第19卷 建模专辑

工程数学学报

2002年02月

JOURNAL OF ENGINEERING MATHEMATICS

Vol. 19 Supp. Feb 2002

文章编号:1005-3085(2002)05-0067-08

公交车调度的规划数学模型

薄立军, 要尉鵬, 王艳辉 指导老师: 刘红卫 (西安电子科技大学,西安 710071)

编者按:本文建立了两种优化模型来研究公交车调度问题。第一种模型中使用 Fisher 聚类算法对客流分布进行了优化分类,这使得客流时间段的划分更为合理。第二种模型基于随机服务系统,主要利用了 GI/M/n 排队系统的平均队长及平均等待时间等基本公式。因城市交通客流是随机的,利用排队理论来研究公交车调度问题更能刻划问题的实质。但单交通线上的公交车具有串联服务的性质,这与 GI/M/n 系统不大符合。第二种模型有明显的不足。

摘 要: 本文根据有序样本聚类的 Fisher 算法,给出一种峰值曲线的优化方法,通过该方法我们得出了上行客流峰值为 5 个,其峰值区间为:5:00-6:00,6:00-9:00、9:00-16:00,16:00-18:00,18:00-23:00;下行客流峰值为 5 个,其峰值区间为:5:00-7:00,7:00-9:00,9:00-16:00,16:00-19:00,19:00-23:00。

然后,依据峰值区间建立确定发车间隔的算法 I 模型和算法 II 模型,对两种算法模型计算结果进行比较分析,得出结论:两个间隔高峰类时间段用算法 I 进行求解,其余 3 个类时间段用算法 II 进行求解。在各个时间段结合处用光滑法进行优化处理,并以处理后的数据为基础制定出两个起点站的发车时刻表,并求出全线共需要 47 辆车,乘客对方案的满意程度为 98.2%、公交公司的满意程度为 76.23%。

最后,运用随机服务系统的相关理论建立随机规划模型,给出概率灵敏度和误差分析,进而得出采集运营数据的较好方案。

关键词: 有序样本聚类; 客流; 峰值; 车次; 平滑法; 随机服务

分类号: AMS(2000) 90C08

中图分类号: TB114.1

文献标识码:A

1 问题重述(略)

2 基本假设

- 1) 公交车在该线路行进中以 20 公里/小时的速度匀速运行,即不考虑启动和停车,每一站停车延迟及其他因素的影响
 - 2) 公交车按发车时刻表顺次发车,准时到达每个站点
 - 3) 乘客候车时间一般不超过10分钟,早高峰时一般不超过5分钟
 - 4) 满载率不要超过120%,一般也不应低于50%

3 符号说明

 P_{i} : i 时段内的配车数(时段配车数)(车次);

 ρ_i : i 时段内的期望满载率;

 H_{i} : i 时段内的小时最高断面的通过量(人);

N: i 时段内的期望占用量(人);

C: 车容量(C=车型定员+最大允许站人数)(人); L: 路线长度(km);

 Q_i : i 时段内的乘客周转量(人 km);

 δ_i : i 时段内乘客的满意率:

δ: 乘客的平均满意率:

W.; i 时段内乘客的平均等待时间

η: 公交公司的平均满意率;

4 问题(1)模型的分析、建立及求解

下面我们逐步以两种不同的方法对公交调度方案进行讨论,第一种方法对公交调度峰值 曲线进行优化,第二种方法对公交调度发车间隔进行确定,进而制定出公交调度方案。最后, 依据制定方案过程中的相关参数的随机特性,抽象出明确完整的随机规划模型。

方法 I 优化公交调度峰值曲线

公交调度人员在制定线路配车计划时,最为重要的依据是线路客流的每日时段分布曲线。调度人员进行这样的峰值划分主要依据以下两个原因:①对于客流大小相似而且相邻的时段配置相同的运力;②划分为若干峰值区间,便于进行驾乘人员的班次安排。(明确一点:公交调度峰值曲线中峰值代表一个区间,而不是一个点)。

公交调度峰值曲线的优化过程实际上是有序样品的聚类问题。

所谓有序样品是指,样品按照一定的要求排成序,分类时不能打破这种次序。设 x_1 , x_2 , …, x_n 表示一组有序的样品,则每一类必须呈 $\{x_i, x_{i+1}, \dots, x_j\}$ (i < j)形态。n 个有序样品分成 k 类的一切可能的分法有 C_n^{k-1} 种,这个数比 C_n^k 要小得多。因此在某种损失函数下,有可能求得最优解。费歇(Fisher)发展了一个有序样品的聚类算法,它可保证求得最优解。

下面给出费歇(Fisher)聚类算法:

① 基础算法

首先定义 D(i,j)表示类 $|i,i+1,\cdots,j|$ 的直径。类的直径 D(i,j)这里采用该类的值与 类均值差的平方和来表示直径的大小。

用 $b_{n,k}$ 表示 n 个样本分成 k 类的一种分法, 即

$$b_{n,k}$$
: $\{i_1 = 1, i_1 + 1, \cdots, i_2 - 1 | , \{i_2, i_2 + 1, \cdots, i_3 - 1\}, \cdots, \{i_k, i_k + 1, \cdots, n\}$ 其中, $i_1 = 1 < i_2 \cdots < i_k < n_0$

定义 6。4分类下的损失函数为:

$$L(b_{n,k}) = \sum_{i=1}^{k} D(i_i, i_{j+1} - 1)$$
 (1)

其中, $i_{k+1} = n+1$

损失函数值越小、分类越合理。设 $b_{n,k}^*$ 为使式(1)达到极小的解

费歇(Fisher)的计算方法使用下面两个递推公式:

$$L(b_{n,2}^*) = \min_{2 \le j \le n} |D(1,j-1) + D(j,n)|$$
 (2)

$$L(b_{n,2}^*) = \min_{k \in \mathbb{Z}} \{L(b_{j-1, k-1}^*) + D(j, n)\}$$
 (3)

在 k=2 时、 $b_{n,2}:\{1,2,\cdots,j-1\},|j,j+1,\cdots,n\},2\leq j\leq n$ 。

由式(1)得 $L(b_{n,2}) = D(1,j-1) + D(j,n)$

最优的分法是上式对 $j(2 \le j \le n)$ 求极小,得到式(2)。为证式(3),只需注意到 n 个有序样品分成 k 类,这等价于将它们先分成两部分: $\{1,2,\cdots,j-1\}$, $\{j,j+1,\cdots,n\}$

其中 $\{1,2,\dots,j-1\}$ 将分成 k-1 类,而 $\{j,j+1,\dots,n\}$ 独成一类、显然 $k \leq j \leq n$,于得到式(3)

②算法实现

若 k(1 < k < n)已知,应求分类 $b_{n,k}^*$,使它的损失函数最小

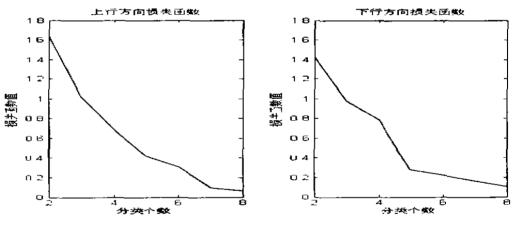
由(3)式,如果 k > 2, 找 j_k , 使 $L(b_{n,k}^*) = L(b_{j_k-1,k-1}) + D(j_k,n)$

于是得到第 k 类 $P_k^* = \{j_k, j_k + 1, \dots, n\}$,然后找 j_{k-1} ,使它满足 $L(b_{j_{k-1},k-1}) = L(b_{j_{k-1},j_k-1}) + D(j_{k-1},j_k-1)$,得到第 k-1 类 $P_{k-1}^* = \{j_{k-1},j_{k-1}+1,\dots,j_k-1\}$

依次类推得到所有类 $(P_1^*, P_2^*, \dots, P_k^*) = b_{n,k}^*$ 。由式(2)和(3)可看出 $b_{n,k}^*$ 是最优解我们应用上面的费歇(Fisher)算法对公交调度峰值曲线进行聚类处理,具体步骤如下:

- 1) 先将各时段客流量转化为其占全线路客流的比例。本题中有 18 个时间段,即 18 个有序样本。
 - 2) 计算所有可能类的直径 $D(i,j) = \sum_{l=1}^{j} (x_l \overline{x_{ij}})^2$, 其中, $\overline{x_{ij}} = \frac{1}{j-i+1} \sum_{l=1}^{j} x_l$
- 3) 计算最小损失函数。用 $b_{i,j}^*$ 表示用前i 个样品分成j 类的最优解,它的最优损失函数为 $L(b_{i,j}^*)$ 。

当 $j \le i \le 18$, $2 \le j \le 8$ 时,利用费歇算法得到上下行方向的最小损失函数值变化曲线图 (1)



图[1] 最小损失函数值变化曲线

经分析计算得出:在分类数为 5 时,损失函数的值为 0.004204 和 0.002762,损失函数已经下降的足够小了。在分类数为 6 时,损失函数的值为 0.003076 和 0.002211。分为 5 类与分为 6 类,两者之间损失函数的差别不大。从方便调度管理的角度出发,期望划分尽可能少的类,所以取曲线最优聚类个数为 5 个。

在聚类个数为 5 时,可以得出,上行方向的最优分类为 5:00-6:00.6:00-9:00.9:00-16:00.16:00-18:00.18:00-23:00; 下行方向的最优分类为 5:00-7:00.7:00-9:00.9:00-16:00.16:00-19:00.19:00-23:00。

方法Ⅱ 确定公交调度发车间隔

我们通过引入时段配车数的概念,来探讨在不同客流状态时如何确定时段配车数和发车 间隔。

定义 在某一时间段内需求的车辆数称之为**时段配车数**。确定原则是,既保证有足够的服务质量,又保证配车数最小。

下面给出两种算法模型,

算法模型
$$I$$
 $P_i = \frac{H_i}{\rho_i \times C} = \frac{H_i}{N_i}$ 算法模型 I $P_i = \max \left| \frac{Q_i}{\rho_i \times C \times L}, \frac{H_i}{C} \right| = \max \left| \frac{Q_i}{N_i \times L}, \frac{H_i}{C} \right|$

我们对确定发车间隔的模型采用两种不同的间隔确定方法进行求解,综合评价后得出综合算法模型:(假设每小时被调查的上车人数基于均匀的达到率)

i) 参数分析

 H_i : 对上行方向和下行方向分别计算,将每一时间段内每一站点的上车人数减去下车人数得到该时间段内该站点净上车人数,然后从每一时间段起始站点开始累加得到每一站点的小时通过量 D_{ij} (i 时间段内第j 个站点的小时通过量),那么i 时间段内的 $H_i = \max_{1 \le j \le n} (D_{ij})$,n 是上行或下行方向的车站总数

$$C:$$
 依据定义有 $C=100\times120\%=120(人)$ $N_i:$ $N_i=C\times\rho_i$ $L:$ 上行方向 $L=14.58(km)$,下行方向 $L=14.61(km)$

$$Q_i: Q_i = \sum_{j=1}^n d_{ij}D_{ij}$$
, d_{ij} 是 i 时间段内第 j 个站点与其相邻的前一个站点间的距离

 ρ_i : 我们对 ρ_i 的估计值是按综合考虑每一时间段内的总客流量在全天的总客流量所占的比率以及一些可以查到的经验值拟合出 ρ_i 与 H_i 的分布关系函数: $\rho_i = 0.0173 \times H_i + 46.6919$

ii) 结果分析

分别按照算法模型 I 和算法模型 II 计算每个时间段内的发车次数。对这些数据分析可知: 算法 I 和算法 II 各自在不同的类时间段会出现不稳定的结果,在每个类时间段单独使用一种算法得出的结果都会具有一定的不可信度。以不同类时间段上的算法模型 I 及算法模型 II 的适应程度作为评价标准,得出确定发车间隔的综合算法模型。

综合算法模型:在两个间隔高峰类时间段:上行方向类时间段 | 2, 3, 4 | , | 12, 13 | 及下行方向类时间段 | 3, 4 | , | 12, 13, 14 | 采用算法 Ⅰ进行求解,而在上下行方向的其余各 3 个类时间段则采用算法 Ⅱ求解。

优化模型Ⅲ平滑法

为确定相邻两个时段之间的转换段内的发车时间,我们采用平滑法(所谓平滑法就是根据计算的时段配车数,在前一时段内确定第一辆车的发车时间,在转换段内综合前后两种配车数,设置平均期望占用量而不是不均匀间隔)。

例如我们所得结果中、上行方向 5:00-5:59, 6:00-6:59 两个时段内(第一辆车为 5:00 发车),时段配车数和发车间隔分别为 8.4260 车次,7.12min;29.9505 车次,2.00min。前一时段所需要的配车数的 0.4260 车被留在 5:57 之后与下一时段的 0.5740 车结合。此外,0.4260 车的期望占用量为 $N_i = \rho_i \times C$,0.5740 车的期望占用量为 $N_{i+1} = \rho_{i+1} \times C$,后一时间段每分钟需求的配车数(斜率)为 29.9505/60,相应的 0.5740 车要运行 0.5740/(26.9505/60) = 1.17min。因此第二时段内第一辆车的发车时间为 6:01:17,其期望占用量为 $0.4260 \times N_i + 0.5760 \times N_{i+1}$ 。

通过以上的分析及数据求解,我们制定出上行方向和下行方向两个起点站的发车时刻表(略),并求出需要的车辆总数(即最小配车数)。我们求得:上行方向为:47 下行方向为:

维普资讯 http://www.cqvip.com

35 故所需车辆总数 $N = \max(47,35) = 47$

根据题意,我们定义 i 时段内乘客的满意率为;

$$\delta_{i} = \begin{cases} 100\% & W_{i} < 5 \\ \frac{5}{W_{i}} & W_{i} \geqslant 5 \end{cases}$$
 高峰期
$$\begin{cases} 100\% & W_{i} < 10 \\ \frac{10}{W_{i}} & W_{i} \geqslant 10 \end{cases}$$
 非高峰期

乘客的平均满意率 $\bar{\delta} = \frac{\sum H_i \cdot \delta}{\sum H_i}$, 计算得 $\bar{\delta} = 98.2\%$

公交公司的平均满意率
$$\overline{\eta} = \frac{\sum H_i \cdot \rho_i}{\sum H_i}$$
, 计算得 $\overline{\eta} = 0.5(\overline{\eta_{up}} + \overline{\eta_{down}}) = 76.23\%$ 其中,上行 $\overline{\eta_{up}} = 78.36\%$,下行 $\overline{\eta_{down}} = 74.09\%$

5 问题(2)基本假设

1) 某站乘客到达为泊松过程,设上(下)行各站乘客的到达过程为平稳的泊松过程,乘客到达率为λ,τ;为第i个乘客到达时刻,则乘客到达的间隔为

$$P\{\tau_{t+1} - \tau_{t} < t \mid = \begin{cases} 1 - e^{-\lambda t}, & t \ge 0 \\ 0, & t < 0 \end{cases}$$
 (4)

又设公交车每批运送的乘客数量为 k,该系统可变换为一等效的新系统。由爱尔兰分布与负指数分布之间的关系可得,新系统的乘客到达间隔服从参数 λ 的 k 阶爱尔兰分布:

$$P\{\tau_{m+1} - \tau_m < t \mid = \begin{cases} 1 - \sum_{i=0}^{k-1} e^{-\lambda t} \frac{(\lambda t)^i}{i!}, \ t \ge 0 \\ 0, \ t < 0 \end{cases}$$
 (5)

式中, τ_m 为新系统第m 组乘客到达时刻。

2) 公交车对每批乘客服务时间为相互独立的负指数分布,每批服务时间是指公交车从起始站出发再返回起始站所需的时间为 t_{RT} 。设公交车成批服务时间 t_{RT} 服从参数 μ 的负指数分布,即

$$P\{t_{RT} < t \mid = \begin{cases} 1 - e^{-\mu t}, \ t \geqslant 0 \\ 0, \ t < 0 \end{cases}$$
 (6)

式中, μ 与 t_{RT} 的关系为; $\mu = \frac{1}{E[t_{RT}]} = \frac{1}{T_{RT}}$, T_{RT} 为运行一个行程的平均时间。

6 问题(2)模型建立与求解

定理 1 在 GI/M/n 系统中,设各顾客的服务时间相互独立且具有公共的以 μ 为参数的负指数分布,则在该顾客等待的时间内,每台服务设施的输出过程(即服务完成离开服务机构的顾客)是一个以 μ 为强度的泊松过程。(参考文献[1])

k 阶爱尔兰分布的分布函数 A(t) 见(5)式,由(5)得到 k 阶爱尔兰分布的概率密度为

$$\alpha(t) = \dot{A} = \frac{\lambda(\lambda t)^{k-1}}{(k-1)!}e^{-\lambda t}$$
 (7)

则到达间隔的期望值为:

$$E[\tau_{m+1} - \tau_m] = \int_0^\infty t dA(t) = \frac{k}{\lambda}$$
 (8)

由(5),(8)得概率密度的拉氏变换为

$$A^*(s) = \int_0^\infty e^{-st} dA(t) = \left(\frac{\lambda}{\lambda + s}\right)^k, \ res(s) > 0$$
 (9)

由排队理论, 可得到公交车服务系统的服务强度:

$$\rho = \frac{E[\nu_t]}{E[\tau_{m+1} - \tau_m]} = \frac{T_{RT}\lambda}{nk}$$
 (10)

其中: k 为爱尔兰分布的阶数(公交车每批运送的乘客数), n 为公交车数量, $E[\nu_k]$ 为系统服务时间的期望值, $E[\tau_{m+1}-\tau_m]$ 为第 n+1 批乘客与第 m 批乘客的到达间隔的期望值。

对于服务强度 ρ<1有下面的定理:

定理 2 1) 设 q_m 为第 m 个乘客到达时系统的队长,于是队长的分布极限存在且与初始条件无关,其表达式为:

$$\pi_{j} = \begin{cases} \sum_{r=j}^{n-1} (-1)^{r-j} {r \choose j} U_{r}, & j = 0, 1, \dots, n-1; \\ K\theta^{j-n}, & j = n, n+1, \dots, \end{cases}$$

其中 θ 为方程 $A^*(n\mu(1-\theta)) = \left(\frac{\lambda}{\lambda + n\mu(1-\theta)}\right)^k$ 在(0,1)之间的唯一解。

利用(7)式将上式化为 ρ 的方程:

$$\theta \left(1 + \frac{1-\theta}{k\rho}\right)^k = 1 \tag{11}$$

且

$$U_{r} = KC_{r} \sum_{k=r+1}^{n} \frac{\binom{n}{k} \left[n(1-\varepsilon_{k}) - k \right]}{C_{k}(1-\varepsilon_{k}) \left[n(1-\theta) - k \right]}, \quad r = 0, 1, \dots, n-1$$
 (12)

$$K \equiv \left[\frac{1}{1-\theta} + \sum_{k=1}^{n} \frac{\binom{n}{k}}{C_{k}(1-\varepsilon_{k})} \frac{n(1-\varepsilon_{k})-k}{n(1-\theta)-k} \right]^{-1}$$
 (13)

$$\begin{cases} C_0 \equiv 1, \\ C_k \equiv \prod_{l=1}^k \frac{\epsilon_l}{1 - \epsilon_l}, & k = 1, 2, 3, \dots, n \\ \\ \epsilon_l \equiv A^*(l\mu) = \left(\frac{\lambda}{\lambda + l\mu}\right)^k = \left(\frac{\rho}{\rho + \frac{l}{kn}}\right)^k \end{cases}$$
(14)

2) 该系统平稳分布下的平均队长为:

$$E[q] = \sum_{j=1}^{\infty} j\pi_j = K \frac{\theta + n(1-\theta)}{(1-\theta)^2} + (1-\delta_{n1})U_1$$
 (15)

其中: U_1 , K, θ 分别由(12), (13), (14)决定, $\delta_{n1} = \begin{cases} 1, & n = 1 \\ 0, & n > 1 \end{cases}$

31 系统的平均等待队长为:

$$E[qw] = \frac{K\theta}{(1-\theta)^2} \tag{16}$$

4) 系统服务台平均占有数:

$$E[l] = E[q] - E[qw]$$

$$(17)$$

定理 3 对 GI/M/n 系统,在平衡状态下,顾客到达时不需等待的概率为:

$$W(\theta) = 1 - \frac{K}{1 - \hat{\theta}} \tag{18}$$

平均等待时间为:

$$\overline{W} \equiv \int_{0}^{\infty} x dW(x) = \frac{K}{n\mu (1-\theta)^2}$$
 (19)

其中 $K \cdot \theta$ 分别由(13), (11)式确定。

我们用系统的平均等待队长作为衡量乘客满意度的标准,用系统服务台平均占有数作为 衡量公交公司利益的标准,因此,在任意时间段内都可以建立目标规划模型:

- 1) 设置变量: n——所需车辆数; K——车容量;
- 选择目标: P₁——第一目标: 尽可能使顾客满意。
 P₂——第二目标: 尽量使公司利益最大。
- 3) 约束条件的制定:
- i) 乘客在非高峰期等待时间不超过 10min, 在高峰期等待时间不超过 5min。
- ii) 车容量在区间(50,120)

综上我们得到随机规划模型为:

$$\min_{\mathbf{E}[qw]} E[qw] = \frac{K\theta}{(1-\theta)^2};$$

$$\min_{\mathbf{E}[t]} E[q] - E[qw];$$

$$K \in (50, 120);$$

$$\overline{W} \equiv \int_0^\infty x dW(x) = \frac{K}{n\mu(1-\theta)^2} \leq 5($$
 10);

其中 K, θ 由式(13),(11)决定

对上述模型进行系统分析得到以下结论:

- 1) 基于对模型结果的分析,服务强度 ρ 对系统的性能指标起着至关重要的作用。由于 $\rho = \frac{T_{RT}\lambda}{nk}$. 而采集数据对 T_{RT} 无影响,故所采集的数据应尽量体现 λ 的实际情况即某站在某期间内指定时间段内的到达人数最大值。
- 2) 对不同的系统参数下得出的重要指标比较后,得出:系统公交车数量对系统性能的改善的灵敏度比公交车容量对系统性能改善的灵敏度大。
- 3) 我们所建立的公交车服务系统模型还可应用于系统的优化配置,只要赋予明确的优化目标函数,就可得到相应的配置参数。同时为计算机仿真提供理论基础。

7 模型的推广

考虑约束条件下的间隔确定。例如以下两种约束条件

- i) 资金有限,应配备的车辆数不足或当日应配车辆发生事故不能上岗又无替补车辆供给时,调度员应及时考虑时刻表的变动。
 - ii) 当运力充足时,可以增加总发车次数,提高对乘客的服务质量。

参考文献:

- [1] 徐光輝. 随机服务系统(第二版)[M]. 北京:科学出版社,1486
- [2] 周义仓、赫孝良、数学建模实验[M] 西安:西安交通大学出版社,1999
- [3] 李 涛、贺勇军、刘志俭 Matlab 工具箱应用指南——应用数学篇[M], 北京:电子工业出版社、2000

The Optimum Mathematical Model on the Bus Dispatch

BO Li-jun, YAO Wei-peng, WANG Yan-bui Adviser: LIU Hong-wei (XiDian University, Xi'an 710071)

Abstract: This paper presents an optimal method of peak curve according to the Fisher algorithm of serial specimen clustering. We conclude 5 uphill passenger-flow peak ranges: 5:00-6:00, 6:00-9:00, 9:00-16:00, 16:00-18:00, 18:00-23:00, and 5 downhill passenger-flow peak ranges: 5:00-7:00, 7:00-9:00, 9:00-16:00, 16:00-19:00, 19:00-23:00.

Then, under the peak ranges, two algorithm models, I and II, are established. Comparing the calculated results of the foregoing models, we conclude: model I is applied to two interval high peaks, and model II is applied to three others. With the smooth method between every two time-sections, we make the bus time schedule of two starting stations, and get 47 needed buses at least. In this scheme, passengers' satisfaction rate is 98.2%, and the bus ecompany's is 76.23%

By the end, we set a random optimum model by the theory of random service system, and give the probability sensitivity and error analysis. Further, we get a better scheme for collecting operation data

Key words: serial specimen culstering; passenger-flow; peak; bus number; smooth method; random service

(上接 53 页)

Reestablishment of three dimensional blood vessel

DING Feng-ping, ZHOU Li-feng, LI Xiao-peng
Adviser: Mathematical Modeling Tutor Group
(Zhejiang University of Technology, Zhejiang Hangzhou 310032, China)

Abstract: The reestablishment of the three dimensional blood vessel is presented in the article. According to the information given by the problem, 100 pieces of sliced sheet of blood vessel are inputted into the program and transformed into data matrixes. Then three steps are given to reestablish the blood vessel. Firstly, the radius of the blood vessel is obtained by searching the biggest inscribed circle of the sliced sheet and here two solutions are given by using tangent method and the biggest overlay. Secondly, the track of the centre of the scrolling ball is hunted by grid method, Monte Carlo method and non – linear optimization method respectively. Thirdly, the projection of the central axes is positioned precisely on three planes. At last verifying of the reestablished blood vessel and error analysis are carried out to test the precision of the model.

Key words: reestablishment of three dimensional image; track; overlay