, 确定同名点, 实验证明该配准方法可以达到亚像素级精度。配准过程不是本文的重点, 在此不作详细说明。

### 1.3 初始化

初始化分为两个阶段: 图像的复原和插值。子像素精度的运动估计是保证所重构的 SRR 图像精度的前提条件。

TIN 在纹理表现方面具有独特优势, 合理的三角网划分可使每个三角分块中的绝大部分像素具有相同的运动规律。因此, 本文利用同名点, 为图像构建 TIN, 假设各三角形内的像素具有相同的运动。下面首先对三角网技术作一简单说明。

#### 1.3.1 Delaunay 三角网及其优势

TIN 被视为最基本的一种网格, 是由一组无规则散落在

空间的点组成。这些点各自与其邻近两点相连, 生成三角面片, 它可以适应规则和不规则的分布数据。根据不同区域的平滑陡峭, 可以用大小各异、疏密不同的 TIN 描述互不相同的纹理形状。具有冗余数据少的特点。

DTN 是一类特殊的 TIN。DTN 具有三个特征: 唯一性, 即对于任何点集  $T$ , 其 DTN 是唯一的, 这一性质是 DTN 独有的; 空外接圆准则, 即对于点集  $P$  为顶点集的 DTN 中每个三角形的外接圆内均不包含点集  $P$  中其他任意点; 最大最小角规则, 三角形的最小内角是相对其他可能的三角网中最大的, 这一准则保证了两个相邻三角形构成的凸四边形的对角线在相互交换后, 6 个内角中的最小者不会增大。以上特征保证了图像可以方便地实现局部插值。

文献[11]将 DTN 用于序列图像的超分辨率重建过程, 利用具有噪声和模糊的图像序列, 采用基于 DTN 的局部插值方法估计初始 SR 图。证实从非一致性采样空间获得一致性采样的过程中, 基于 DTN 的插值方法具有低计算复杂性、可并行执行和局部调整的优势, 最重要的是, 这一方法适合在硬件上实现, 可以用于对实时图像序列的处理。

#### 1.3.2 基于 DTN 的 SR 图像初始化

假定图像配准精度达到亚像素级, 如前所述, 本文中的特征点选自边缘点, 经过配准从中获得同名点集, 合理选择特征点分布密度, 构造出的 TIN 网可以将绝大多数的同类点(前景点或背景点)分割在一个三角形中, 并且认为每个三角形中像素符合一致的几何关系如仿射变换, 少数误分割的像素对计算变换参数及复原图像影响不大。

利用配准结果构建同名三角形的过程是这样的: 分别在主、从图中, 利用配准得到的同名点集构造三角形, 其中, 为从图构造 DTN, 为主图构造对应的 TIN, 构造规则应保证从图中的每个三角形在主图中有且只有一个三角形与之对应, 其中主图中同名三角形的三个顶点分别是从图中对应三角形三个顶点的同名点, 这里称为同名三角形, 同名三角形之间采用简单仿射模型计算相对运动关系。这一策略在简化问题的同时, 也能保证精度要求。

在构建同名三角形的基础上生成 SR 图, 主要初始化处理步骤可以总结如下。

1) 假定由 HR 图经过降采样并加噪获得主图, 对于主图中的每个像素, 其坐标按尺度因子放大, 得到其在 HR 估计图中的坐标。

2) 假定从图由 HR 图经过几何变换、降采样并加噪获得, 则当明确了主、从图间的几何变换关系, 从图中的所有像素可以通过反变换、尺度放大估计得到其在 HR 图中的坐标。只不过在我们的算法中, 每个像素根据它所在三角形遵循不同的变换参数。多数情况下, 变换后的坐标都不在整数坐标位置。

3) 按照以上方法把所有的 LR 图像素全部投影到主图坐标系后, 得到一幅像素非均匀分布的采样空间点集, 用双线形插值(Bilinear Interpolation, BI) 算法获得在每个整数坐标位有值得分布均匀采样空间, 作为初始 HR 估计图。

#### 1.4 基于 DTN\_POCS 的迭代优化

我们知道,SRR 是一个病态的逆问题。在图像采样过程高频部分损失和图像降质是不可能仅仅利用图像插值的方法恢复。因此,图像插值常被看作是 SRR 方法的一种,不是真正意义上的重建技术。

要想进一步改善重建效果,最便捷的方法就是引入先验知识。比如,利用针对同一景物的,相互间具有子像素位移的多帧图像间的信息互补,以获得更多关于场景的信息。在不同 SRR 算法中,基于凸集投影(Projection Onto Convex Sets, POCS)的方法具有均方误差最小的优势,因此很容易在算法中增加新的限制条件。将 POCS 方法引入超分辨率重构过程,增加限定条件约束集合,即规定在对源 HR 图作了模糊和降采样后,得到的图像必须包含在观测得到的 LR 图中,可以保证数据的连续性,并由此得到简化的解空间。本文利用 POCS 方法并加以改进,用一段视频复原出一幅更高分辨率的图像。

##### 1.4.1 基于三角网的 POCS 算法

传统基于 POCS 的重建算法处理过程中采用主从图之间全图一致的运动变换关系。对于存在局部运动的视频,这一处理存在缺陷。本文基于 DTN 分块,对 POCS 的执行过程算法,这里将这一算法称为 DTN-POCS 算法。算法还利用边缘检测结果,采用 POCS 和局部残差均方差计算对 HR 图边缘做了进一步细化。

局部细化过程针对所有边缘点。边缘点的确定根据对 LR 序列边缘检测。当把 LR 图像素投影到 HR 坐标时,对原 LR 中的边缘点做标注。初始化产生 HR 估计图后,在 HR 每个整数坐标像素的单位像素邻域内,搜索是否存在标注的边缘点,存在,则认为该 HR 像素属于边缘点,需要优化。

令  $f$  表示放大因子,对 HR 内每个边缘点,计算图像  $(2 \times f + 1) \times (2 \times f + 1)$  窗口与各 LR 的残差均方差,如低于阈值,则在下一次循环中不再处理此点;否则将残差结果累积到 HR 上,在下一循环时再次计算。

##### 1.4.2 算法的实现

算法的实现过程如下:

- 1) 对于每个 LR 帧,根据 TIN 分割,以所在三角形的变换参数并结合尺度因子得到其在 HR 中的同名点坐标。待所有帧均变换到 HR 图空间后,以不规则插值得到 HR 图,在 HR 图中分别构建与各 LR 对应的同名三角网,计算变换参数。

- 2) 采用假设的点扩散函数,由 HR 生成低分辨率图序列;其中 HR 中像素到 LR 的变换按照其所在同名三角形的变换关系计算。

- 3) 计算 HR 与各 LR 的残差。当所有残差的均方差都小于指定阈值或循环次数超限,退出循环;执行 5)。

- 4) 所有 LR 的残差结果分别累积到 HR 上,得到新的 HR 估计,继续循环。

- 5) 利用 LR 边缘检测结果,对 HR 图边缘局部修正,修正方法见第 1.4.1 节,当所有边缘点都修正到残差均方差小于阈值或循环次数超限,退出循环。

其中:1)~4) 为迭代优化过程,5) 为细化过程。

## 2 试验结果及评价

### 2.1 试验说明

为了证实所提出方法的可行性,本文采用图像序列“Foreman”中连续 4 帧,来构建一帧 HR 图。其中每帧尺寸为  $144 \times 176$ ,重建的尺度因子为 2,即放大一倍。实验利用 Matlab2007 编程,运行环境 Windows XP PC (CPU 为 Intel Core2 2.6 GHz, RAM 为 2.0 GB)。

理论上,希望 HR 估计与实际之间的残差修正到 0,实际上做不到的,需要设定上限阈值。

残差均方差阈值的设定直接影响到图像重建效果。设置太高,不利于图像细节恢复,难以获得好的图像;设置太低,容易造成噪声。本实验参照文献[12],阈值采用  $10^{-3}$ 。为防止残差的均方差无法收敛而造成执行陷入死循环,限定迭代优化过程迭代次数在 3~40;细化阶段需要针对每个边缘像素处理,比较费时,因此限定迭代次数在 3~15。

为了比较算法的有效性,本文还给出另外三种方法的试验结果作比较。

第一种方法为单帧图像简单插值算法。从 LR 帧中提取一帧,以双线性插值方法提高图像分辨率一倍。前面提过,简单插值方法并不是真正意义上的 SR 重建方法,这里只是作为效果的参照。简称方法 1。本文选择第 30 帧即图 1(a)中的第一帧插值。

方法二在图像间的变换采用全图一致仿射参数,首先采用本文提出的配准方法找到同名点集,认为从图所有像素运动一致,由同名点集分别计算每幅从图到主图的变换参数,接着将所有 LR 图中的像素按照这一参数变换到 HR 图中,用双线性插值算法得到均匀采样的初始 HR 估计图,再利用 POCS 算法迭代优化,简称为方法 2。

第三种方法则引入分块运动估计策略:采用等大小块分割和新三步搜索法实现块运动估计其中块大小为  $8 \times 8$ ,行、列最后不足 8 像素的归到前一块,默认每个块内的像素的运动一致。重建仍采用 POCS,只是每个  $8 \times 8$  子块遵循各自的变换参数。简称方法 3。

### 2.2 实验结果及评价

下面来分析比较实验结果。实验从序列图像中抽取连续 4 帧。

图 1(a)是获取的 LR 帧,从左至右分别是视频序列中第 30~33 帧。图 1(b)~(e)分别是三个对比方法和本文方法的实验结果(右侧是实验结果的局部,用来从视觉结果分析不同方法的重建效果)。

从图 1(a)可以看出四 LR 帧之间存在全局运动,即图像背景整体向右移动;同时存在局部运动,即人物向左移动;同时注意到在第 32 帧(即四帧中的第三帧)存在一个人物闭眼的动作。

对比三个重建结果,方法 2 的重建结果图像的左侧有明显的黑边,这是由于采用全局一致的变换模型,当计算运动参数时,参与计算的点集包含了分别属于前、背景的点,使得获得的运动参数对于前、背景像素的运动描述均不正确,导致重

建误差。

方法 3 由于采用了分块运动估计策略,当某一块中包含的绝大部分像素同属于一类时,重建效果明显好于方法 2,但从细部观察可以看到在前、背景交界部分(如帽檐部分),由于子块中包含两类像素,因此运动估计还是有很大误差,影响了重建精度。

本文采用的方法以边缘点作为特征点,因此构建 TIN 时,前、背景像素更容易被分在不同三角形中,加之本文提出的重建思想使得不同三角形中的像素遵循各自的变换参数,进一步提高了重建的准确性。因此本文方法重建结果基本没有黑边,前、背景交界部分重建效果改善明显。

再来比较各算法获得的细节特征。由复原结果中的建筑边缘分析,方法 1 由于针对单帧视频图像,复原过程不受配准精度影响,图像边缘相对完整;也同样是由于这一方法针对单帧图像,图像能够提供的细节信息不充分,结果图像的轮廓比较模糊;而三种重建算法能够得到更为清晰的图像轮廓。同时,由于前、背景运动不一致,在三种复原结果中,人物的帽檐均出现模糊,在三幅复原结果图中,本文提出的方法帽檐更清晰,接近单帧插值的效果。这一实验结果进一步证实了本文提出的重建方法具有更高的重建精度。

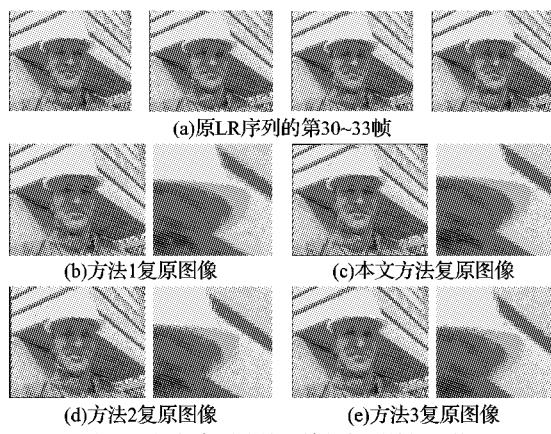


图 1 视频序列连续四帧的复原结果比较

从细节还可以看到,由于在 32 帧人物存在闭眼的动作,因此在恢复的结果中人物眼睛的复原效果并不理想。这是由于本文选择的特征对于这一类动作捕获能力不足。

以上是从直观表现分析了不同重建方法的优缺点,下面从统计学角度对三种方法做定量分析。

表 1 给出了插值方法和三种重建方法的峰值信噪比 PSNR。

表 1 不同方法的 PSNR 结果比较

方法	PSNR	方法	PSNR
本文方法	13.345 0	方法 2	13.183 3
方法 1	13.231 3	方法 3	13.216 4

可以看出,本文提出方法 PSNR 最好,甚至高于基于插值的复原方法,而作为对比的两种重建方法由于存在比较大的边缘模糊,PSNR 都比较低,其中基于以全局一致仿射变换得到的重建结果为最差,说明亚像素级的运动估计对图像的重建的确至关重要。同时,视频序列中,合理提取特征点对分割图像前背景,进而获取正确的运动信息,也是提高图像超分辨率重建精度的关键。

### 3 结语

本文采用基于边缘点的特征选择策略,在此基础上提出了一种基于 DTN 分块估计运动参数和 POCS 迭代优化的超分辨率重建思想。实验和统计对比分析证实,对于前背景运动不一致的视频序列,这一方法的优势在于直接利用特征点构建三角网,能根据图像纹理分布自动调整分块密度,可以有效降低系统负担,提高效率;同时,采用三角网更有利图像纹理描述,改善重建效果。这一优势使所提出的方法可直接用于存在远程传输和压缩需要的应用中。

本文提出的方法中,特征点分布密度对于重建效果有一定的影响。确定出合适的特征点选择密度是本文下一步工作的重点。同时,本文的三角分割思想也适合并行处理,为实时视频处理以及大规模图像处理提供了可能,这也将作为下一步的研究方向。

### 参考文献:

- [1] PARK S C, PARK M K, KANG M G. Super-resolution image reconstruction: A technical overview [J]. IEEE Signal Processing Magazine, 2003, 20(3): 21 - 36.
- [2] Michael Elad [EB/OL]. [2009 - 08 - 18]. <http://www.cs.technion.ac.il/~elad/>.
- [3] Peyman Milanfar [EB/OL]. [2009 - 05 - 11]. <http://www.ee.ucs.edu/~milanfar/>.
- [4] 袁琪,赵荣椿.遥感高光谱图像超分辨率重建[C]//信息与信息处理技术论文集. 大连:信号处理学会, 2007: 89 - 94.
- [5] 禹晶,苏开娜.块运动估计的研究进展[J].中国图象图形学报, 2007, 12(12): 2032 - 2041.
- [6] 王素玉,卓力,沈兰荪,等.一种用于动态视频超分辨率的多尺度最小二乘仿射块匹配图像配准方法[J].电子与信息学报, 2009, 31(3): 542 - 545.
- [7] PATANAVIJIT V, SERMWUTHISARNTT P. A robust iterative super-resolution reconstruction of image sequences using a Turkey's bi-weight Bayesian approach with fast affine block-based registration [C]// ICME: IEEE International Conference on Multimedia & Expo. Beijing: IEEE, 2007: 480 - 483.
- [8] 沈兰荪,卓力.小波编码与网络视频传输[M].北京:科学出版社, 2005: 146 - 164.
- [9] YU LE, ZHANG DENG-RONG, HOLDEN E-J. A fast and fully automatic registration approach based on point features for multi-source remote-sensing images [J]. Computers & Geosciences, 2008, 34(7): 838 - 848.
- [10] AGOSTINI V, DELSANTO S, KNAFLITZ M. et al. Noise estimation in infrared image sequences: A tool for the quantitative evaluation of the effectiveness of registration algorithms [J]. IEEE Transactions on Biomedical Engineering, 2008, 55(7): 1917 - 1920.
- [11] LERTRATTANAPANICH S, BOSE N K. High resolution image formation from low resolution frames using Delaunay triangulation [J]. IEEE Transactions on Image Processing, 2002, 11(12): 1427 - 1441.
- [12] Laboratory of audiovisual communications (LCAV) [EB/OL]. [2009 - 05 - 20]. <http://lcavwww.epfl.ch/software/superresolution>.