

A Type Design Method of Large Hydraulic Turbine Based on The Gaussian Mixture Model

Abstract. Because of the complex fluid motion in hydraulic turbine and the imperfect design theory, the selection design of Large-scale hydraulic turbine is achieved based on the calculation and analysis on the synthetic characteristic curves, which is subjectivity and low efficiency. To solve this problem, the Gaussian mixture model is used to extract the geometric features from the synthetic characteristic curves so that the retrieval process of the model wheel can be achieved by these geometric features. The search model of the running area from the synthetic characteristic curves is build based on the contour curve similarity transformation method. In the paper, the Monte Carlo method is adopted to obtain the mean values of the synthetic characteristics in the running area so that the evaluation targets can be established by combining the mean values with the hydraulic turbine design experiences. Finally the validity of running area can be evaluated by the evaluation targets. The test results show that the accuracy and efficiency of the selection design of Large-scale hydraulic turbine can be improved by the proposed method.

Streszczenie. W artykule przedstawiono problemy projektowania dużych turbin hydraulicznych. Zaproponowano model matematyczny bazujący na mieszanym modelu Gaussa w celu wydobycia parametrów geometrycznych. Zaadaptowano metodę Monte Carlo do określania średnich wartości syntetycznych charakterystyk w obszarze pracy. (Metoda projektowania dużych turbin hydraulicznych bazująca na mieszanym modelu Gaussa)

Keywords: Hydraulic turbine, Selection design, Synthetic characteristic curves, Gaussian mixture model.

Słowa kluczowe: turbiny hydrauliczne, metoda Monte Carlo, projektowanie.

1. Introduction

In the design of Large-scale hydraulic turbine, the selection design is the most important step. Because of the complex fluid motion in the turbine and imperfect design theory, it is still necessary to combine the theoretical with the experiments to design the over-fluid of the turbine. Based on the turbine design similarity theory, with ensuring geometric similarity, the fluid motion similarity and dynamic similarity, the structure of Large-scale prototype turbine is reduced to the Small-scale model turbine[1]-[2]. The experimental data and results on the model turbine are obtained first, and then they are used for the design of the prototype turbine. Model turbine experimental data is obtained by testing the model wheel under the any working conditions, and then the data is analyzed based on the similarity rules to achieve the synthetic data, which is used to plot the synthetic characteristic curves (SCC) of the model wheel. Therefore, a wheel model corresponds to a set of SCC. The selection design of the turbine is to retrieve a wheel model used in the history power station, which is similar to the current power station. The validity of the wheel model for the current power are analyzed and evaluated.

The analysis includes two steps: first, the running area is searched in the SCC; secondly, the running area is analyzed whether it is validity for the new power. The running area in SCC is an area surrounded by the extreme condition points retrieved by the search algorithm. Because of the incapacity of establishing the mathematics model for the SCC, the expert experiences and the calculation on the working conditions are used for the evaluation of the running area, which is not only time consuming, but also because of subjecting to the experts own experience, the human error may be introduced.

These are two aspects studied in this paper: the graphics features of SCC are extracted by the Gaussian mixture model, and the features are used for the evaluation of SCC. The Monte Carlo method is used to calculate the mean efficiency $\bar{\eta}$, the mean opening \bar{o} and the mean cavitation coefficient $\bar{\sigma}$ within the area enclosed by the 3D surface rebuilt by the graphics features. Considering the design experts' experience on judging the validity of the running area, based on the evaluation targets established by $\bar{\eta}$, \bar{o} and $\bar{\sigma}$, the judging process can be achieved automatically. Tests show that the results of selection design using the

proposed method are same as the ones of expert selection. More importantly, the efficiency is greatly improved.

2. The Analysis of the SCC

2.1 The Description of the SCC .

As mentioned above, the experimental data of hydraulic turbine model wheel is used to plot the SCC. Fig.1 shows the SCC which includes the efficiency curves, the opening curves, the output limiting curves and the cavitation coefficient curves. These curves describe the synthetic performance of the model wheel.

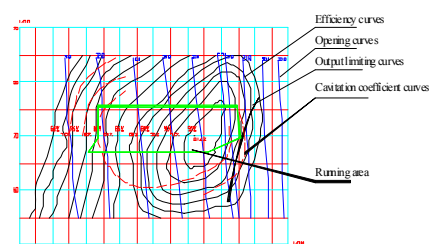


Fig.1. the SCC of turbine model wheel

Each type of the curves in SCC expresses the relationship between a characteristic parameter and the speed and flow. For example, the efficiency curves express the relationship between the efficiency and unit speed and unit flow.

$$(1) \quad \eta_m = f(Q, n)$$

where η_m - the efficiency of model wheel ; Q - the unit flow ; n - the unit speed.

The SCC describes, with the way of the contour curves, the distribution of the wheel characteristics within the area produced by the turbine flow and speed. Therefore they are the 3D traces of the body. The efficiency curves are composed of n contour curves, that is,

$$(2) \quad \eta_m = \{ \eta_i | n = y_i(Q) \} \quad i=1 \dots n$$

The SCC is the basis for the selection design. General design methods are: the main parameters of the power stations are saved in the database, such as design head, the water head scope and the capacity and so on. For the new design, the history power with the similar parameters is

found, then, the model wheel used in the history power is retrieved, finally, the SCC of the model wheel is analyzed whether the model wheel is applicable for the new power. This approach is of large limitations, although some wheels are applied in few powers, even more applicable in the current powers, they may not be selected because of the experts' error. If the geometrical features of the SCC are used to search the model wheel, there is the widest range to search the high-performance wheel for the current power.

2.2 The Feature Extraction of the SCC by the Gaussian Mixture Model.

It is necessary that the geometrical features are extracted for each type of curves respectively. Because the SCC is composed of a set of contour curves, there are unavoidable a large amount of the geometrical feature parameters. It can be derived from the continuity of 3D surface described by the contour curves that 2 contour curves with the small high differences are of the smaller shape and the size differences between them. The shape of a contour curve is of a certain similarity with the two adjacent contour curves. Therefore, the shape can be obtained by weighting the shapes of the two adjacent contour curves. Finally, because the designer usually concerned with the high-performance area in the SCC, only parts of the SCC are chosen for the feature extraction. Some algorithms can be used to extract the geometrical features of the curves, and then the geometrical features can be used to rebuild SCC.

Gaussian mixture model[3]-[4] is a strong tool for the probability modeling. For the feature extraction of the curve, Gaussian mixture model can be used to decompose curve into the linear combination of several Gaussian distributions. Dempster [5] proposed the EM algorithm in 1977, and then, followed by Redner and Walker[6] and other studies, EM algorithm became the main algorithm for the maximum likelihood estimation of Gaussian mixture model.

In the selection design process, generally, the designers pay more attention to the regional features of the SCC rather than to the accurate features of the SCC. Gaussian mixture model can be used to extract the features with less parameter, which can better describe the regional features of the SCC. Supposing $S = \{s_1, s_2, \dots, s_n\}$ is a set of samples with d -dimensional random distribution, $s_i = [s_i^1, s_i^2, \dots, s_i^d]$ is a sample in S , supposing the probability distribution of the sample obeys $f(s)$, then $f(s)$ can be expressed by.

$$(3) \quad f(S) = \sum_m \alpha_m p(s | \theta_m)$$

Among them, $\alpha_1, \alpha_2, \dots, \alpha_m$ is the probability of each distribution composition, θ_m is the distribution parameter of the composition m , $\theta_m = \{\theta_1, \dots, \theta_k, \sigma_1, \dots, \sigma_k\}$ is the set of the distribution parameters. If

$$\alpha_m \geq 0, m = 1, \dots, k \text{ and } \sum_{m=1}^k \alpha_m = 1, p(s | \theta)$$

obeys Gaussian distribution, then $f(S)$ means the Gaussian distribution model while the distribution is decided by α_m, θ_m . It is proposed that the minimum length of information (MML) criteria[7] is used as a variable of EM algorithm. For example, the 93% ~ 64% efficiency curves of

the SCC shown in Figure 2 are used for geometrical feature extraction. The results are shown in Table 1.

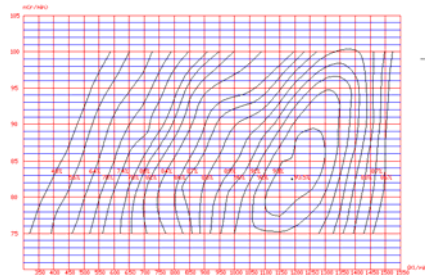


Fig.2. Efficiency curve in the turbine model

All the Gaussian models are saved in the database, which are used for the efficiency curve retrieval and comparison. Each efficiency curve can be rebuilt by Eq. (4).

$$(4) \quad f(n, Q) = \begin{cases} \frac{1}{2\pi\sigma_1\sigma_2} \exp\left(-\frac{1}{2}\left[\left(\frac{n-\mu_1^1}{\sigma_1^1}\right)^2 + \left(\frac{Q-\mu_2^1}{\sigma_2^1}\right)^2\right]\right) \\ \frac{1}{2\pi\sigma_1^2\sigma_2^2} \exp\left(-\frac{1}{2}\left[\left(\frac{n-\mu_1^2}{\sigma_1^2}\right)^2 + \left(\frac{Q-\mu_2^2}{\sigma_2^2}\right)^2\right]\right) \\ \dots \\ \frac{1}{2\pi\sigma_1^3\sigma_2^3} \exp\left(-\frac{1}{2}\left[\left(\frac{n-\mu_1^3}{\sigma_1^3}\right)^2 + \left(\frac{Q-\mu_2^3}{\sigma_2^3}\right)^2\right]\right) \end{cases}$$

Definition 1. Given 2 curve functions: $f_1(x), f_2(x)$, then

$d(f_1, f_2) = \int_{c_0}^{c_1} |f_1(x) - f_2(x)| dx$ is defined as the distance between $f_1(x)$ and $f_2(x)$. $[c_0, c_1]$ is the domain. If given ε , $d(f_1, f_2) \leq \varepsilon$, then $f_1(x)$ is similar with $f_2(x)$.

During the retrieval process of the SCC established by Eq.(4), only the areas surrounded by a certain range of flow and speed are used for the SCC comparison. It means only the points in the interval $[c_0, c_1]$ are selected for the comparison. $d(f_1, f_2) \leq \varepsilon$ is used to obtain the comparison between the different SCC.

2.3 The Establishment of the Running Area in the SCC.

The running area of the SCC is determined by the efficiency curves, that is, the products of the efficiency and the flow equal to the given values are used to determine the boundary points of the running area. The solving process of the point is, Given the corresponding parameters, the head H_i and the wheel diameter D_1 under the maximum head H_{max} , design head H_r and Minimum head H_{min} , the output P_i of turbine in the condition i is,

$$(5) \quad P_i = 9.81\eta_i H_i Q_i' D_1^2 \sqrt{H_i}$$

where η_i — the efficiency of the condition i ; Q_i' — the unit flow of the condition i ; n_i' — the unit speed of the condition i . The unit speed n_i' of the condition i is,

$$(6) \quad n_i' = n D_1 / \sqrt{H_i}$$

where n — given Rated speed. The $\eta_{i0} \times Q_{i0}'$ and the unit speed n_{i0}' can be obtained under the given parameters P_i, H_i and D_1 by the Eqs.(5)-(6). The horizontal line ($n_i' = n_{i0}'$) is called Unit Speed Line(USL).

The vertical line ($Q = Q_i'$) is called the Unit Flow Line(UFL), which intersects the USL at η_i . The Eq. (7) is established.

$$(7) \quad \eta_i \times Q_i' = \eta_{i0} \times Q_{i0}'$$

Table .1 The Gaussian Models of Efficiency curves

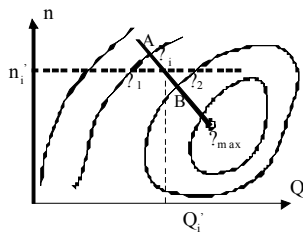
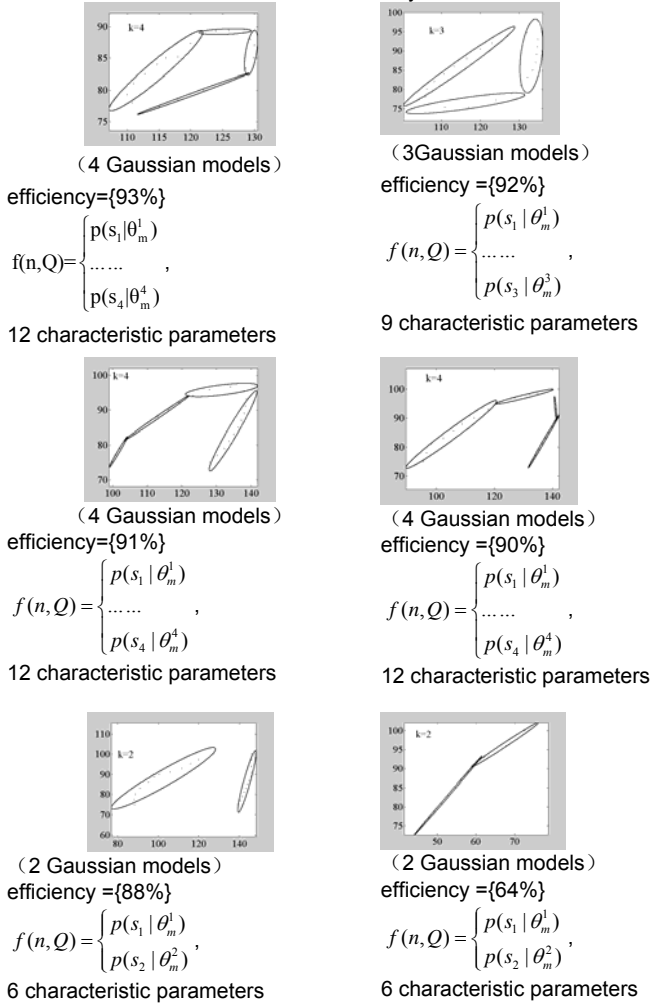


Fig.3. Running point in running area

The point $(\eta_i \times Q_i', n_i')$ in the SCC is the correct running condition point for the turbine. Several running condition points can surround the running area of SCC for the current power station. Because $\eta_{i0} \times Q_{i0}'$ are the fixed value under the condition parameters P_i , H_i and D_i , and the efficiency m is respectively increasing and decreasing at the right and left of the highest efficiency. The boundary points of the running condition are searched from the right and left sides respectively in the paper. During left searching, due to

Q and η increasing, the intersection point η_i can be located between the efficiency curves η_{i-1} and η_{i+1} by comparing the multiple efficiency curves, that is $\eta_{i-1} < \eta_i < \eta_{i+1}$ shown in Fig.3.

It is acknowledged by the similar shape of the adjacent contour lines, that, supposed there are 2 adjacent contour curves, $\eta_{i-1}(n, Q)$ and $\eta_{i+1}(n, Q)$, the contour curve $\eta_i(n, Q)$ locating between the contour lines η_{i-1} , η_{i+1} can be expressed by

$$(8) \quad \eta_i(n, Q) = W\eta_{i-1} + W'\eta_{i+1}$$

Where W and W' are weight function. The Eq. (9) is obtained by putting the Eq. (7) into Eq. (8).

$$(9) \quad (W\eta_{i-1} + W'\eta_{i+1}) \times Q = \eta_{i0} \times Q_{i0}'$$

In order to obtain the left-most point of the working conditions, Let $\frac{\partial Q}{\partial W} = 0$ and $\frac{\partial Q}{\partial W'} = 0$, then the Eqs. (10)-(11) can be obtained.

$$(10) \quad (\eta_{i-1}(n, Q) + W'\eta_{i+1}(n, Q)) \times Q = 0$$

$$(11) \quad (W\eta_{i-1}(n, Q) + \eta_{i+1}(n, Q)) \times Q = 0$$

Let $n_i = n_{i0}$, the W , W' and Q can be solved by simultaneous Eqs. (9)-(11). The running area of SCC in the working conditions is the area surrounded by the blue lines, as shown in Fig.4.

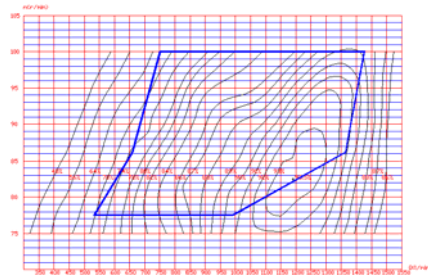


Fig. 4. the working conditions of turbine model wheel

3.The Comprehensive Evaluation Based On the Feature of SCC

As mentioned in Section 1, after the running area of the model wheel for the new power being obtained, the expert experiences are combined with the calculation to judge whether the area meets the requirements so that the best validity wheel is retrieved for the current power station. However, due to the complexity of flow movement, the generalized mathematical model can be established between speed n and flow Q and water head H , output P and efficiency η , but the objective function cannot be established.

Usually, the experiences of experts are: the higher average efficiency ($\bar{\eta}$), the lower average opening ($\bar{\sigma}$) and lower cavitation coefficient ($\bar{\sigma}$) in running area show that the wheel is of the better performance in the running area. In addition, the values of the parameters within the area should not be out of the threshold which is the certain value for the certain type of hydraulic turbine. efficiency>[E0], opening \rightarrow flying speed<[n0], Output<[P0], cavitation coefficient<[σ0].

Therefore, when all the values of the parameters are within the range of threshold, the following expression can be used to determine the performance of the model wheel for the current power.

$$(12) \quad \lambda(\bar{\eta}, \bar{o}, \bar{\sigma}) = w_1 \bar{\eta} + w_2 \frac{1}{\bar{o}} + w_3 \frac{1}{\bar{\sigma}}$$

where: $\bar{\eta}$ —mean efficiency; \bar{o} —mean opening; $\bar{\sigma}$ —mean cavitation coefficient.

Because the SCC is a set of contour curves, the mean values of the running area are the values of the volume divided by the area covered by the 3D surface within the running area. There are respectively:

$$(13) \quad \bar{\eta} = \frac{\bar{V}_\eta}{\bar{S}_\eta}; \quad \bar{o} = \frac{\bar{V}_o}{\bar{S}_o}; \quad \bar{\sigma} = \frac{\bar{V}_\sigma}{\bar{S}_\sigma}$$

In this paper, the Monte Carlo method is adopted to solve the volume enclosed by the SCC within the running area. In order to ensure the higher accuracy, the new contour curves should be created to make the SCC dense. In the last section, the expressions of the SCC are established so that the new contour curves can be created according to the Eq. 8, where $W = W' = 0.5$. Because it is easier to calculate the projected area of the SCC in running area, it is not repeated here.

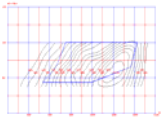
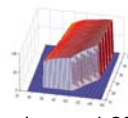
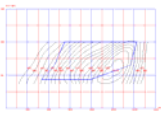
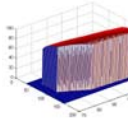
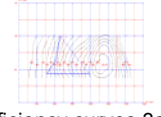
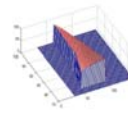


Fig. 5. Query interface



Fig. 6. Display settings

Table .2 The efficiency surface conditions and mean

 Efficiency curves 1 and corresponding running area	 volumn = 1.2334e+005 mean efficiency = 86.0339
 Efficiency curves 2 and corresponding running area	 volumn = 1.2332e+005 mean efficiency = 86.0192
 Efficiency curves 3 and corresponding running area	 volumn = 4.1210e+004 mean efficiency = 83.4607

4. The Case of the Turbine Selection Design

Here, the parameters of a power are given: maximum head $H=43.90\text{m}$, wheel diameter $D_1=6.85\text{m}$ and output $N=122\text{Mw}$. The power station parameters are input through the interface shown in Figure 5 and the display parameter

of characteristic curves is set in the interface shown in Figure 6. The system can retrieve the model wheel applicable to the current power according to the given query parameter. This example shows that the system finds 3 model wheels whose geometric features of the SCC are similar to those of the first wheel. The volume covered by the efficiency curve in running area and the mean efficiency are shown in the table 2. The first model wheel is of the highest mean efficiency based on the evaluation by the Eq. (13) in Section 4 so that it is of the better performance, which is also consistent to the results selected by the turbine design expert. In summary, the correct selection of the hydraulic turbine can be achieved by the proposed method. It takes less than one minute to retrieve the correct SCC from the library while it takes two minutes for the experts to evaluate a SCC. Therefore, the efficiency advantage increases continuously with the increased number of the SCC.

5. Conclusion

Based on the graphical features theory, the geometric similarity of the SCC of the model wheel is obtained, which supports the model retriever based on the geometric features of the curve. The method proposed by the paper greatly expands the search scope of the model and full use of the history model. However, the above method has no good solution in determining the reasonable extraction density curves. Therefore, when the 3D shapes of the curves are very similar, the method is also of large limitation, which is also the next issue to be solved.

Acknowledgment

The authors would like to thank the supports by National natural science foundation of China (No: 51075083).

REFERENCES

- [1]Liangjun Cheng, Turbine, China Machine Press , (1981)
- [2]Shenglin Ji, Guozhu Liu. Turbine. China Hydraulic and Electric Press, 1985,5
- [3]Peng Yu, Xinwei Tong and Jufu Feng. Unsupervised classification based on penalized maximum likelihood of Gaussian mixture models, Chinese Journal of Applied Probability and Statistics, 24,(2008), No. 5, 475-482
- [4]Paul D.McNicholas, Model-based classification using latent Gaussian mixture models. Journal of Statistical Planning and Inference, 140, (2010), 1175–1181
- [5]Dempster,A., Laird,N. and Rubin,D., Maximum likelihood estimation from incomplete data via the EM algorithm, J.Royal Statistical Soc.B, 39,1977, 1–38
- [6]Redner,R.A., and Walker,H.F., Mixtures densities, maximum likelihood and the EM algorithm, SIAM Review, 26, (1984), 195–239
- [7]Figueiredo, M.A.T., and Jain,A.K.. Unsupervised learning of finite mixture models, IEEE-PAMI, 24, 2002, No. 3, 381–396.

Authors

Assistant Prof. Lin LIN, Prof. Shisheng ZHONG and Doc. Xiaolong XIE are with School of Mechanical Engineering, Harbin Institute of Technology, Harbin, 150001, China. (E-mail: waiwaiyi@163.com; zhongss@hit.edu.cn; axiexiaolong@163.com)