

# 用神经网络恢复残差的图像超分辨率算法

潘逢治,张立明

(复旦大学电子工程系图像与智能实验室,上海 200433)

**摘 要:** 本文提出了一种基于残差线性恢复的图像超分辨率算法框架,以已有的图像插值算法为基础,利用人工神经网络恢复图像残差,相迭加得到高分辨率图像的估计.在数学上论证了基于神经网络的图像插值算法的依据,并证明了本算法比以往的算法具有更好的性能.理论上,任何一种单帧的图像插值算法都可以被引入本算法框架.同时给出了算法实例和测试结果.

**关键词:** 图像超分辨率; 人工神经网络; 线性恢复; 残差

**中图分类号:** TN911. 73      **文献标识码:** A      **文章编号:** 0372-2112 (2004) 01-0161-05

## Super-Resolution of Images Based on Restoration of Residual Errors by Neural Networks

PAN Feng-zhi, ZHANG Li-ming

(Dept. of Electronic Engineering, Fudan University, Shanghai 200433, China)

**Abstract:** A super-resolution scheme frame based on linear restoration of residual errors is proposed in this paper, which combines an existing image interpolation algorithm with an artificial neural network (ANN) modeling the residual errors between the interpolated image and the real high-resolution image. Mathematical proof is presented that the performance of our method is better than existing interpolation algorithms. Theoretically, any intra-frame interpolation algorithm can be combined into our scheme to improve its performance. Example applications and their results demonstrate the effectiveness of our approach compared with traditional methods.

**Key words:** super-resolution; artificial neural network (ANN); linear restoration; residual error

### 1 引言

图像超分辨率技术通常是指从原始图像获得更高分辨率图像的技术.许多实际应用要求精确的图像放大,例如需要跟踪小目标的监视系统和需要进行病灶识别的医学图像系统,因此图像超分辨率技术一直得到普遍的重视.然而,事实上如果只有一幅低分辨率的图像的信息,要精确地获得高分辨率图像是不可能的,因为人们无法恢复因低采样率而丢失掉的高频信息.同时,如果用各种不同的函数对低分辨率图像进行插值,可以得到无限多个被放大的图像,它们在降采样后等于原低分辨率图像.

常用的图像放大方法是用帧内插值.传统的双线性插值、三次样条插值等方法都因存在上述问题而不能真正地提高分辨率.90年代以后,许多基于新的假设的图像插值算法被提出来.如方向性插值方法<sup>[1]</sup>、多卷积核和非线性滤波法<sup>[2]</sup>、正交变换域方法<sup>[3,4]</sup>和 Bayesian 分析法<sup>[5]</sup>等.由于这些方法所基于的假设与实际情况有一定差距,因此无法保证恢复出高分辨率图像的准确性.多帧插值方法利用带全局运动的多帧低分辨率图像建立一组方程,来解出高分辨率未知像素的灰度值.由于带有亚像素(subpixel)运动的多帧低分辨率图像,其

中每一帧都包含其他帧所没有的有关高分辨率图像的信息,因此这种方法得到的效果比较好<sup>[6]</sup>.但这一类方法计算量很大,而且需要进行多帧之间的全局运动估计.如果估计不准确,其效果会急剧下降,从而无法实用.

1995年以后,人工神经网络被应用于图像超分辨率<sup>[7,9]</sup>,其优点在于利用神经网络的学习能力,抽取并逼近低分辨率图像与高分辨率图像间的内在联系,进而比较准确地恢复出低分辨率图像中已经丢失的原高分辨率图像的信息. Ahmed等<sup>[7]</sup>采用一个径向基函数(RBF)网络,为每一个输入模板建立一个隐单元;Plaziac<sup>[8]</sup>采用一个固定大小的反向传递(BP)网络;Candocia等<sup>[9]</sup>首先采用 Kohonen 自组织网络对图像子块进行聚类,然后对每一类图像子块用一个线性或非线性神经网络进行插值.要建立一个可以泛化的神经网络,或者需要数量巨大的神经元<sup>[7]</sup>,或者需要对图像进行分类<sup>[9]</sup>,计算量都非常大.在神经网络规模较小的情况下,即使图像训练集很大,其泛化效果亦难得到保证<sup>[8]</sup>.

本文提出了一种将现有的帧内插值算法与残差线性恢复相结合的图像超分辨率算法框架.在数学上证明了本方法的性能高于现有的算法,并且残差的线性恢复比直接对图像进行线性恢复具有更广泛的适应能力.在实际应用中,我们利用

一个神经网络逼近和模拟残差的线性恢复关系,实验结果证明其具有较强的泛化能力,并具有训练时间短,计算量小的特点,易于实时处理.

## 2 原理

假设需要进行放大的低分辨率图像是由高分辨率的图像经过低通滤波加降采样以后得到的.那么图像的超分辨率问题就可以归结为在已知低分辨率图像的情况下,如何估计降采样前的高分辨率图像的问题.如果低分辨率图像和高分辨率图像之间存在普遍的内在联系的话,那么我们就可以用神经网络通过学习的方法得到这种联系,来恢复原始的高分辨率图像.现有的神经网络方法就是建立在此基础上的<sup>[8,9]</sup>,并给出了实验结果,但没有给出这种方法的依据.本文给出了神经网络方法的理论依据,指出了它们的局限性,然后提出了我们的方法,并证明了其优势.

### 2.1 低分辨率图像与高分辨率图像的关系

由于传感器分辨率低,加上聚焦不准等原因造成的图像清晰度下降,会使图像的分辨率降低.不失一般性,我们把低分辨率图像看作是高分辨率的图像经过低通滤波加上降采样而得到的.将  $LN \times LN$  的高分辨率图像所有的像素经过行堆叠后排列成一列,然后将  $C$  幅高分辨率图像的列向量横向排列成为矩阵  $f$ ,  $C$  幅  $N \times N$  的低分辨率图像经过行堆叠后排列得到矩阵  $g$ .可以把二者的关系看作是一种映射关系.

**定义 1(正向映射 H)** 正向映射  $H: f \rightarrow g$  为高分辨率图像经过低通滤波和降采样获得低分辨率图像的过程.  $H$  表示滤波算子和抽取算子共同作用的结果.

**定义 2(逆向映射 G)** 逆向映射  $G: g \rightarrow f$  为低分辨率图像通过某种方法恢复高分辨率图像的过程.

**定理 1** 如果  $g$  是  $f$  通过某一种低通滤波和  $L$  点降采样获得的低分辨率图像集,在已知  $f$  和  $g$  的情况下,当  $\text{rank}(f) = KM$ ,  $K = L \times L$ ,  $M = N \times N$ ,则可从  $f$  和  $g$  得到正向映射  $H$  在最小均方误差意义上的线性估计  $\hat{H}$ .

**证明** 不失一般性,把  $f$  和  $g$  之间的正向映射关系写为线性关系

$$g = Hf + n, f \in R_+^{KM \times C}, n \in R_+^{M \times C}, H \in R^{M \times KM} \quad (1)$$

其中  $n$  表示噪声.在已知  $f$  和  $g$  的条件下,求解  $H$ ,使下列目标矩阵  $W(H)$  达到最小

$$W(H) = (g - Hf)(g - Hf)^T \quad (2)$$

对式(2)求导,令其导数为 0.当  $\text{rank}(f) = KM$  时,可以求得  $H$  的最小均方估计为<sup>[10]</sup>

$$\hat{H} = gf^T(ff^T)^{-1} \quad (3)$$

证毕.

由于低通滤波和  $L$  点降采样是客观存在的物理过程,对任意图像都可操作,  $H$  是低通滤波和  $L$  点降采样共同作用的线性估计,因此当此物理过程近似为线性过程时,  $H$  与  $f$  和  $g$  的类别无关.

**定理 2** 如果  $g$  是  $f$  通过某一种低通滤波和  $L$  点降采样获得的低分辨率图像集,在已知  $f$  和  $g$  的情况下,当  $\text{rank}(g) = M$ ,则可从  $f$  和  $g$  得到逆向映射  $G$  在最小均方误差意义上

的线性估计  $\hat{G}$ .

定理 2 的证明过程与定理 1 的证明类似,把  $f$  和  $g$  的逆向映射关系写为线性关系

$$f = Gg + v, f \in R_+^{KM \times C}, g \in R_+^{M \times C}, G \in R^{KM \times M} \quad (4)$$

其中  $v$  表示噪声.使下列将目标矩阵  $W(G)$  达到最小

$$W(G) = (f - Gg)(f - Gg)^T \quad (5)$$

当  $\text{rank}(g) = M$  时,可以求得  $(f, g)$  上  $G$  的最小均方估计

$$\hat{G} = fg^T(gg^T)^{-1} \quad (6)$$

为了避免计算量庞大的矩阵求逆,  $G$  的线性估计  $\hat{G}$  可以用线性神经网络对数据集  $(f, g)$  进行学习来得到(或用多层非线性神经网络得到非线性的  $\hat{G}$ ).文献[8]和[9]就是在已知  $f$  和  $g$  的条件下,用神经网络学习得到  $\hat{G}$ ,然后对低分辨率图像通过  $\hat{G}$  得到高分辨率图像的估计.但是,由于逆向映射是一对多的映射,因此  $\hat{G}$  和  $H$  不同,不具有与  $f$  和  $g$  无关的性质.从式(6)看,  $\hat{G}$  是依赖  $f$  的,由于不同的  $f$  经过相同的低通滤波加降采样后可能得到相同的  $g$ ,因此会得到不同的  $\hat{G}$ .显然,不同类型的  $f$  和  $g$  的逆向映射是有很大不同的,它与  $f$  有关.因而文献[8]得不到好的泛化能力.文献[9]对  $f$  和  $g$  先分类,然后对不同类型的图像求出不同的  $\hat{G}$ ,使计算的复杂度大大增加.

### 2.2 恢复残差获得高分辨率图像估计的算法

用  $Z$  表示某一种已知的图像插值算法,通过  $Z$  得到的  $C$  幅高分辨率图像的估计为

$$\tilde{f} = Z(g), \tilde{f} \in R_+^{KM \times C}, g \in R_+^{M \times C} \quad (7)$$

$\tilde{f}$  表示用插值算法  $Z$  得到的对高分辨率图像  $f$  的估计.  $\tilde{f}$  与真实高分辨率图像  $f$  的残差为

$$e_f = f - \tilde{f} = f - Z(g) \quad (8)$$

我们将式(8)称为高分辨率残差图像集.用降采样方法  $H$ ,对高分辨率残差图像集  $e_f$  进行降采样得到对应的低分辨率残差的图像集为

$$e_g = He_f = H[f - Z(g)] = Hf - HZ(g) \quad g - HZ(g) \quad (9)$$

将  $e_g$  向  $e_f$  的逆向映射关系近似为线性映射,用  $G$  表示,可写为

$$e_f = G e_g + v_e, e_f \in R^{KM \times C}, e_g \in R^{M \times C}, G \in R^{KM \times M} \quad (10)$$

同定理 2,当  $\text{rank}(e_g) = M$  时,可以得到  $G$  的最小均方估计

$$\hat{G} = e_f e_g^T (e_g e_g^T)^{-1} \quad (11)$$

从  $\hat{G}$  得到高分辨率残差图像的估计为

$$\hat{e}_f = \hat{G} e_g = e_f e_g^T (e_g e_g^T)^{-1} e_g \quad (12)$$

根据式(8),用函数  $S$  表示我们的算法对高分辨图像的估计为

$$S(g) = \hat{f} = \hat{e}_f + \tilde{f} \quad (13)$$

将式(7)、(12)代入式(13),并结合式(9),可以得到高分辨率图像的估计公式(14)

$$S(g) = Z(g) + \hat{G} e_g = Z(g) + \hat{G} [f - HZ(g)] \quad (14)$$

根据式(14),基于残差线性恢复的图像超分辨率算法框架如图 1 所示.

### 2.3 恢复残差获得高分辨率图像估计算法的优势

**定义 3(误差矩阵)** 对于任意一个从低分辨率图像估计高分辨率图像的方法  $X$ ,它从  $g$  得到对  $f$  的估计为  $\hat{f} =$

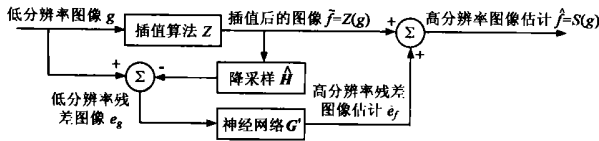


图 1 用神经网络恢复残差的图像超分辨率算法框架

$X(g)$ , 那么定义该算法在图像集  $f$  和  $g$  上的误差矩阵为

$$P(X, f, g) = ee^T = (f - \tilde{f})(f - \tilde{f})^T \quad (15)$$

**定理 3** 对于高分辨率图像集  $f$  和低分辨率图像集  $g$ , 任意图像插值算法  $Z$ , 而  $e_r, e_g$  和  $S$  是由式 (8)、(9) 和式 (14) 定义, 当  $\text{rank}(e_r) = KM$  且  $\text{rank}(e_g) = M$  时, 有

$$P(S, f, g) = P(Z, f, g)$$

**证明** 根据定义 3 和式 (8), 有

$$P(Z, f, g) = [f - Z(g)] [f - Z(g)]^T = e_r e_r^T \quad (16)$$

根据定义 3 和式 (8)、(13), 有

$$P(S, f, g) = [f - \tilde{f} - e_r] [f - \tilde{f} - e_r]^T = [e_r - e_r] [e_r - e_r]^T \quad (17)$$

将式 (12) 代入式 (17), 合并同类项后, 得到

$$P(S, f, g) = [e_r - e_r e_g^T (e_g e_g^T)^{-1} e_g] [e_r - e_r e_g^T (e_g e_g^T)^{-1} e_g]^T = e_r e_r^T - e_r e_g^T (e_g e_g^T)^{-1} e_g e_r^T \quad (18)$$

令  $A = e_r e_r^T, B = e_r e_g^T (e_g e_g^T)^{-1} e_g e_r^T, C = [e_r - e_r] [e_r - e_r]^T$ . 代入式 (16)、(18), 有

$$P(Z, f, g) = A \quad (19)$$

$$P(S, f, g) = A - B = C$$

当  $\text{rank}(e_r) = KM$  且  $\text{rank}(e_g) = M$  时, 易得  $A, B$  是正定的,  $C$  是半正定的, 因此  $C \leq A$ . 证毕.

定理 3 说明, 不管用什么插值算法作为我们方法中的  $Z$ , 我们的方法对图像集  $f$  进行估计的均方误差总不大于原算法  $Z$ . 也就是说, 对于任意一种插值算法, 只要在我们方法中把它当作  $Z$ , 那么我们的方法总在均方误差意义上比  $Z$  性能更优.

由于  $Z$  是某一种已知的插值算法, 一般来说, 经过它插值得到的  $\tilde{f}$  与真实高分辨率图像  $f$  的差别在于丢失了高频信息. 在  $f$  中只含有的低频信息的部分 (平滑区域), 残差图像  $e_r$  的值都接近于 0; 在  $f$  中含有的高频信息的部分 (边缘),  $e_r$  才有显著的值. 因此,  $e_r$  的种类比  $f$  少得多. 在 2.1 节我们已经提到, 不同种类的  $f$  会得到不同的  $G$ , 由于  $e_r$  的种类比  $f$  少得多, 因此通过式 (11) 估计出的  $G$  的种类就比  $G$  少得多, 那么每一个  $G$  就比  $G$  适应更多的图像. 只用一个神经网络来模拟  $e_g$  到  $e_r$  的逆向映射, 将比用一个神经网络来模拟  $g$  到  $f$  的逆向映射具有更好的泛化能力.

### 3 算法实现

在实际应用中, 我们采用自适应线性神经网络 (ALNN) 来学习逆向映射矩阵  $G$ . 根据式 (10), 神经网络的训练过程如图 2 所示. 对照图 1, 增加了学习的教师  $e_r$  及 LMS 学习算法.

下面给出图 1、2 中各方框的具体实现.

#### 3.1 插值算法 $Z$

从 2.2 节的讨论中可以看到, 本算法对已知插值算法  $Z$

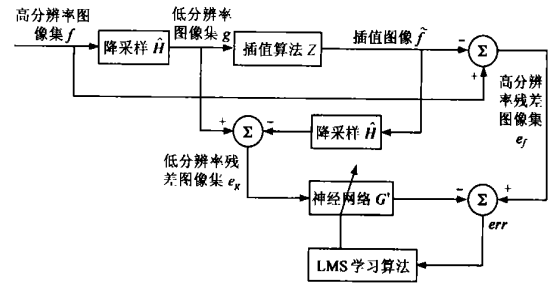


图 2 神经网络的训练过程

没有任何限制, 因此可以选用任何已有的图像插值算法. 在本例中选用双线性插值 (bilinear interpolation) 算法. 之所以选择双线性插值算法, 是因为它是一种常用的多项式插值算法, 并经常被用作其他图像插值算法的比较标准.

#### 3.2 降采样方法 $H$

为了使低分辨率图像在一定程度上保持清晰并避免迭混, 先将高分辨率图像沿列方向用下面式 (20) 所示的低通滤波器滤波<sup>[11]</sup>:

$$p_{lv}(x, y) = \frac{64p(x, y) + 38[p(x, y-1) + p(x, y+1)]}{128} - \frac{6[p(x, y-3) + p(x, y+3)]}{128} \quad (20)$$

式中  $p$  表示原图像像素的灰度值,  $p_{lv}$  表示经过列方向低通滤波后的像素的灰度值. 所得的结果再沿行方向用下面式 (21) 所示的低通滤波器滤波:

$$p_l(x, y) = \frac{64p_v(x, y) + 38[p_v(x-1, y) + p_v(x+1, y)]}{128} - \frac{6[p_v(x-3, y) + p_v(x+3, y)]}{128} \quad (21)$$

然后抽取每偶数行偶数列的像素, 得到只有原图像  $1/4$  大小的低分辨率图像.

#### 3.3 神经网络 $G$ 的结构

自适应线性神经网络的输入输出关系为

$$y = Wx + b, \quad x \in R^i, y \in R^j, W \in R^{j \times i} \quad (22)$$

$i$  为输入  $x$  的维数,  $j$  为输出  $y$  的维数,  $W$  为互联权值矩阵,  $b$  为阈值矢量. 比较式 (22) 与式 (10), 可以发现以下对应关系: 低分辨率残差图像像素  $e_g$  对应神经网络的输入  $x$ , 高分辨率残差图像像素  $e_r$  对应输出  $y$ , 逆向映射矩阵  $G$  对应权值矩阵  $W$ , 噪声  $v_e$  对应阈值矢量  $b$ .

可以看到, 定理 3 要求  $\text{rank}(e_r) = KM$ , 那么它的必要条件是  $C \leq KM = L^2 N^2$ . 这就意味着需要极大量的低分辨率和高分辨率图像对才能估计出逆向映射矩阵  $G$ . 如此大量图像的需求是难以满足的. 我们知道, 如果降采样过程  $H$  是局部的空间平移不变过程, 对整幅图像进行低通滤波和降采样等效于将图像分成相互交叠的 (overlapping) 图像子块分别进行模糊和降采样, 然后再进行拼合, 其条件是图像子块的大小必须不小于低通滤波的支撑域. 我们同样采用分块处理的方法来完成图像超分辨率, 这样既可以减小逆向映射矩阵  $G$  的规模, 又可以减少可能的残差图像的种类, 而且使得定理 3 的条件易于得到满足. 本论文中图像子块的大小为  $5 \times 5$ , 因此定理 3 的必要条件是  $C \leq 25L^2$ , 任意一对实际的低分辨率和高

分辨率图像分块组成的  $f$  和  $g$  都能满足这个条件. 为进一步减小神经网络的规模, 我们将神经网络的输出设计成该子块中近中心的一小部分. 我们设计输入的低分辨率残差图像子块的大小为  $5 \times 5$ , 输出为对应高分辨率残差图像子块近中心的  $3 \times 3$  个像素, 对应的神经网络有 25 个输入、9 个输出, 在神经网络的训练阶段, 我们从低分辨率残差图像中逐像素取出有交叠的图像子块组成  $e_g$ , 从高分辨率残差图像中逐像素取出有交叠的图像子块组成  $e_r$ , 对神经网络进行训练. 在神经网络的使用阶段, 用上述同样的方法取出低分辨率残差图像的图像子块作为神经网络的输入, 将输出拼合成高分辨率残差图像, 位置重叠的输出作平均以减少噪声的影响.

我们选择的学习算法是 Widrow-Hoff 学习算法, 它的基本优化步骤为

$$\begin{aligned} W_{k+1} &= W_k + (y - W_k x) x \\ b_{k+1} &= b_k + (y - W_k x) \end{aligned} \quad (23)$$

式中  $k$  表示训练的轮数,  $W$  表示权值矩阵,  $b$  表示阈值矢量,  $x$  表示输入矢量,  $y$  表示期望的输出矢量 (教师),  $\eta$  表示学习步长. 由于 Widrow-Hoff 学习算法的目标函数是均方误差, 因此又被称为最小均方 (LMS) 学习算法, 它与我们求  $G$  的最小二乘估计的要求是一致的.

#### 4 实验测试结果

我们分别以  $256 \times 256$  的 Lena 图像和 Pep 图像作为原始的高分辨率图像对神经网络进行训练. 以  $128 \times 128$  的 Ape、Barbara、Bird、Camera、Goldhill、Harbor、Lena、Pep 图像作为测试集, 测试结果如表 1 所示. 与其相比较的, 是传统的双线性插值算法和 bicubic 算法, 以及文献[8]中介绍的用 BP 网络的方法.

表 1 与传统及文献[8]方法结果的比较(从  $128 \times 128$  到  $256 \times 256$ )

图像名称	以下方法所得结果的 PSNR (dB)				
	Bilinear	Bicubic	[8]的方法*	我们的方法**	我们的方法***
Ape	24.81	25.35	25.20	25.97	26.02
Barbara	29.12	30.09	29.32	31.09	31.04
Bird	37.54	39.68	38.99	41.77	41.68
Camera	30.62	32.24	31.96	34.25	34.30
Goldhill	29.83	30.75	30.44	31.77	31.71
Harbor	26.41	27.09	26.72	27.97	27.98
Lena	30.29	31.59	31.32	33.31	33.06
Pep	29.69	30.88	30.83	32.06	32.20

\*用 Lena、Pep、Ape、Harbor 四幅图像作为训练集; \*\*用 Lena 图像作为训练集; \*\*\*用 Pep 图像作为训练集

表 2 是我们的方法与文献[9]中介绍的通过分类的算法的比较. 由于文献[9]未对算法进行泛化测试, 因此只得到一项数据.

我们用训练好的神经网络, 将  $256 \times 256$  的测试图像放大到  $512 \times 512$ , 其结果见表 3.

表 2 与分类方法[9]的比较(从  $128 \times 128$  到  $256 \times 256$ )

图像名称	以下方法所得结果的 PSNR (dB)			
	文[9]的方法*	文[9]的方法**	我们的方法***	我们的方法****
Lena	32.63	31.78	33.31	33.06

\*用 Lena 图像作为训练集, 30 类; \*\*用 Pep 图像作为训练集, 30 类; \*\*\*用 Lena 图像作为训练集; \*\*\*\*用 Pep 图像作为训练集

表 3 与传统方法结果的比较(从  $256 \times 256$  到  $512 \times 512$ )

图像名称	以下方法所得结果的 PSNR (dB)			
	Bilinear	Bicubic	我们的方法*	我们的方法**
Ape	27.53	29.47	33.41	33.44
Barbara	24.88	25.19	25.47	25.49
Bird	41.85	43.52	44.97	44.89
Camera	33.95	35.47	36.88	36.82
Goldhill	34.48	36.89	41.26	41.03
Harbor	23.71	24.10	24.54	24.55
Lena	32.66	33.96	35.57	35.40
Pep	33.89	36.05	40.05	40.03

\*用 Lena 图像作为训练集; \*\*用 Pep 图像作为训练集

从表 1、2、3 的数据, 我们可以看出: 我们算法的性能比传统的算法有较大的提高, 并高于文献[8]中用神经网络直接对图像进行插值的方法. 与文献[9]中经过分类的方法相比也毫不逊色. 在本例中我们以双线性插值算法作为  $Z$ , 根据定理 3 可以推断, 如果采用 bicubic 算法或其他性能更好的算法作为  $Z$ , 能够使本算法的性能得到进一步的提高. 实验数据表明训练集图像的选择对本方法的性能的影响是微小的, 符合 2.3 节中的讨论. 而且只用一幅图像进行训练, 神经网络就具有较强的泛化能力.

图像实例见图 3. 图 3(a) 是低分辨率的文本图像, 图 3(b) 是经过双线性插值得到的图像, 图 3(c) 是我们的方法 (用 Lena 图像训练) 获得的图像. 可以看到, 虽然所用的神经网络是用与文本图像完全不同的图像集进行训练的, 但仍然比原来用双线性插值得到的模糊文本变得清晰, 可分辨性增强了.

#### 5 计算复杂度分析

由式 (22) 可以得到, 神经网络得到每一组输出所需要的计算量为  $25 \times 9 = 225$  次乘法和加法. 相比之下, 文献[8]中采用了有 24 个输入、16 个隐单元、5 个输出的 BP 网络, 计算一组输出所需要的计算量为  $24 \times 16 + 16 \times 5 = 464$  次乘法和加法, 以及 16 次非线性函数的运算. 文献[9]中在分类结束后采用线性联想记忆的方法进行插值的计算量与我们的方法相当, 但它在插值之前先要进行分类, 增加了额外的计算量, 同时也大大增加了训练阶段的时间.

#### 6 结论

本文提出了一种用线性神经网络恢复图像残差的超分辨

率算法框架,并在数学上证明了这种方法比已有图像插值算法具有更好的性能.此方法将已有的图像插值算法与神经网络相结合,利用神经网络恢复插值图像与实际高分辨率图像之间的残差,从而得到更真实更清晰的高分辨率图像估计.实验结果证明了本方法训练得到的神经网络具有较好的泛化能力,可以用于各种图像的精确放大.而且由于采用了单层线性

神经网络,使整个方法的计算量非常有限,易于实时处理.

应该注意到,不仅仅图像插值算法,任何一种单帧的图像超分辨率算法都可以当作我们的算法框架中的 Z,然后通过恢复残差来得到改进后的超分辨率算法.然而当 Z 的估计越来越精确时,式(11)估计存在所需要的  $e_g$  满秩的条件就越来越难满足,需要的训练图像集就越来越大.



图 3 (a)原始低分辨率图像;(b)双线性插值得到的图像;(c)用我们的方法得到的图像

参考文献:

[ 1 ] K Jensen ,D Anastassiou. Subpixel Edge Localization and the Interpolation of Still Images[J]. IEEE Trans Image Processing ,1995 ,4(3) :285 - 295.

[ 2 ] A M Darwish ,M S Bedair. An Adaptive Resampling Algorithm for Image Zooming[A]. Proc. SPIE[C]. San Jose ,CA ,USA ,1996 ,2666 :131 - 144.

[ 3 ] S A Martucci. Image Resizing in the Discrete Cosine Transform Domain [A]. Proc. Int. Conf. Image Processing [ C]. Washington ,DC ,USA , 1995 ,2 :244 - 247.

[ 4 ] W K Carey ,D B Chuang ,S S Hemami. Regularity-preserving image interpolation [J]. IEEE Trans Image Processing ,1999 ,8 (9) : 1293 - 1297.

[ 5 ] R R Schultz ,R L Stevenson. A bayesian approach to image expansion for improved definition[J]. IEEE Trans Image Processing ,1994 ,3 :233 - 242.

[ 6 ] 邵凌 ,丁佩律 ,张立明 ,胡波. 从多帧低分辨率图像序列中获取高分辨率图像的算法研究[J]. 电子学报 ,2002 ,30(1) :58 - 61.

[ 7 ] F Ahmed ,S C Gustafson ,M A Karim. High-fidelity image interpolation using radial basis function neural networks [A]. Proc. IEEE National Aerospace and Electronics Conf [C]. Dayton ,OH ,USA ,1995 ,2 :588 - 592.

[ 8 ] N Plaziac. Image interpolation using neural networks[J]. IEEE Trans Image Processing ,1999 ,8(11) :1647 - 1651.

[ 9 ] F M Candoncia J C Principe. Superresolution of Images based on Local

Correlations[J]. IEEE Trans. Neural Networks ,1999 ,10 (2) :372 - 380.

[ 10 ] T Kailath ,A H Sayed ,B Hassibi. Linear Estimation[M]. Upper Saddle River ,NY :Prentice Hall ,2000.

[ 11 ] ISO/ IEC/JTC1/ WG11 MPEG3/ 143[S]. 1993 ,Rome ,Italy.

作者简介:



潘逢治 男,1978 年生于江苏无锡,2000 年于复旦大学电子工程系获学士学位,2003 年获复旦大学电子工程系硕士学位,现在 Sony 公司工作,目前研究方向主要为图像处理.



张立明 女,1943 年 2 月生于浙江临海,1965 年毕业于复旦大学物理系,现任复旦大学电子工程系教授、博士生导师,发表著作 4 本,论文 90 多篇,目前研究方向主要包括人工神经网络、图像处理和识别.