

基于泛函网络的非线性回归预测模型及学习算法

何登旭¹, 李艳芳¹, 刘向虎², 周永权¹

HE Deng-xu¹, LI Yan-fang¹, LIU Xiang-hu², ZHOU Yong-quan¹

1. 广西民族大学 数学与计算机科学学院, 南宁 530006

2. 运城学院 应用数学系, 山西 运城 044000

1. College of Maths and Computer Science, Guangxi University for Nationalities, Nanning 530006, China

2. Department of Applied Mathematics, College of Yuncheng, Yuncheng, Shanxi 044000, China

E-mail: dengxu@126.com

HE Deng-xu, LI Yan-fang, LIU Xiang-hu, et al. Non-linear regression forecast model based on functional networks and learning algorithm. Computer Engineering and Applications, 2008, 44(24): 74-77.

Abstract: Fitting of forecast function is very difficult and important in non-linear regression forecast problems. The accuracy is directly affected by the fitting of forecast function. Non-linear model in the traditional method is difficult to solve the system whose non-linear is stronger, and the result of fitting and forecast is not ideal. Function network is a recently introduced extension of neural networks. It has certain advantages solving non-linear problems. Non-linear regression forecast model and learning algorithm based on functional networks are proposed in this article. Some examples about one-dimensional non-linear regression forecast and multi-dimensional non-linear regression forecast are provided. The simulation results demonstrate that forecast model based on functional networks whose accuracy of fitting and forecasting is more than some traditional methods has some value about theory and application.

Key words: functional networks; non-linear regression; forecast; learning algorithm

摘要: 在非线形回归预测中, 预测函数的拟合是其难点和关键, 直接影响预测精度。当系统非线性较强时, 传统方法不易于处理, 拟合和预测结果不理想。泛函网络是最近提出的一种对神经网络的有效推广, 在处理非线性问题时有一定的优势。为此提出了基于泛函网络的非线性回归预测模型和相应的学习算法。并分别就一元非线性回归预测和多元非线性回归预测给出了相应的实例。计算机仿真结果表明, 泛函网络预测模型拟合度和预测精度都明显高于某些传统的方法, 有较好的理论和应用价值。

关键词: 泛函网络; 非线性回归; 预测; 学习算法

DOI: 10.3778/j.issn.1002-8331.2008.24.021 **文章编号:** 1002-8331(2008)24-0074-04 **文献标识码:** A **中图分类号:** TP18

1 引言

回归分析是工程应用和科学实验等领域广泛应用的方法之一, 主要应用于预测、描述、控制等。在预测问题中, 常用多元回归模型反映预报量与各因素之间的依赖关系, 其中线性回归分析有着广泛的应用。由于自然规律的复杂性, 以现实世界为研究对象的目标系统, 绝大多数为非线性系统。解决非线性回归问题, 传统的方法一般是通过变量替换法转化为线性回归^[1], 利用线性回归理论来确定回归系数, 然后通过“反变换法”或“两阶段法”得到原非线性回归预测模型的参数估计值, 这样就可能不再具有最优性(残差平方和最小), 若用于实际预测有时会产生较大的预测误差。对此, 汤振民、陶菊春等人分别于1998、2003年在文献[2-3]中对上述方法进行了一定的改进, 但

是对非线性关系进行“线性化”处理的缺陷仍未克服: 一是变量替换依赖于预测者的观察和经验, 当模型涉及多个变量时, 要找出适当的变量替换是相当困难的; 二是近似计算要产生误差, 当误差较大时, 可能导致预测的失败。

本文希望利用泛函网络在解决非线性问题方面的优势, 建立一种基于泛函网络的非线性回归预测模型和相应的学习算法, 实例验证表明, 该模型和算法具有明显的优势。

2 泛函网络的结构

泛函网络^[4]是1998年由E-Castillo提出的, 是人工神经网络的一般化推广, 泛函网络处理的是一般泛函模型, 各个神经元之间的连接无权值, 并且神经元函数不固定, 而是一个给定

基金项目: 国家自然科学基金(the National Natural Science Foundation of China under Grant No.60461001); 国家民委自然科学基金(No.05GX06); 广西省自然科学基金(the Natural Science Foundation of Guangxi Province of China under Grant No.桂科自0728054); 广西民族大学研究生教育创新计划项目(No.gxun-chx0749)。

作者简介: 何登旭(1963-), 男, 教授, 从事数学建模方法、优化计算及计算机应用研究; 李艳芳(1981-), 女, 硕士研究生, 主要从事智能优化算法方面的研究。

收稿日期: 2007-10-25 **修回日期:** 2008-01-31

的基函数族的组合。人们可根据特定问题来选择不同的基函数族(如多项式、三角函数、Fourier 展开级数等)来满足不同系统问题的建模和逼近。在文献[5]中,作者通过具体实例把泛函网络与神经网络进行比较,从仿真结果可以看出,泛函网络的性能优于神经网络。泛函网络已经成功地被应用于很多方面,在文献[6]中作者用泛函网络来进行混沌时间序列的预测,提取被混沌信号屏蔽的信息,在文献[7]中作者用泛函网络来逼近微分、差分 and 泛函方程的解。泛函网络在上述的实验中都表现出了较好的性能。

在图1中画出了一个典型的泛函网络模型图,其主要组成部分如下:

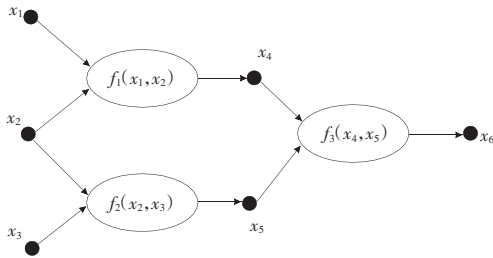


图1 一个典型的泛函网络模型图

(1) 一个输入单元层。其功能是输入信息,在图1中输入层包含 $\{x_1, x_2, x_3\}$ 。输入单元以带有相应名字的实心圆来表示。

(2) 若干中间存储单元层。它存储由神经元产生的信息,在图1中只有一个中间存储单元层,包含 $\{x_4, x_5\}$ 。也以带有相应名字的实心圆来表示。

(3) 一层或多层泛函神经元(或称处理单元)。每个泛函神经元是一个计算单元,它计算的是一组来自前一层神经元或输入单元的输入值,并给下一层神经元或输出单元提供数据。计算单元相互连接,每一个神经元的输出可作为另一个神经元或者输出单元输出数据的一部分,一旦给定输入值,输出便由神经元的类型来确定,它有一函数定义。在图1中有两层泛函神经元,分别为 f_1, f_2 和 f_3 。

(4) 一层输出层。这是最后一层单元,它输出网络的结果是数据,输出单元也以带有相应名字的实心圆来表示,图1中为 $\{x_6\}$ 。

(5) 一些直接连接。它们连接输入层、中间层神经元和输出层,在图1中用箭头表示,箭头的方向表示信息流的方向。

所有这些元素一起形成了泛函网络的结构,它确定了网络的泛函能力。

3 基于泛函网络的非线性回归预测模型

既然泛函网络作为神经网络的一种推广,与神经网络一样,泛函网络也有各种各样的结构,不可能用一个统一的通用的结构来描述所有的泛函网络,也不可能用一个统一的泛函方程来表示所有的泛函网络。根据这个特性,以结合实际问题的“先验知识”作为指导,可以设计出所求问题的泛函网络结构。

3.1 基于泛函网络的一元非线性回归预测

3.1.1 一元非线性回归预测模型

根据一元非线性回归的特点,可以设计出一个单输入、单输出和三层中间处理单元的泛函网络模型,如图2所示,其中, x 为自变量, y 为因变量。神经元函数 $f_i(x)$, $(i=1, 2, \dots, n)$,不是固定的,而是可以学习的,是一个给定的函数族的组合。与图2相对应的泛函方程为:

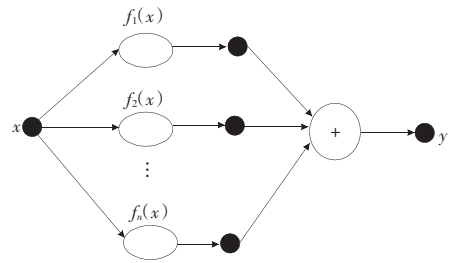


图2 一元非线性回归预测模型

$$y = \hat{F}(x) = \sum_{i=1}^n f_i(x) \quad (1)$$

在泛函网络中一个重要的问题是表达式的唯一性,文献[5]已经证明这种单输入、单输出网络模型的表达式的唯一性。在图2的模型中要得到回归预测模型等价于学习函数 $\hat{F}(x)$,而学习 $\hat{F}(x)$ 等价于学习函数 $f_i(x)$ 。这样,可以用一组已知基函数 $\{\phi_{ij}(x); j=1, 2, \dots, m, i=1, 2, \dots, n\}$ 的线性组合来表示 $f_i(x)$:

$$\hat{f}_i(x) = \sum_{j=1}^m a_{ij} \phi_{ij}(x) \quad (2)$$

令 $\boldsymbol{\phi}_i(x) = (\phi_{i1}(x), \phi_{i2}(x), \dots, \phi_{im}(x))^T$, $\mathbf{a}_i = (a_{i1}, a_{i2}, \dots, a_{im})^T$, 则式(3)可以写为:

$$\hat{f}_i(x) = \sum_{j=1}^m a_{ij} \phi_{ij}(x) = \mathbf{a}_i^T \boldsymbol{\phi}_i(x) \quad (3)$$

把式(4)代入式(2)就可以得到:

$$y = \hat{F}(x) = \sum_{i=1}^n \hat{f}_i(x) = \sum_{i=1}^n \sum_{j=1}^m a_{ij} \phi_{ij}(x) = \sum_{i=1}^n \mathbf{a}_i^T \boldsymbol{\phi}_i(x) \quad (4)$$

其中 $\{\phi_{ij}(x); j=1, 2, \dots, m, i=1, 2, \dots, n\}$ 是一些给定的适合逼近 f_i 到期望精度线性独立的函数集。 $\{a_{ij}; j=1, 2, \dots, m, i=1, 2, \dots, n\}$ 为泛函网络的参数。

3.1.2 一元非线性回归预测模型的学习算法

对于有教师的学习算法应当事先给定一组训练样本,比如:

$$\{(x^{(k)}, y^{(k)}) | k=1, 2, \dots, q\} \quad (5)$$

采用一种简单、收敛速度快的学习算法——基于最小二乘法思想的学习算法。

定义误差代价函数:

$$e_k = y^{(k)} - \hat{y}^{(k)} = f(x_k) - \hat{F}(x_k) = f(x_k) - \sum_{i=1}^n \sum_{j=1}^m a_{ij} \phi_{ij}(x_k) \quad (6)$$

为了找到最优的网络参数,需要最小化误差平方和:

$$Q = \sum_{k=1}^q e_k^2 = \sum_{k=1}^q \left[f(x_k) - \sum_{i=1}^n \sum_{j=1}^m a_{ij} \phi_{ij}(x_k) \right]^2 \quad (7)$$

为了保证网络表达式的唯一性,设置网络的初始值为:

$$\hat{f}_i(x_0) = \sum_{j=1}^m a_{ij} \phi_{ij}(x_0) = u_{0i} \quad i=1, 2, \dots, n \quad (8)$$

其中 u_{0i} 是给定的一些常数。为了找到最优化网络参数,应用Lagrange 乘数法,考虑下面的目标函数:

$$Q_\lambda = Q + \sum_{i=1}^n \lambda_i \left(\sum_{j=1}^m a_{ij} \phi_{ij}(x_0) - u_{0i} \right) = \sum_{k=1}^q \left[f(x_k) - \sum_{i=1}^n \sum_{j=1}^m a_{ij} \phi_{ij}(x_k) \right]^2 + \sum_{i=1}^n \lambda_i \left(\sum_{j=1}^m a_{ij} \phi_{ij}(x_0) - u_{0i} \right) \quad (9)$$

式(18)中分别对 a_{ij}, λ_i 求其偏导得:

$$\begin{cases} \frac{\partial Q_\lambda}{\partial a_{lr}} = -2 \sum_{k=1}^p e_k \phi_{lr}(x_k) + \lambda a_{lr} \phi_{lr}(x_{i0}) = 0 \\ r=1, 2, \dots, n, l=1, 2, \dots, m \quad (10) \\ \frac{\partial Q_\lambda}{\partial \lambda_i} = \sum_{i=1}^n a_{ij} \phi_{ij}(x_0) - u_{i0} = 0 \end{cases}$$

在一组给定的基函数下,通过式(10),可解出最优泛函参数,从而得到一元非线性回归预测模型的有关参数。

3.2 基于泛函网络的多元非线性回归预测

3.2.1 多元非线性回归预测模型

根据多元非线性回归预测的特点,设计出一个多输入、单输出和三层中间处理单元的泛函网络模型,如图3所示。

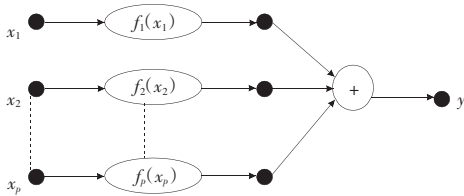


图3 多元非线性回归预测模型

在图3所示的模型中, x_1, x_2, \dots, x_p 为自变量,记 $X=(x_1, x_2, \dots, x_p)$, y 为因变量, $f_l(x_l), l=1, 2, \dots, p$,为神经元函数,不是固定的,而是可以学习的,是一个给定的函数族的组合。多元非线性回归预测模型可以看作是由一元非线性回归预测模型“拼装”而成的,可以由多个一元非线性回归预测模型来逼近。图3所示中 $f_l(x_l)$ 等同于图2中的输出 y ,与图3相对应的泛函方程为:

$$y = \hat{F}(x) = \sum_{l=1}^p c_l f_l(x_l) = \sum_{l=1}^p c_l \sum_{i=1}^n f_{li}(x_l) \quad (11)$$

其中,每个神经元函数可看作是一个一元非线性回归预测模型,要得到多元非线性回归预测模型的相关参数等价于学习 $\hat{F}(x)$,而学习 $\hat{F}(x)$ 等价于学习函数 $f_l(x_l)$,而学习 $f_l(x_l)$ 等价于学习 $f_{li}(x_l)$,可以用一组已知基函数 $\{\phi_{ij}(x_l); j=1, 2, \dots, m, i=1, 2, \dots, n, l=1, 2, \dots, p\}$ 的线性组合来表示 $f_{li}(x_l)$:

$$\hat{f}_{li}(x_l) = \sum_{j=1}^m a_{ij} \phi_{ij}(x_l) \quad (12)$$

令 $\phi_{li}(x_l) = (\phi_{li1}(x_l), \phi_{li2}(x_l), \dots, \phi_{lim}(x_l))^T, \mathbf{a}_{li} = (a_{li1}, a_{li2}, \dots, a_{lim})^T$,则式(12)可以写为:

$$\hat{f}_{li}(x_l) = \sum_{j=1}^m a_{ij} \phi_{ij}(x_l) = \mathbf{a}_{li}^T \phi_{li}(x_l) \quad (13)$$

把式(4)代入式(2)就可以得到:

$$y = \hat{F}(x) = \sum_{l=1}^p c_l \hat{f}_l(x_l) = \sum_{l=1}^p c_l \sum_{i=1}^n \sum_{j=1}^m a_{ij} \phi_{ij}(x_l) = \sum_{l=1}^p c_l \sum_{i=1}^n \mathbf{a}_{li}^T \phi_{li}(x_l) \quad (14)$$

其中 $\{\phi_{ij}(x_l); j=1, 2, \dots, m, i=1, 2, \dots, n, l=1, 2, \dots, p\}$ 是一些给定的适合逼近 f_{li} 到期望精度线性独立的函数集。 $\{a_{ij}(x_l); j=1, 2, \dots, m, i=1, 2, \dots, n, l=1, 2, \dots, p\}$ 为泛函网络的参数。

3.2.2 多元非线性回归预测模型的学习算法

对于多元非线性回归预测模型,事先给定一组样本,比如

$$\{(X^{(k)}, y^{(k)}) | k=1, 2, \dots, q\} \quad (15)$$

其中, $X^{(k)} = [x_1^{(k)}, x_2^{(k)}, \dots, x_p^{(k)}]$ 。采用简单的、收敛速度快的学习算法——基于最小二乘思想的学习算法,就可以得到泛函网络的相关参数,与上面的一元非线性回归预测相似,在此不再赘述。

4 应用研究实例

为了验证本文提出的基于泛函网络的回归预测模型及学习算法的有效性,定义均方误差(RMSE):

$$RMSE = \sqrt{\frac{1}{q} \sum_{k=1}^q \|y_k - \hat{y}_k\|^2} \quad (16)$$

式中 y_k 表示函数的真实值, \hat{y}_k 表示网络的输出值, q 表示学习样本数。

以下考虑两个具体例子,并对结果进行分析。

4.1 一元非线性回归预测实例

选用文献[8]的一组数据见表1,该数据表示不同年龄组与其远视率之间的关系。由于其关系的复杂性,二者之间是一种非线性关系。文献[8]通过作散点图,采用变量代换的方法,得到拟合曲线为:

$$\hat{y} = e^{5.730487858 - 0.313960199x} \quad (17)$$

其拟合优度: $R^2 = 1 - \frac{\sum (y - \hat{y})^2}{\sum (y - \bar{y})^2} = 0.779055$,其拟合值

见表1。本文采用表1中的原始数据,利用图2所示的一元回归预测模型和相应的学习算法,取基函数为 $\{1, e^{mx}\}$ 。运用MATLAB编程,在 $RMSE=0$ 时,经过化简可以得到不同年龄组与其远视率之间关系的表达式。6~12为拟合值,13~18为预测值。 $\hat{y} = e^{6.563248251 - 0.421436012x}$ 根据定义计算其拟合优度为: $R^2 = 0.952362$,其拟合值见表1。

表1 不同年龄组与远视率之间的原始数据与拟合预测结果

x	y	变量代换法	泛函网络法
6	63.64	46.839410	69.789523
7	61.06	34.218501	50.502364
8	38.84	24.998261	35.245423
9	13.75	18.262430	16.524215
10	14.50	13.341585	15.051275
11	8.07	9.746670	9.251532
12	4.41	7.120411	6.061524
13	2.27	5.021803	2.982651
14	2.09	3.800167	3.225394
15	1.02	2.776225	1.952631
16	2.51	2.028151	2.325427
17	3.12	1.481662	2.943251
18	2.98	1.082425	3.025413

从表1可以看出,泛函网络预测模型拟合和预测的精度明显高于变量代换的方法。并且求解算法简单。在求解一些非线性问题时,可以采用这种模型。

4.2 多元非线性回归预测实例

根据表2资料建立全国工商税收(Y)关于工业净产值(x_1)、运输业净产值(x_2)、商业净产值(x_3)的回归预测模型,原始数据和具体结果见表2。1978~1986为拟合数据,1987~1992为预测数据。

在表2中, Y_{c1} 为线性化回归模型的预测拟合结果, Y_{c2} 为泛函网络模型的预测拟合结果。如建立多元线性回归模型,则为:

$$Y_{c1} = 213.2169 + 0.1566x_1 - 0.6220x_2 + 0.6798x_3$$

$$R^2 = 0.9797, SE = 102.1299, DW = 0.7600, F = 177.0355$$

表2 1978-1992年全国工商税收
预测资料及其拟合结果 亿元

年份	Y	x_1	x_2	x_3	Y_{c1}	Y_{c2}
1978	451.29	1 487	118	294	493.29	448.38
1979	472.69	1 628	121	245	510.41	472.53
1980	501.35	1 804	126	247	532.42	512.42
1981	538.40	1 840	131	268	571.26	541.25
1982	604.59	1 948	147	231	537.25	596.31
1983	657.81	2 136	166	254	689.34	658.35
1984	770.51	2 516	205	377	726.41	785.23
1985	1 147.82	3 163	259	697	1 184.26	1 123.54
1986	1 248.56	3 573	320	732	1 332.15	1 305.26
1987	1 330.35	4 262	384	876	1 379.24	1 348.57
1988	1 541.62	5 416	460	1 261	1 505.14	1 572.31
1989	1 814.83	6 241	547	1 405	1 734.17	1 798.61
1990	1 895.35	8 610	787	1 148	1 928.15	1 902.34
1991	2 024.65	7 703	887	1 689	2 065.57	2 043.25
1992	2 277.32	9 805	957	1 813	2 231.84	2 254.23

注:表中数据来自于中国统计年鉴。

式中 Y_c 为 Y 的估计值,由于 x_2 前出现负号,经济意义得不到解释,且残差存在着自相关,此模型有待改进。

用本文提出的如图3所示的多元非线性回归预测模型和相应的学习算法。取基函数分别为: $\varphi(x_1)=\varphi(x_2)=\{1, \ln x\}$, $\varphi(x_3)=\{1, x\}$ 。运用 MATLAB 编程,求其泛函网络相关系数,在 $RMSE=0$ 时,经过化简可以得到:

$$f_1(x_1)=-6\ 734.498\ 2+970.254\ 6 \ln x_1$$

$$f_2(x_2)=-3\ 516.547\ 1+826.406\ 7 \ln x_2$$

$$f_3(x_3)=305.992\ 9+1.099\ 7 x_3$$

它们之间的关系表达式:

$$Y_{c2}=a+b \times f_1(x_1)+c \times f_2(x_2)+d \times f_3(x_3)=$$

$$\begin{matrix} -21.505\ 8 & +0.469\ 1 & f_1(x_1) & +0.233\ 2 & f_2(x_2) & +0.316\ 3 & f_3(x_3) & = \\ (-0.796\ 7) & (2.079\ 5) & & (0.962\ 4) & & (4.105\ 2) & & \\ -3\ 904.116\ 2 & +455.156\ 0 & \ln x_1 & +192.747\ 5 & \ln x_2 & +0.347\ 9 & x_3 & \\ R^2=0.995\ 5, SE=48.133\ 8, DW=2.179\ 4, F=809.853\ 9 \end{matrix}$$

本文提出的模型和学习算法与线性化模型得出的结果比较如表2,从上述表达式可以看出,本文基于泛函网络的模型和学习算法,修正了线性回归预测的不合理系数,在预测上,更重要的是各统计检验量均有所改进,相关系数变大,标准误差减少,残差无自相关。它们的拟合图见图4。从表2、图4可以看出本文提出的模型和学习算法的拟合预测结果明显高于线性化模型,因而此模型比较理想,可用此模型进行预测。

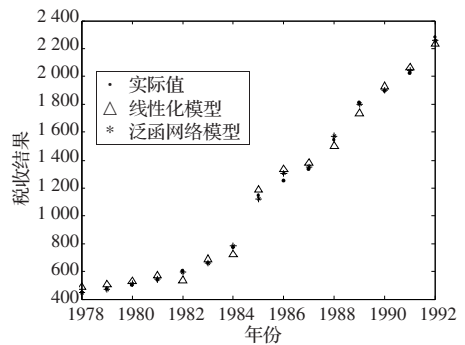


图4 预测结果比较图

5 结束语

传统回归预测在处理非线性回归预测时用线性模型逼近,在系统非线性比较强时,这种方法受到限制,结合泛函网络在处理非线性问题时的优势,分别就一元非线性回归预测和多元非线性回归预测提出一种基于泛函网络的非线性回归预测模型并给出相应简单的、收敛速度快的学习算法。并分别与传统处理方法作比较,验证了该算法的有效性。从以上实例可以看出:对于一些复杂模型的回归预测问题能够用泛函网络方法去解。复杂模型是指模型中变量的关系无法确定或很难确定,用传统方法很难求解。但是也应该看出,在处理一些简单的回归预测问题如简单的一元线性回归预测问题时,传统方法有一定的优势。本文只是对泛函网络用于回归预测进行了初步的尝试,还有许多问题如用到具体的预测问题中将在后续工作中完成。

参考文献:

- [1] 约翰·内特.应用线性回归模型[M].张勇,译.北京:中国统计出版社,1990.
- [2] 汤振民.非线性回归线性化处理中的加权[J].数理统计与管理,1998,2:41-47.
- [3] 陶菊春.可线性化非线性回归预测模型的剖析与改进[J].数学的实践与认识,2003,33(2):7-12.
- [4] Castello E.Functional networks[J].Neural Proc Lett,1998,7:151-159.
- [5] Castello E,Gutierrez J M.Nonlinear time series modeling and prediction using functional networks,extracting information masked by chaos[J].Phys,Lett A,1998,244:71-84.
- [6] Castello E,Cobo A,Gutierrez J M.Working with differential functional and difference equations using functional networks[J].Appl Math Model,1999,23:89-107.
- [7] Castillo E,Cobo A,Gutierrez J M.Functional networks with applications[M].[S.l.]:Kluwer Academic Publishers,1999.
- [8] 韦博成.非线性模型[J].数理统计和应用概率,1986(1):105-144.

(上接73页)

误差的原因,提出了一种根据运动矢量的空间相关性,确定误差较大的运动矢量,并对其进行自适应修正的算法,改善了图像质量,降低了码率。

参考文献:

- [1] Youn J Y,Sun M T,Lin C W.Motion vector refinement for high-performance transcoding[J].IEEE Trans Multimedia,1999,1:30-40.
- [2] Chen Mei-Juan,Chu Ming-Chung,Pan Chih-Wei.Efficient motion-

estimation algorithm for reduced frame-rate video transcoder[J].IEEE Trans Circuits Syst Video Technol,2002,12:269-275.

- [3] Yusuf A A,Murshed M,Dooley L S.An adaptive motion vector composition algorithm for frame skipping video transcoding[C]//Electrotechnical Conference,MELECON 2004,Proceedings of the 12th IEEE Mediterranean,12-15 May 2004,1:235-238.
- [4] Patil V,Kumar R.An effective motion re-estimation in frame-skipping video transcoding [C]//ICCTA'07,International Conference on Computing:Theory and Applications,March 2007:655-659.