

文章编号: 1001-3806(2006)01-0107-03

基于 NSA 多尺度模型的人脸识别

赵明华¹, 游志胜^{1*}, 赵永刚², 吕学斌¹

(1. 四川大学 计算机学院 图形图像研究所, 成都 610064 2. 西南石油学院, 成都 610500)

摘要: 提出了 NSA 多尺度模型。该模型摒弃了 LSA 模型中不同尺度的图像间具有线性映射关系的假设。首先利用神经网络建立不同尺度图像间的映射关系; 其次使用反向传播算法训练神经网络确定这种映射关系; 最后根据该映射关系由低分辨率图像估计高分辨率图像。利用对比度相似性量化估计图像与目标图像间的相似程度。将该模型用于人脸识别, 提出利用梯度算子进行图像分割提高识别的准确性。实验结果表明, 以该模型分析得到的对比度相似性为 95.3634%; 以该模型为基础的人脸识别系统对光照具有很好的鲁棒性。

关键词: 多尺度分析; 线性尺度自回归 (LSA); 神经网络尺度自回归 (NSA); 小波变换

中图分类号: TP391.41 **文献标识码:** A

Face recognition based on NSA multiscale model

ZHAO Ming-hua¹, YOU Zhi-sheng¹, ZHAO Yong-gang², LYU Xue-bin¹

(1. Institute of Image & Graphic Department of Computer Science, Sichuan University, Chengdu 610064 China 2. Southwest Petroleum Institute, Chengdu 610500 China)

Abstract A multiscale model named neural networks scale autoregressive (NSA) is presented. The model uses neural networks to represent the map between the pixels residing at images of various resolutions, instead of the assumption of linearity in linear scale autoregressive (LSA) model. Then back-propagation algorithm is used in the neural networks training to decide the map. Finally, the identified map is used to estimate images at finer resolution from coarser versions. Similarity in contrast is used to gauge the degree of similarity between estimated images and target images. The model is applied in face recognition and image segmentation. Approach using gradient operator is proposed to increase the recognition accuracy. Experimental results show that similarity between image estimated by NSA model and target image is 95.3634%. The NSA-based face recognition system is robust to illumination.

Key words multiscale analysis; linear scale autoregressive; neural networks scale autoregressive; wavelet transform

引 言

计算机人脸识别是由计算机分析人脸图像, 从中提取出有效识别信息, 用以辨认身份的一种技术。由于人脸图像这种多维信号的处理, 计算效率十分重要。目前, 多尺度分析这种高效的图像处理方法已被用于人脸识别。本文中首先讨论多尺度分析的特点, 然后指出 LSA 多尺度模型中假设不同尺度的图像间具有线性映射关系的不足之处, 在此基础上构建了基于神经网络的 NSA 多尺度模型, 最后将 NSA 多尺度模型用于人脸识别实验, 通过比较估计图像和目标图像的相似程度将其分类。

1 多尺度分析和 LSA 模型

一般来说, 图像的信息依其结构的不同, 在人们感

知过程中的重要性也会不同, 即图像中不同结构的信息, 客观上就应该出现在感知的不同层次上。人的感知过程是一个由粗到精的分层处理过程, 即首先通过大范围特征来进行粗分析, 然后在这一基础上, 通过捕捉细节, 使分析得以越来越精细, 最终获得对感知对象的准确理解^[1]。单尺度分析难以很好地描述对象多变的结构和细节, 而且不符合人的感知机制。而多尺度分析正是与这一机制相匹配的多层次描述和多层次处理方法。

多尺度分析是用不同的分辨率或者不同的尺度逐级逼近函数 $f(x)$ ($f(x) \in L^2$), 然后选择合适的分辨率或者同时在各级分辨率上处理信号。本文中主要研究小波多尺度分析。每一次小波变换将原始图像分解成一个低频图像和 3 个高频图像^[2,3]。该过程可以用树结构表示, 树的每一层代表一个尺度。树上结点 s 的状态向量可以表示为:

$$X(s) = f(X(\bar{y}_s), \text{app}(\bar{y}_s), \text{det}(\bar{y}_s)) \quad (1)$$

式中, \bar{y}_s 表示 s 的父结点, \bar{y} 表示向上变化的算子, $X(s)$ 表示在尺度 s 上的状态向量, $\text{app}(\bar{y}_s)$ 和 $\text{det}(\bar{y}_s)$

作者简介: 赵明华 (1979-), 女, 博士研究生, 主要从事数字图像处理 and 模式识别等研究。

* 通讯联系人。E-mail: zsyout@mail.sc.cninfo.net

收稿日期: 2004-08-02 收到修改稿日期: 2004-10-12

分别表示尺度 \bar{y}_s 上的近似信息和细节信息。 f 表示处于不同尺度的图像间的映射函数。随着函数 f 的结构和状态向量 $X(s)$ 的不同, 可产生不同的多尺度模型。

LSA 多尺度模型以白小波系数假设和处于不同尺度的图像间具有线性映射关系的假设为前提^[4,5]。然而, 尺度之间线性关系的假设使得不能从模型的高阶导数中提取有用信息(高阶导数为 0); 在处理图像等复杂信号时, 以这种假设为基础模型不能并行对所有像素进行运算, 运算效率不高。

2 NSA 模型

NSA 多尺度模型的基本思想是: 用前馈神经网络来确定不同尺度的图像之间的映射关系而保持 LSA 模型中小波系数的假设。构造具有 1 个隐含层的 NSA 模型, 包含 H ($H = 40$) 个隐单元。其内部结构见图 1。

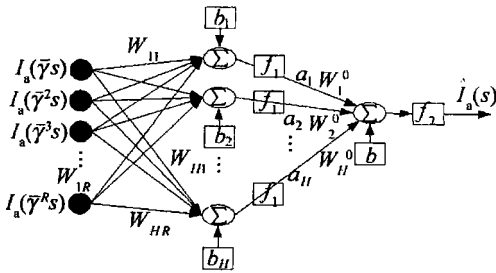


Fig 1 Internal structure of the NSA model

图 1 中, $I_a(\bar{y}^s), I_a(\bar{y}^{2s}), \dots, I_a(\bar{y}^R s)$ 表示较低分辨率图像的像素信息; $I_a(\hat{s})$ 表示高分辨率估计图像的像素信息; $W_{11}, W_{12}, \dots, W_{1R}, \dots, W_{H1}, W_{H2}, \dots, W_{HR}$ 表示从输入层到隐含层的权值; $W_1^0, W_2^0, \dots, W_H^0$ 表示从隐含层到输出层的权值; b_1, b_2, \dots, b_H, b 是相应的偏置; f_1 和 f_2 是激活函数, 设 $f_1(x) = \frac{e^{-x} - e^x}{e^{-x} + e^x}$, $f_2(x) = x$ 。NSA 多尺度模型的训练步骤是: (1)对原始图像进行小波变换; (2)使用反向传播训练算法, 根据上述小波变换结果和原始图像设置神经网络中合适的权值, 从而确定不同尺度上图像像素间的映射关系; (3)重复步骤 2 找到图像中所有像素的这种映射关系; (4)根据上述映射关系, 从低分辨率的图像中估计得到高分辨率的图像。

第 k 个隐单元的输出可以表示为:

$$a_k = f_1 \left[\sum_{j=1}^R (W_{kj} I_a(\bar{y}^j s) + b_k) \right] \quad (2)$$

高分辨率的估计像素值通过以下公式计算:

$$\hat{I}_a(s) = f_2 \left[\sum_{k=1}^H (W_k^0 a_k) + b \right] \quad (3)$$

由 (2) 式、(3) 式和激活函数, 可以得到:

$$\hat{I}_a(s) = \sum_{k=1}^H \left\{ W_k^0 \tanh \left[\sum_{j=1}^R (W_{kj} I_a(\bar{y}^j s) + b_k) \right] \right\} + b \quad (4)$$

训练完成后, 估计高分辨率图像就是并行对所有像素进行运算, 计算效率提高。

3 实验与分析

3.1 估计效果实验

利用 ORL 人脸数据库对所构造模型的估计效果进行测试。实验结果如图 2 所示, 图 2a 是从该库选取的一个原始图像, 图 2b 和图 2c 分别是图 2a 经过一级和两级小波分解得到的图像, 图 2d 和图 2e 分别表示利用 LSA 模型和 NSA 模型得到的估计图像。



Fig 2 Images estimated by LSA model and NSA model respectively

由图 2 可见, 利用 LSA 模型得到的图像在口鼻部分不清晰; 利用 NSA 模型得到的图像从清晰性上更加接近原始图像。为了量化模型的估计效果, 使用以下标准来测量估计图像与目标图像之间的相似程度:

$$\Omega(M, \hat{M}) = 100 \times \left\{ 1 - \frac{1}{N_1 N_2} \times \sum_{k=1}^{N_1} \sum_{l=1}^{N_2} [(I_a^{\text{diff}}(k, l) - \bar{I}_a^{\text{diff}})^2] \right\} \quad (5)$$

式中, M 和 \hat{M} 表示目标图像和估计图像, Ω 表示两幅图像间的对比度相似程度, $I_a^{\text{diff}}(k, l)$ 表示这两幅图像在 (k, l) 处像素值的差别, \bar{I}_a^{diff} 表示差别图像的平均灰度值, N_1 和 N_2 表示图像的维数。使用 NSA 模型得到的估计图像与目标图像之间的 Ω 值达到 95.3634%, 而使用 LSA 模型得到的 Ω 值 93.1123%。NSA 模型得到的估计图像都更加接近目标图像。然而, 由于缺乏细节图像信息, 使用 NSA 模型估计某些像素时出现不确定现象, 表现为图 2e 中分散的白斑, 实验测得该白斑的比率为 1.7%。

3.2 人脸识别实验

分别选取 ORL 人脸数据库的前 20 个人的第 1 个人脸图像组成训练集, 将训练集中的每个图像在以 2% 为步长、从 -10% ~ +10% 的光照变化下得到的所有图像组成测试集。这样, 训练集共有 20 幅图像组成, 测试集共有 220 幅图像组成。实验步骤为: (1)分别对每个训练样本和测试样本进行两级小波分解; (2)使用训练样本的小波分解结果和目标图像为神经网络设置合适权值, 从而得到处于不同尺度的图像像素间的映射关系; (3)根据测试样本的小波分解结果和上述映射关系估计高分辨率的图像; (4)使用 Ω 测量估计图像与目标图像的相似程度, 进而将其分类。

当光照变化为 $\alpha\%$ 时, 像素值的变化如下:

$$I_{\text{new}}(s) = (1 + \alpha\%) \times I_{\text{old}}(s) \quad (6)$$

图 3a和图 3b表示测试样本是光照为 - 10% 的第 1 个人脸图像时, 分别利用训练出的 20 个 LSA 模型和 20 个 NSA 模型得到的估计图像。由图可见, 使用 NSA 模型得到的估计图像之间的差别较为显著, 除了第 1 个图像外, 只有第 17 和第 19 个估计图像与目标图像比较接近。图 4 表示测试样本是光照为 - 10% 的第 1 个人脸图像时, 分别利用训练出的 20 个 NSA 模型得到的估计图像与目标图像之间的对比度相似性。

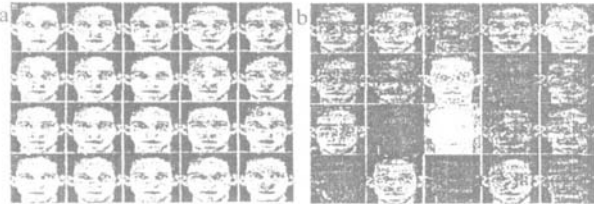


Fig 3 Estimation of the face number 1 using the 20 LSA and NSA models

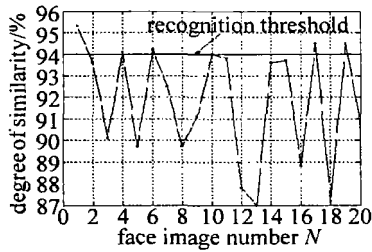


Fig 4 Degree of similarity between estimates and target image measured by Ω

表 1 是使用对比度相似性标准测量得到的全体测试样本在不同光照下的识别率。由表可见, 识别率基本不随光照变化而变化, 这说明以 NSA 模型为基础的人脸识别系统对光照具有很好的鲁棒性。以 NSA 模型为基础的人脸识别系统的识别率达到 95%, 高于 LSA 模型。

Table 1 Recognition results of NSA-based face recognition system under various illumination

illumination /%	- 10	- 8	- 6	- 4	- 2	0	2	4	6	8	10
criterion Ω /%	95	95	95	95	95	95	95	95	95	95	95

为了提高识别的准确率, 使用梯度算子图像分割方法对图像进行预先筛分。训练出的 NSA 模型对输入像素的偏微分如下:

$$\frac{\partial [\hat{I}_a(s)]}{\partial [I_a(\bar{y}^j s)]} = \sum_{k=1}^H \{W_k^0 \times W_{k1} \times [1 - (\tanh(\sum_{j=1}^R W_{kj} I_a(\bar{y}^j s) + b_k))^2]\} \quad (7)$$

第 K 阶梯度运算可以表示为:

$$I_a^{\nabla K}(s) = \|\nabla K(\hat{I}_a(s))\|_2 = \sqrt{\sum_{j=1}^R \left\{ \frac{\partial^K [I_a(s)]}{[\partial (I_a(\bar{y}^j s))]^K} - \bar{I}_j^{\nabla K} \right\}^2} \quad (8)$$

式中, $\bar{I}_j^{\nabla K} = \frac{1}{N_1 N_2} \times \sum_s \left\{ \frac{\partial^K [I_a(s)]}{[\partial (I_a(\bar{y}^j s))]^K} \right\}$ 。利用下式

将图像像素值 $I_a^{\nabla K}(s)$ 转化成布尔型值 $\beta_a^{\nabla K}(s)$:

$$\begin{cases} \beta_a^{\nabla K}(s) = 1 & I_a^{\nabla K}(s) \geq \bar{I}^{\nabla K} \\ \beta_a^{\nabla K}(s) = 0 & I_a^{\nabla K}(s) < \bar{I}^{\nabla K} \end{cases} \quad (9)$$

式中, $\bar{I}^{\nabla K} = \frac{1}{R} \times \sum_{j=1}^R \bar{I}_j^{\nabla K}$ 。通过这种方法, 图像转化成更适用于计算机处理的二进制图像。图 5 表示测试样本是第 1 个图像时, 分别使用 20 个 NSA 模型的权和基进行 5 阶梯度运算的结果。由图可见, 第 1 个图像的面部特征最清晰。利用这种方法可以使识别准确率提高到 96.3%。



Fig 5 Segmentation of subject number 1 by the 20 NSA models using 5th order gradient operator

4 结论

在分析 LSA 多尺度模型的基础上, 摒弃 LSA 模型中不同尺度的图像间具有线性映射关系的假设, 构建了基于神经网络的 NSA 多尺度模型。将该模型用于从低分辨率图像中估计高分辨率图像, 利用对比度相似性量化估计图像与目标图像的相似程度。经测试发现, 使用该模型分析得到的对比度相似性高于 LSA 模型的分析结果。将该模型用于人脸识别处理, 用处于不同光照下的 220 幅人脸图像组成测试集对识别效果进行实验, 结果表明, 以该模型为基础的人脸识别系统对光照具有很好的鲁棒性。为了提高系统识别的准确率, 提出使用梯度算子图像分割方法对图像进行预先筛分, 效果良好。

参 考 文 献

- [1] MARAGOS P. Pattern spectrum and multiscale shape representation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence 1989, 11(7): 701~ 716
- [2] MALLAT S. A wavelet tour of signal processing [M]. 2nd ed. Beijing: China Machine Press, 2002. 230~ 238 (in Chinese).
- [3] BURRUS C S, GOPNATH R A, GUO H T. Introduction to wavelets and wavelet transforms [M]. New Jersey: Prentice Hall, 1998. 159~ 164
- [4] BANHAM M R, KATSAGGELOS A K. Spatially adaptive wavelet-based multiscale image restoration [J]. IEEE Transactions on Image Processing 1996, 5(4): 619~ 634
- [5] DAUDI K, FRAKT A B, WILLSKY A S. Multiscale autoregressive models and wavelets [J]. IEEE Transactions on Information Theory 1999, 45(3): 828~ 845
- [6] BOUMAN C, LIU B. Multiple resolution segmentation of textured images [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence 1991, 13(2): 99~ 113.