

# 联合运动估计与基于模式的超分辨率图像重构

张 地<sup>1,2</sup>, 彭 宏<sup>1</sup>

(1. 华南理工大学计算机科学与工程学院, 广东广州 510641; 2. 韶关学院计算机科学系, 广东韶关 512005)

**摘 要:** 超分辨率图像重构是利用关于同一场景的多帧低分辨率图像重构出一幅具有更高分辨率图像的过程。已有的超分辨率图像重构算法对于人工模拟所得到的低分辨率图像序列具有很好的效果, 但对于拍摄到的真实低分辨率图像序列而言, 重构后的图像往往比较模糊, 有时甚至仍然无法分辨。为此, 本文提出了一个联合运动估计与基于模式的超分辨率图像重构算法。实验结果表明, 该算法能够得到优于常规算法的高分辨率图像。

**关键词:** 超分辨率; 模式识别; 图像序列; 图像重构

中图分类号: TN911.73 文献标识码: A 文章编号: 0372-2112(2008)01-0180-04

## Joint Image Registration and Pattern Based High Resolution Image Reconstruction

ZHANG Di<sup>1,2</sup>, PENG Hong<sup>1</sup>

(1. College of Computer Science and Engineering, South China University of Technology, Guangzhou, Guangdong 510641, China;

2. Department of Computer Science, Shaoguan University, Shaoguan, Guangdong 512005, China)

**Abstract:** Super resolution image reconstruction is the process of producing a high resolution image from a set of low-resolution images of the same scene. Previously published techniques usually perform well on simulated low resolution image sequences, but the quality of the reconstructed images from real low resolution videos, is always blurring or even unrecognizable. A joint image registration and pattern based high resolution image reconstruction algorithm was thus proposed in this paper. Experimental results demonstrated that the proposed algorithm could produce superior results against normal algorithms.

**Key words:** super resolution; pattern recognition; image sequence; image reconstruction

### 1 引言

传统图像恢复问题是对一幅输入图像进行恢复。如果有一组彼此相似而又不完全相同的图像序列, 它们描述的是同一场景, 各自含有一些不同的亚像素细节, 便可以考虑将它们全体作为输入图像序列, 利用多幅图像所提供的额外信息, 重构出一幅具有更高分辨率的图像, 这个过程就是超分辨率图像重构。通常, 超分辨率图像重构包含两个步骤, 第一个步骤是将各个低分辨率图像进行对准, 也就是通常的运动估计; 第二个步骤是重构过程, 将已经对准的各个低分辨率图像融合为一幅高分辨率图像。在理想情形下, 要放大  $q \times q$  倍, 至多有  $q^2$  帧非冗余(指具有不同的运动矢量, 因而能提供不同的亚像素信息)的低分辨率图像。但对于实际应用而言, 除去一些专用的受控成像系统外, 虽然可获得的低分辨率图像数目可远远多于  $q^2$ , 但大多数是具有相同运动矢量的冗余图像, 基本上不可能获得  $q^2$  帧非冗余的低分

辨率图像。

迄今已提出的超分辨率图像重构算法包括有频域算法, 凸集投影算法<sup>[1]</sup>、最大后验概率(MAP)算法<sup>[2~8]</sup>。其中, 由于 MAP 算法具有完善的理论框架, 灵活多变的运动模型, 方便地引入各种图像先验知识, 并能得到优于其他算法的重构结果等优点而受到普遍关注。但所有这些算法对于模拟所得到的低分辨率图像序列具有很好的放大效果, 但对于实际拍摄到的真实低分辨率图像序列而言, 重构后的图像往往比较模糊, 有时甚至仍然无法分辨。究其原因, 主要有两点: 一是运动估计的病态性, 二是真实低分辨率图像序列所提供的亚像素信息远远不够, 也就是非冗余的低分辨率图像数目远远低于  $q^2$  帧。为此, 本文提出了一个联合运动估计与基于图像模式(本文是基于文本图像)的超分辨率图像重构算法。与已有的算法不同, 该算法先构造一个基于运动矢量和欲求的高分辨率图像的目标函数, 然后同时对运动矢量和高分辨率图像进行求解。

## 2 MAP 超分辨率图像重构算法

设有  $M$  帧分辨率为  $S \times S$  的低分辨率图像  $\{y_i, i = 1, 2, \dots, M\}$ , 用向量  $y$  来表示. 假设所有这些低分辨率图像都是由一幅分辨率为  $qS \times qS$  的高分辨率图像经过平移、模糊和降抽样后所产生, 并用向量  $z$  来表示该高分辨率图像. 则通用超分辨率图像重构模型可以表示为:

$$y = Wz + n \quad (1)$$

式(1)中  $y$  和  $z$  分别是  $MS^2 \times 1$  和  $q^2S^2 \times 1$  的向量,  $W$  是一  $MS^2 \times q^2S^2$  矩阵, 表示作用于  $z$  上的平移、模糊及亚抽样算子,  $n$  是零均值的白高斯噪声, 其概率密度函数具有以下形式:

$$\Pr(n) = \frac{1}{(2\pi)^{p/2} \sigma_n^p} \exp\left[-\frac{1}{2\sigma_n^2} n^T n\right] \quad (2)$$

式中  $p = MS^2$ . 高分辨率图像  $z$  的 MAP 估计值由下式决定:

$$\arg \max_z \{\log \Pr(z|y)\} \quad (3)$$

利用贝叶斯法则  $\Pr(z|y) = \Pr(y|z)\Pr(z)/\Pr(y)$ , 式(3)可写成:

$$\arg \min_z \{-\log \Pr(y|z) - \log \Pr(z)\} \quad (4)$$

由式(1)及式(2), 可得式(4)的第一项为一平方函数, 如下式所示:

$$-\log \Pr(y|z) = \frac{1}{2\sigma_n^2} \|y - Wz\|^2 \quad (5)$$

将式(5)代入式(4), MAP 超分辨率图像重构算法归结为求下面函数的极小值:

$$E(z) = \frac{1}{2\sigma_n^2} \|y - Wz\|^2 + \log \Pr(z) \quad (6)$$

## 3 新算法

### 3.1 联合 Bayes 运动估计

在已有的重构算法中, 矩阵  $W$  是通过计算各帧低分辨率图像之间的、精确至亚像素精度的运动矢量来预先确定. 然而, 由于运动估计本身的病态性, 除去受控成像系统或人工模拟的情形(即低分辨率图像序列是通过某一高分辨率图像按事先设定的方案进行平移、模糊和亚抽样所生成), 并不能确保对实际图像序列进行估计后所获得的运动矢量是准确可靠的. 因此, 对于超分辨率图像重构来讲, 更应从概率的角度来看待运动估计. 也就是说, 当给定一系列低分辨率图像, 不仅仅要重构出最可能的高分辨率图像, 还要确定每帧低分辨率图像最可能的运动矢量, 也就是:

$$\arg \max_r \{\log \Pr(z, r|y)\} \quad (7)$$

上式中  $r$  为运动矢量. 将式(1)重新写成以下的形式:

$$y = W(r)z + n \quad (8)$$

$W(r)$  表明  $W$  是  $r$  的函数. 采用与上一节类似的推导过程, 可得到包含有高分辨率图像  $z$  和运动矢量  $r$  的目标函数为:

$$E(z, r) = \frac{1}{2\sigma_n^2} \|y - W(r)z\|^2 + \log \Pr(z, r) \quad (9)$$

### 3.2 基于模式的图像先验模型

一般将式(9)右边第二项选为不依赖于  $r$  的马尔可夫随机场模型. 在这里, 我们将之选择为一种既依赖于  $r$ , 又依赖于图像模式的模型. 具体地说, 就是通过模式识别, 将低分辨率图像序列  $\{y_i, i = 1, 2, \dots, M\}$  划分为一系列子集  $\{S_{i,k}, |k = 1, 2, \dots\}$  后, 将该模型定义为:

$$\Pr(z, r) = \prod_i \left( \sum_k \Pr(z, r_i | y_i \in S_{i,k}) \Pr(y_i \in S_{i,k}) \right) \quad (10)$$

上式中  $r_i$  为低分辨率图像  $y_i$  的运动矢量. 一旦有了输入低分辨率图像  $y_i$ , 利用模式识别来确定  $y_i$  属于哪一个子集后,  $\Pr(z, r)$  就退化为求  $\Pr(z, r_i | y_i \in S_{i,k})$ . 为了求出该概率, 我们先给出本文中图像模式识别的定义.

#### 3.2.1 图像模式识别

进行图像模式识别就是通过表征图像内在特性的特征矢量将与低分辨率图像最为相似的高分辨率图像样本找出来, 其目的是为了在重构过程中补充低分辨率图像序列中所欠缺的亚像素信息, 这对于无法获得足够多的非冗余的低分辨率图像的大多数实际应用来讲尤其重要. 由于  $y_i$  可以看成是  $z$  的降采样版本, 受到文献[3, 9]的启发, 则可以对  $z$  不断地进行降采样以形成一个图像的降采样金字塔  $P_0(z), \dots, P_N(z)$ . 原始图像  $z$  位于该金字塔的第 0 层, 金字塔第  $j$  层的图像  $P_j(z)$  是  $z$  在  $x$  轴和  $y$  轴都进行  $2^j$  倍数降采样后的版本.

随后, 我们再构造该降采样金字塔的水平一阶导数  $H_0(I), \dots, H_N(I)$ , 垂直一阶倒数  $V_0(I), \dots, V_N(I)$  和对角一阶倒数  $DX_0(z), \dots, DX_N(z), DY_0(z), \dots, DY_N(z)$  (图 1 标出了这四个方向上的一阶倒数). 我们将对应于图像  $z$  的特征金字塔定义为:

$$F_j(z) = (H_j(z), V_j(z), DX_j(z), DY_j(z)), \quad \text{for } j = 0, 1, \dots, N \quad (11)$$

特征金字塔中第  $j$  层像素  $(m, n)$  的识别特征 (RF) 矢量定义为:

$$RF_j(z)(m, n) = (F_j(z)(m, n), F_{j+1}(z)\left(\left[\frac{m}{2}, \frac{n}{2}\right]\right), \dots, F_N(z)\left(\left[\frac{m}{2^N}, \frac{n}{2^N}\right]\right)) \quad (12)$$

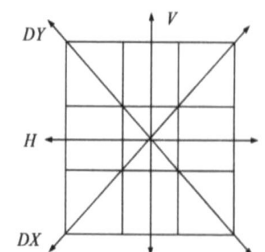


图 1 四个方向上的一阶倒数

设有一帧低分辨率图像  $y_i$  和由  $B$  帧高分辨率训练样本图像  $Z_j(j=1, \dots, B)$  所组成的训练集, 且低分辨率图像的分辨率比训练样本低  $2^k$ . 于是对于训练样本, 可以构造出从 0 层到  $N$  层的特征金字塔, 即  $F_0(Z_j), \dots, F_N(Z_j)$ ; 对于低分辨率图像, 可以构造出从  $k$  层到  $N$  层的特征金字塔, 即  $F_k(y_i), \dots, F_N(y_i)$ .

将低分辨率图像  $y_i$  中每一像素  $(m, n)$  的识别特征矢量  $RF_k(y_i)(m, n)$  与训练样本中处于同一  $k$  层的所有像素的识别特征矢量  $RF_k(Z_j)(p, q)$  进行比较, 得到最佳匹配图像  $IM_i(m, n) = j$  和最佳匹配像素  $PI_i(m, n) = (p, q)$ . 由于此时的最佳匹配像素仅位于训练样本中的最佳匹配图像  $Z_{IM_i(m, n)}$  的特征金字塔的第  $k$  层上, 要得到位于特征金字塔第 0 层上的最佳匹配像素, 仅需将  $PI_i(m, n)$  乘以  $2^k$  即可. 为了标记上的方便, 将最佳匹配图像和特征金字塔第 0 层上的最佳匹配像素重新定义为:

$$\begin{aligned} \overline{PI}_i(m, n) &\equiv 2^k PI_i \left( \left[ \frac{m}{2^k} \right], \left[ \frac{n}{2^k} \right] \right) \\ \overline{IM}_i(m, n) &\equiv IM_i \left( \left[ \frac{m}{2^k} \right], \left[ \frac{n}{2^k} \right] \right) \end{aligned} \quad (13)$$

### 3.2.2 图像先验模型

若  $(m, n)$  是低分辨率图像  $y_i$  特征金字塔第 0 层上的像素, 则不难得到该像素在欲重构的高分辨率图像  $z$  中所对应的像素为:

$$f(r_i) \left( \left[ \frac{m}{2^k} \right], \left[ \frac{n}{2^k} \right] \right) \quad (14)$$

式中  $f(r_i)(\cdot)$  是  $y_i$  到  $z$  的映射函数. 对于  $z$  中的该像素, 我们希望施加这样的约束: 它的水平和垂直一阶导数应该等于训练集最佳匹配图像中最佳匹配像素的水平和垂直一阶导数. 如果进一步假设欲重构的高分辨率图像与训练样本中最佳匹配图像之间的一阶导数的误差服从方差为  $\sigma_\mu^2$  的高斯分布, 有:

$$\begin{aligned} \log \Pr(z, r_i | y_i \in S_{i,k}) &= \frac{1}{2\sigma_\mu^2} \left\{ \sum_{m,n} [H_0(z)(f(r_i) \left( \left[ \frac{m}{2^k} \right], \left[ \frac{n}{2^k} \right] \right)) - H_0(Z_{\overline{IM}_i(m,n)})(\overline{PI}_i(m,n))]^2 + \sum_{m,n} [V_0(z) \right. \\ &\left. (f(r_i) \left( \left[ \frac{m}{2^k} \right], \left[ \frac{n}{2^k} \right] \right)) - V_0(Z_{\overline{IM}_i(m,n)})(\overline{PI}_i(m,n))]^2 \right\} \end{aligned} \quad (15)$$

### 3.2.3 优化算法

将式(15)和式(10)代入式(9), 可得到最终的目标函数  $E(z, r)$ .

为了求得使  $E(z, r)$  取极小值所对应的  $r$  和  $z$ , 我们采用轮流交替对变量  $r$  和  $z$  求极小值的方法. 具体地说, 就是令  $z^{n-1}$  和  $r^{n-1}$  为第  $n-1$  次迭代后的值; 在第  $n$  次迭代时, 先令  $r = r^{n-1}$ , 代入  $E(z, r)$  后有:

$$E(z, r) = E(z, r^{n-1}) \equiv E_1(z) \quad (16)$$

求解  $E_1(z)$ , 可采用 GD、CG 等数值算法. 当采用 GD 算法时, 以  $z^{n-1}$  为初始值, 有<sup>[4]</sup>:

$$z^n = z^{n-1} - \varepsilon \nabla_z E_1(z) \quad (17)$$

再令  $z = z^n$ , 代入  $E(z, r)$  得:

$$E(z, r) = E(z^n, r) \equiv E_2(r) \quad (18)$$

当输入低分辨率图像序列中含有局部位移时, 必须利用数值算法来求解  $E_2(r)$ . 但当仅含有全局性位移时, 就不必采用复杂的数值算法, 而可以利用一般的搜索算法: 即以  $r^{n-1}$  为基点, 搜索使  $E_2(r)$  取极小值的  $r^n$ . 在取得  $z^n$  和  $r^n$  后, 重复上述迭代过程直到收敛条件得到满足.

新算法与传统 MAP 算法相比, 计算复杂性的增加主要在以下两个方面:

(1) 高、低分辨率像素之间的模式匹配: 因用于匹配的识别特征矢量  $RF$  的维数很低, 这一部分的计算复杂性就主要取决于样本空间的大小. 对于本文的应用来讲, 30 帧左右的训练样本已经足够, 由此而增加的计算时间仅相当于迭代一次求解  $E_1(z)$  的时间. 此外, 匹配只需要在迭代前进行一次即可.

(2) 每次迭代时都要进行运动估计: 由于可以采用一般的块匹配算法来进行运动估计, 而块的维数一般远远低于原始图像的维数, 故由此而增加的计算量与迭代时需逐个更新每个图像像素相比, 基本上可以忽略不计, 也就是说求解  $E_2(r)$  的时间与求解  $E_1(z)$  的时间相比, 基本上可以忽略不计.

## 4 实验结果

为了获得训练用的高分辨率文本图像, 我们采用在运行文本处理软件时直接从电脑屏幕上抓取的方法. 在总共抓取 30 帧后, 再全部手工剪切成  $256 \times 256$  像素大小. 为了产生相应的模拟低分辨率图像, 先随机从训练样本中选取一幅, 再对该帧图像进行倍数为 4 的平移、模糊和亚抽样, 并重复上述过程直到产生 16 帧分辨率为  $64 \times 64$  像素、非冗余的低分辨率图像. 针对模拟重构结果, 本文采用 SNR 提升比作为定量比较的依据, 其定义为:

$$\Delta_{\text{SNR}} = 10 \log_{10} \frac{\|z - z_0\|^2}{\|z - \hat{z}\|^2} \quad (19)$$

上式中  $z_0$  是对参考帧进行双线性插值后得到的结果. 所有实验都是在一台配置为 Celeron 2.13GHz CPU 的计算机上进行.

图 2(a) 给出了一帧典型的模拟低分辨率图像, 采用双线性插值法对之放大 4 倍后的结果如图 2(b); 分别采用 MAP 算法和本文提出的算法对模拟得到的全部 16 帧非冗余低分辨率图像进行重构. 所需的重构时间

分别为 1262.3s 和 1448.7s, 重构结果如图 2(c) 和 (d)。从结果可以看出: 虽然新算法比 MAP 算法多用了 14.7% 的时间, 提高了重构图像的定量指标, 但在主观视觉对比上则基本上没有差别。此外, 采用新算法所得到的各帧低分辨率图像的运动矢量与模拟时设定的平移量一致, 这证明联合运动估计与超分辨率图像重构是可行的。我们再对用一个网络摄像头(300,000 像素)所拍摄到的低分辨率“日历”录像序列进行倍数为 4 的超分辨率重构。该录像总共有 20 帧, 其中一帧如图 3(a), 对 (a) 采用双线性插值放大 4 倍后的结果如 (b), 采用 MAP 算法和本文提出的算法所得到的结果分别如图 3(c) 和 (d), 所需要的重构时间分别为 1108.1s 和 1283.5s。可以看出, 与 MAP 算法相比, 新算法在多付出 15.8% 的计算时间后, 不仅能够重构出更明锐的边界, 使得重构出的图像更加清晰; 图 3(c) 中一些仍旧无法分辨的字符, 如图像第二行中的“TUE 二”和“WED 三”, 第四行中的“8”, 在图 3(d) 中都得到了更好地细化。

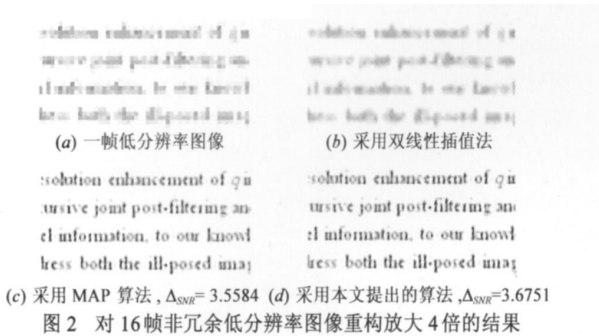


图 2 对 16 帧非冗余低分辨率图像重构放大 4 倍的结果

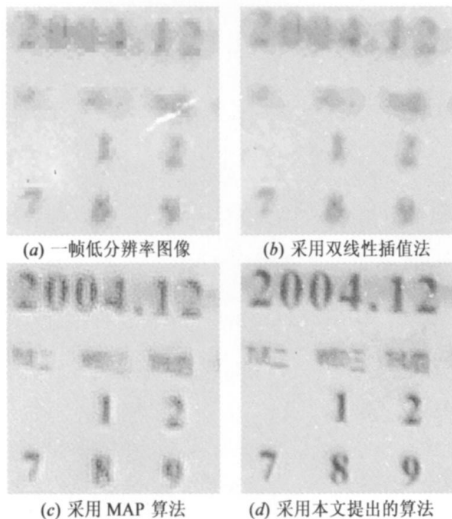


图 3 对实际拍摄到的低分辨率录像序列重构放大 4 倍后的结果

## 5 结论

本文提出了一个联合运动估计与基于图像模式的超分辨率图像重构算法。与已有的算法不同, 该算法先构造一个基于运动矢量和高分辨率图像的目标函数,

然后同时求解运动矢量和高分辨率图像, 使得该目标函数取极小值。需要指出的是, 虽然本文仅对文本图像进行了验证, 但将新算法推广至其他类型的图像, 也是简单和直接的。

## 参考文献:

- [1] A J Patti, Y Altunbasak. Artifact reduction for set theoretic super resolution image reconstruction with edge adaptive constraints and higher order interpolants [J]. IEEE Trans. Image Processing, 2001, 10(1): 179-180.
- [2] N Nguyen, P Milanfar, G Golub. A computationally efficient superresolution image reconstruction algorithm[J]. IEEE Trans Image Processing, 2001, 10(4): 573-583.
- [3] S Baker, T Kanade. Limits on super-resolution and how to break them[J]. IEEE Trans Pattern Analysis and Machine Intelligence, 2002, 24(9): 1167-1183.
- [4] Di Zhang, Minghui Du. Fast hybrid approach to large magnification super resolution image reconstruction[J]. Optical Engineering, 2005, 44(3): 1-9.
- [5] 孟庆武. 预估估计混叠度的 MAP 超分辨率处理算法[J]. 软件学报. 2004, 15(2): 207-214.
- [6] 张新明, 沈兰荪. 基于多尺度边缘保持正则化的超分辨率复原[J]. 软件学报. 2003, 14(6): 1075-1081.
- [7] 苏秉华, 等. 基于 Markov 约束的泊松最大后验概率超分辨率图像复原法[J]. 光子学报. 2002, 31(4): 492-496.
- [8] Di Zhang, Huifang Li, Minghui Du. Fast MAP based multi-frame super resolution image reconstruction[J]. Image and Vision Computing, 2005, 23(7): 671-679.
- [9] P J Burt, E H Adelson. The laplacian pyramid as a compact image code [J]. IEEE Trans Image Communications, 1983, 31(4): 532-540.

## 作者简介:



张 地 男, 1972 年 5 月生于湖南岳阳, 1996 年毕业于武汉大学, 获硕士学位, 2005 年毕业于华南理工大学通信与信息系统专业, 获博士学位, 现为韶关学院计算机科学系教师, 华南理工大学博士后, 主要研究方向为图像处理和模式识别。E-mail: zhangdi@sgu.edu.cn



彭 宏 男, 1956 年 10 月生于重庆, 华南理工大学计算机科学与工程学院教授, 博士生导师, 主要研究方向为智能计算、知识工程。