Academic Journal of Science and Technology
ISSN: 2771-3032 | Vol. 12, No. 3, 2024

Research on Underwater Navigation Adaptation Area Classification Prediction Based on Support Vector Machine Model

Xiangyun Tan[1, *]

School of Electronic and Information Engineering, Liaoning Technical University, Liaoning, China
[*] Corresponding author

Abstract: This paper focuses on the selection of adaptation areas for gravity matching navigation in marine regions, using gravity anomaly benchmark data as the foundation. A Support Vector Machine (SVM) model is established, and the model is verified and studied for its complexity. Initially, the gravity anomaly benchmark data is cleaned and preprocessed to ensure the accuracy and consistency of the information. Subsequently, based on the principal component analysis (PCA) criterion, four gravity field characteristic parameters of the adaptability designation for each area are screened, resulting in two feature attribute indicators for judging regional adaptability: average gravity change and gravity field slope standard deviation. Through feature extraction from the refined gravity anomaly benchmark map, the classification of adaptation areas is predicted using a Support Vector Machine under non-linear separable conditions. Finally, for new gravity anomaly benchmark data, a translatability prediction is conducted on the classification system, achieving a good fitting effect. This indicates that System has significant applicability to new gravity anomaly data. It performs excellently in processing new gravity anomaly data, effectively interpreting and simulating the characteristics and patterns of these data. The system can accurately capture non-linear relationships and complex dynamic changes within the data, demonstrating good adaptability and robustness.

Keywords: Gravity anomaly, navigation adaptation area, gravity field model, principal component analysis, support vector machine.

1. Introduction

Gravity aided navigation is one of the main methods to satisfy the requirements of autonomous, passive, high concealment, unrestricted by region and time domain, and high precision navigation and positioning for underwater vehicles. In the gravity aided navigation system, the adaptive area is the navigation area that affects the reliability and accuracy of navigation, and the calibration and identification technology of the adaptive area is one of the most challenging problems.

Due to the different gravity anomaly and feature distribution in different regions, it is very important to establish a feasible adaptive region classification prediction model to ensure the navigation accuracy of underwater vehicles. Based on the reference data of gravity anomaly, this paper studies the selection of gravity matching navigation adaptation area in ocean area, and establishes a support vector machine model to solve the classification and prediction of underwater navigation and positioning adaptation area.

2. Data Processing

This paper begins by cleaning and preprocessing the gravity anomaly benchmark data to ensure the accuracy and consistency of the information, and then carries out interpolated refinement processing on the gravity anomaly values of the study area. By establishing four gravity field characteristic parameters: gravity standard deviation, gravity field information entropy, average gravity change, and gravity field slope standard deviation [1-4]. The calculation formulas for each characteristic parameter are as follows.

The standard deviation is the square root of the arithmetic mean of the squared differences between each unit value of the population and its mean. It reflects the degree of dispersion of the data sample, and the specific calculation method is as follows:

First, calculate the mean of the gravity anomaly values using the following formula.

$$\overline{g}=\frac{1}{mn}\sum_{j=1}^{m}\sum_{j=1}^{n}g\left(i,j\right) \tag{1}$$

Then, to calculate the gravity standard deviation, the formula is as follows.

$$\sigma=\sqrt{\frac{1}{m(n-1)}\sum_{j}^{m}\sum_{i}^{n}\left(g\left(i,j\right)-\overline{g}\right)^{2}} \tag{2}$$

Information entropy is used to measure the uncertainty or randomness of information. The specific calculation formula for the gravity field information entropy H is as follows:

$$H=-\sum_{j}^{m}\sum_{i}^{n}p_{ij}\log_{2}p_{ij} \tag{3}$$

Where $p_{ij}=\frac{g(i,j)}{\sum_{j}^{m}\sum_{i}^{m}g(i,j)}$ represents the probability of the occurrence of g(i, j) in the local area of the gravity field.

Average gravity change (AGD) refers to the average of the differences in gravity anomalies between adjacent grids in the gravity data, with a larger value indicating more changes in the regional gravity field. The specific calculation formula is as follows:

$$AGD=\frac{\frac{\sum_{j}^{m}\sum_{i}^{n-1}\sigma_{ij}}{m(n-1)}+\frac{\sum_{j}^{m-1}\sum_{i}^{n}\omega_{ij}}{n(m-1)}}{2} \tag{4}$$

Where $\sigma_{ij}=\left|g(i,j)-g(i,j+1)\right|$ and $\omega_{ij}=\left|g(i,j)-g(i+1,j)\right|$ represent the gravity differences between adjacent grids in the x and y directions, respectively.

The standard deviation of the gravity field slope is an indicator that statistically and descriptively measures the degree of change in the slope of the Earth's gravity field. The specific calculation formula is as follows.

$$\sigma_s=\sqrt{\frac{1}{mn-1}\sum_{j}^{m}\sum_{i}^{n}\left[S(i,j)-\overline{S}\right]} \tag{5}$$

$$\overline{S}=\frac{1}{mn}\sum_{j}^{m}\sum_{i}^{h}S(i,j) \tag{6}$$

$$S=\arctan\left(\sqrt{S_{\phi}^{2}+S_{\lambda}^{2}}\right) \tag{7}$$

$$S_{\phi}=\frac{1}{2*{}_{sc}Grid}\left(g(i+1,\ j)+g(i+1,j+1)-g(i,j)-g(i,j+1)\right) \tag{8}$$

$$S_{\lambda}=\frac{1}{2*{}_{sc}Grid}\left(g(i,\ j+1)+g(i+1,j+1)-g(i,j)-g(i+1,j)\right) \tag{9}$$

Where σ_s denotes the standard deviation of the slope, S represents the slope of the gravity field, S_{ϕ} is the slope in the latitude direction, S_{λ} is the slope in the longitude direction, ${}_{sc}Grid$ is the grid side length, and $\overline{S}$ signifies the average slope within the computational window of (1'×1').

3. Adaptation Area Classification Prediction

3.1. Feature attribute index screening

Principal Component Analysis (PCA) is a dimensionality reduction algorithm used to lower high-dimensional data into a lower-dimensional subspace while retaining important information from the data [5]. It identifies the most significant directions within the original data through linear transformation, known as principal components, and projects the data onto these components. The fundamental idea of PCA is to find the directions that explain the maximum variance in the data.

This paper, in order to facilitate the comparative analysis of the magnitude and characteristics of the values among various gravity field characteristic parameters, selects the feature attribute indices X for the adaptation area based on the principal component analysis (PCA) criterion, taking into account the impact of all gravity field characteristic parameters.

Assuming there are k gravity field characteristic parameters, with n data points for each parameter, this can form a sample matrix y of size n×k. We first perform standardization processing on it by calculating the mean values column by column.

$$Y_{ij}=\frac{y_{ij}-\overline{y_j}}{S_j} \tag{10}$$

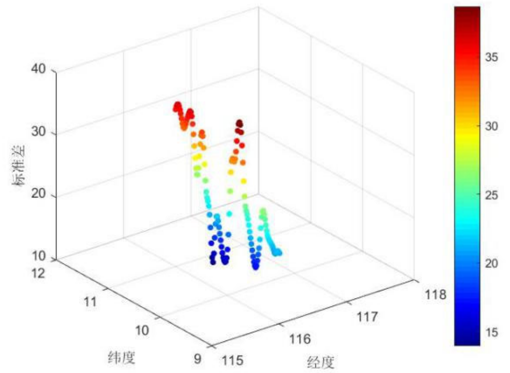
Subsequently, the covariance matrix of the standardized samples is calculated. After calculating the eigenvalues and eigenvectors of R, the contribution rate and the cumulative contribution rate of the principal components are obtained.

$$Z=\frac{\lambda_i}{\sum_{k=1}^{n}\lambda_k}(i=1,2,\cdots,k) \tag{11}$$

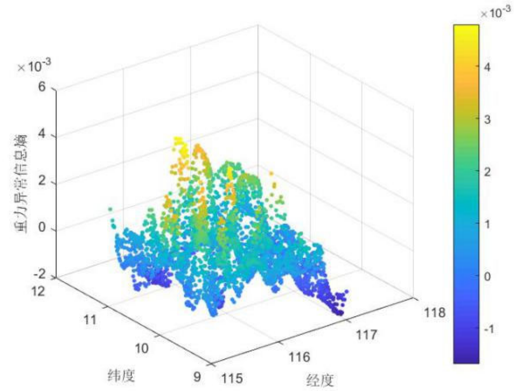
$$T=\frac{\sum_{k=1}^{i}\lambda_k}{\sum_{k=1}^{n}\lambda_k}(i=1,2,\cdots,k) \tag{12}$$

Generally, a number of principal components are selected such that the cumulative contribution rate exceeds 85%, that is as follows.

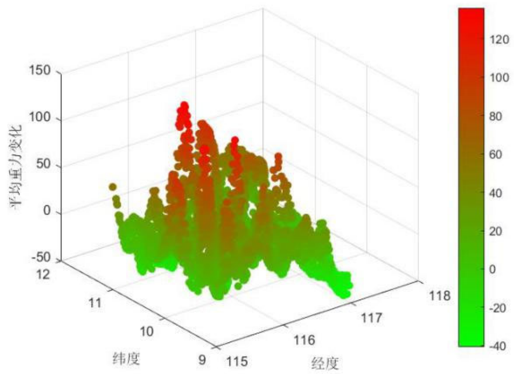
$$T=\frac{\sum_{k=1}^{p}\lambda_k}{\sum_{k=1}^{n}\lambda_k}\geq 0.85 \tag{13}$$



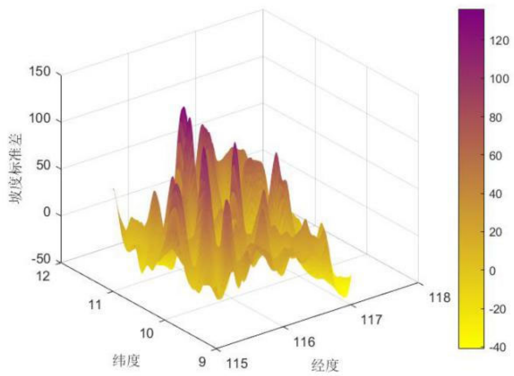
(a) Gravity standard deviation



(b) Gravity field information entropy



(c) Average gravity variation



(d) Gravity field slope standard deviation

Figure 1. Standardized three-dimensional plot of various characteristic parameters of the marine gravity field.

Fig. 1 represents the standardized three-dimensional plots of various characteristic parameters of the gravity field for a study sample size of 122m × 122m. It can be observed that it is not effective to evaluate adaptability based solely on a single gravity field characteristic parameter, as the amount of gravity information contained in a single feature parameter is limited, and the reliability of the results obtained based on this basis is not high. Therefore, it is necessary to integrate information from multiple gravity field characteristic parameters to serve as a basis for judging adaptability.

Integrating the aforementioned four gravity field characteristic parameters, calculations are performed according to the principal component analysis (PCA) criteria, using formulas (10) to (13) to determine the characteristic roots, variance contribution rates, cumulative contribution rates, and loadings, as shown in tables 1 and 2.

Table 1. Characteristic roots, variance contribution rate, and cumulative contribution rate.

Principal component	Characteristic roots	Variance contribution rate(%)	Cumulative contribution rate(%)
First principal component	1.61	40.238	40.238
Second principal component	1.243	31.066	71.303
Third principal component	0.839	20.977	92.281
Fourth principal component	0.309	7.719	100

Table 2. Loadings of each principal component.

	First principal component	Second principal component
Gravity standard deviation	0.125	0.909
Gravity field information entropy	0.443	-0.459
Average gravity variation	0.819	-0.271
Gravity field slope standard deviation	0.853	0.365

According to the table 1, the cumulative probability of the first two principal components is less than 85%, while the cumulative contribution rate of the last two principal components reached 92.281%, which exceeds 85%. This means that the latter two principal components contain 92.281% of the effective information from the original characteristic parameters. Therefore, selecting the last two principal components is sufficient for a comprehensive assessment of the adaptability of the study area, i.e., the characteristic attribute indicators for marine gravity matching navigation adaptability zone are average gravity variation and gravity field slope standard deviation.

The table 2 shows the magnitude of loadings for each characteristic parameter of the gravity field in the first two principal components from table 1. The results indicate that the first principal component has a large absolute value of loadings for average gravity variation and gravity field slope standard deviation. Both the second and third principal components have large absolute values of loadings for gravity standard deviation and gravity field slope standard deviation.

3.2. Establishment of SVM model

Support Vector Machine (SVM) is a machine learning method based on statistical theory [6], and its fundamental idea is to find an optimal classification surface that separates two types of samples, while maximizing the distance between sample points and the classification surface.

For the training set samples: (x_1, y_1), (x_2, y_2), $\cdots$, (x_n, y_n), $x \in R^n, y \in (-1,1)$ the quadratic programming for finding the optimal classification surface is as follows:

$$\min \frac{1}{2}\omega^{T*}\omega + c\sum_{i=1}^{N} x_i \tag{14}$$

$$s.t. y_i\left(\omega^T x_i + b\right) \geq 1 - \xi_i, \xi_i \geq 0, i = 1, 2, \cdots, N \tag{15}$$

The Lagrange function is shown as follows.

$$L(\omega, b, l) = \frac{1}{2}\omega^{T*}\omega + c\sum_{i=1}^{N} z_i - \sum_{i=1}^{N} l_i\left[y_i\left(\omega^T x_i + b\right) - 1 + x_i\right] - \sum_{i=1}^{N} a_i x_i \tag{16}$$

Where $\lambda_i \geq o, i = 1, 2, \cdots, N$ and the dual problem of equation $\min \frac{1}{2}\omega^{T*}\omega + c\sum_{i=1}^{N} x_i$ can be represented as follows.

$$\max Q(l) = \sum_{i=1}^{N} l_i - \frac{1}{2}\sum_{i,j=1}^{N} y_i y_j \lambda_i \lambda_j \left(x_i, x_j\right) \tag{17}$$

$$s.t. \sum_{i=1}^{N} y_i \lambda_i = 0, 0 \leq \lambda_i \leq C \tag{18}$$

The decision function can be derived as follows.

$$f(x, l) = \operatorname{sgn}\left(\sum_{sv} y_i \lambda_i \left(x_i, x_j\right) + b\right) \tag{19}$$

The dual function for linearly non-separable cases is as follows.

$$\max Q(\lambda) = \sum_{i=1}^{N} \lambda_i - \frac{1}{2}\sum_{i,j=1}^{N} y_i y_j \lambda_i \lambda_j \left(\varphi_{xi}{}^{*}\varphi_{xj}\right) \tag{20}$$

In the equation, $\varphi_{xi}{}^{*}\varphi_{xj}$ represents the kernel function, and the decision function is as follows.

$$f(x, l) = \operatorname{sgn}\left(\sum_{sv} y_i \lambda_i K\left(x_i, x_j\right) + b\right) \tag{21}$$

In summary, the process involves using a pre-selected non-linear mapping to transform the input vectors into a high-dimensional feature space, where the optimal classification hyperplane is constructed.

The adaptability zone selection problem based on Support Vector Machine (SVM) is essentially a binary classification problem using SVM. In this paper, the gravity anomaly reference map is processed in blocks, and based on the two characteristic attribute indicators average gravity variation and gravity field slope standard deviation selected through Principal Component Analysis (PCA), each small block on the reference map is classified. Blocks that meet the criteria are designated as adaptability zones, while those that do not are considered non-adaptability zones.

Firstly, the feature data of the gravity anomaly reference map is processed. To avoid the adverse effects of high-dimensional vector data obtained by directly operating on the original reference map data, and to ensure the speed of training calculations and the accuracy of classification when using SVM for high-dimensional vector data classification, this paper directly employs the Radial Basis Function (RBF) kernel, which generally achieves good classification results under normal circumstances.

3.3. Adaptability zone selection and analysis

(1) Selection and division of gravity anomaly reference map

Firstly, a gravity anomaly reference map is selected, in this case, an oceanic gravity anomaly reference map with a grid size of (1' × 1') is chosen, as shown in Fig. 2. According to the first issue, this gravity anomaly reference map is divided into blocks, determining the window size. Here, an area of 4×4 is chosen as the size for each block, resulting in 16 regional sample blocks with a total of 14,884 training samples.

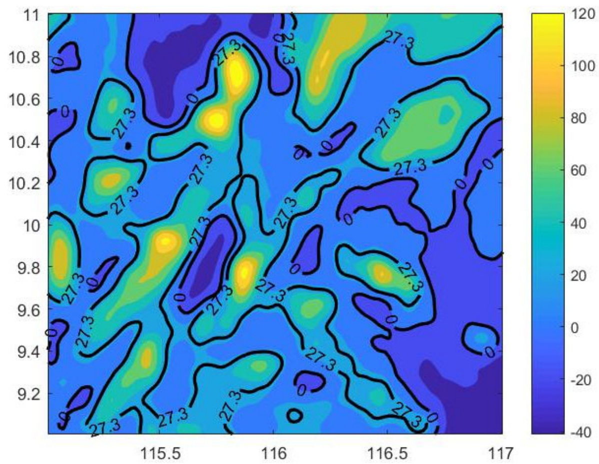


Figure 2. Marine regional gravity anomaly reference map.

Feature extraction of the gravity anomaly reference map.

Firstly, each divided area is numbered from left to right and from top to bottom, marking 0 to 16 sample blocks. Then, using the feature extraction method from Question One, each area is subjected to feature extraction. Based on Principal Component Analysis (PCA), features are screened to obtain the feature attribute indicators of the gravity anomaly reference map as the average gravity variation and the gravity field slope standard deviation. Partial data obtained is shown in the table 3.

Table 3. Partial feature data obtained from the gravity anomaly reference map.

	Average gravity variation.	Gravity field slope standard deviation.
1	20.0776	19.1078
2	20.8999	27.4852
3	20.5292	33.1969
4	20.9012	26.1435
5	20.4753	29.9397
6	20.8983	33.1245
7	21.3161	35.3946
8	20.6575	16.8395
9	20.6094	35.8979
10	21.2786	36.1356
...	...	...

Data normalization processing

Before training and classifying predictions with Support Vector Machines (SVM) on this data, it is necessary to normalize the data. The purpose is to project the value range of a feature's attribute into a specific range to avoid the impact of numerical attributes with different size ranges on the fairness of results from distance-based classification methods. The normalization mapping formula is shown as follows.

$$V' = \frac{v - \min_A}{\max_A - \min_A}\left(\max'_A - \min'_A\right) + \min'_A \qquad (22)$$

Here, $\max_A$ and $\min_A$ represent the maximum and minimum values of attribute A, respectively, while $\max'_A$ and $\min'_A$ are the corresponding maximum and minimum values in the new space after transformation.

Creating the training sample set

To select the adaptability zones for the gravity anomaly reference map, it is first necessary to obtain a standard gravity anomaly reference map that has already defined the adaptability and non-adaptability zones. Therefore, the previously selected gravity anomaly reference map needs to be delineated with adaptability zones. Based on the feature data statistics from the previous section, the criteria for dividing the adaptability zones of the gravity anomaly reference map are as follows.

$$AGD > 20 \cap \sigma_s > 28 \qquad (23)$$

Where AGD represents the average gravity variation and σ_s represents the gravity field slope standard deviation, both of which are values after normalization treatment. As long as the above conditions are met, it is marked as an adaptability zone; otherwise, it is a non-adaptability zone. Using this rule, the gravity anomaly reference map can be divided, and the required training sample set can be obtained.

SVM training and classification testing

The 14,884 samples are divided into two categories based on the adaptability zone selection criteria: those that meet the adaptability zone conditions are labeled as +1, and the rest are labeled as -1. The samples are then randomly divided into two parts, with 7,442 samples in each part. One part is used as the training samples, and the other as the test samples. First, the SVM is trained with the training samples. Once a good training effect is achieved, the well-trained SVM is then used to classify and predict the test samples.

Adaptability zone classification prediction effect and analysis

The gravity anomaly reference map of the test samples is classified and predicted for the adaptability zones using the well-trained Support Vector Machine (SVM). Then, the classification matching effect is tested within the regional adaptability zones. The prediction effect is shown in Fig. 3.

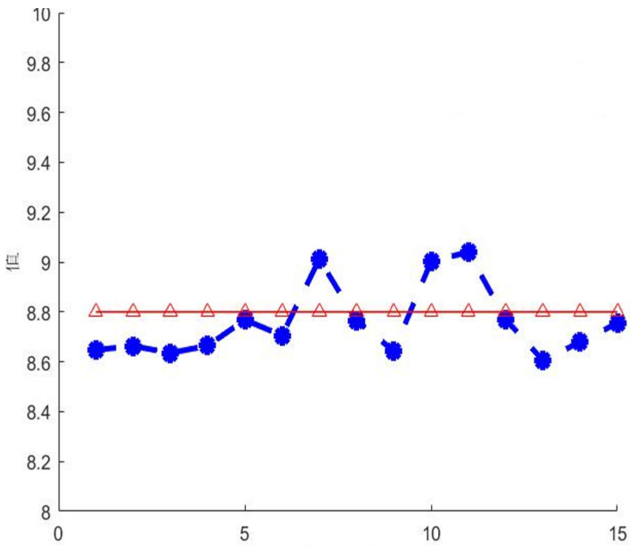


Figure 3. Adaptability zone matching prediction results

From Fig. 3, it can be seen that the adaptability zone classification prediction for the test samples' gravity anomaly reference map using the well-trained Support Vector Machine (SVM) fits the data very well. This indicates that the SVM, based on the training sample set and with the Radial Basis Function (RBF) kernel, can accurately capture the nonlinear relationships and complex dynamics within the data, demonstrating good adaptability and robustness.

3.4. Validation of new gravity anomaly data

In order to ensure the accuracy and consistency of the data, it is necessary to clean and process the gravity anomaly reference data first. The 14,762 samples of data were divided into two categories according to the selection criteria of adaptive area and non-adaptive area. Those that met the conditions of adaptive area were marked as +1, and those that did not were marked as -1. Then the samples were randomly divided into two pieces with 7,381 samples each, one as training sample and the other as test sample. When a good training effect is achieved, the trained support vector machine is used to classify and predict the test samples. The trained support vector machine is used to predict the gravity anomaly reference map of the test sample, and then the classification matching effect is tested in the regional adaptation area. The predicted results are shown in Fig. 4.

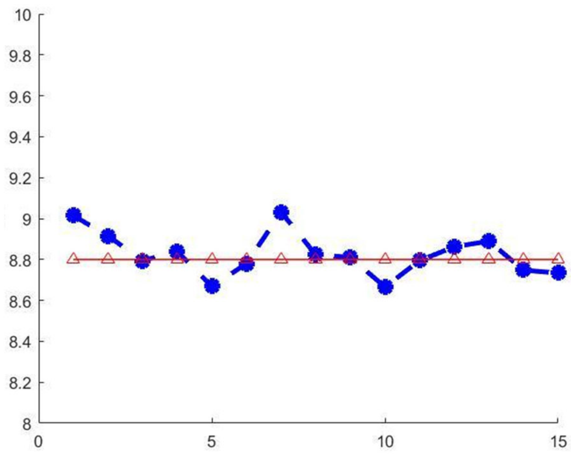


Figure 4. Adaptability zone matching prediction results

As can be seen from Fig. 4, the migration prediction of system established in question 2 with reference data of gravity anomalies has obtained a good fitting effect, indicating that system has significant applicability to new gravity anomaly data, and system has performed well in processing new gravity anomaly data and can effectively interpret and simulate the characteristics and rules of these data. It can accurately capture nonlinear relationships and complex dynamic changes in data, showing good adaptability and robustness.

4. Conclusion

This paper focuses on the prediction of gravity region adaptation areas for underwater navigation technology based on gravity assistance, utilizing the classification method of Support Vector Machines (SVM). SVM, by introducing kernel functions, can accurately capture non-linear relationships and complex dynamic changes in the data, demonstrating good adaptability and robustness. In future research, SVM can be combined with other models and techniques, such as ensemble learning methods (e.g., random forests, gradient boosting trees), feature selection methods, and feature engineering, to further enhance the performance and generalization capabilities of the model.

References

[1] Li, S. S. "Research on the theory and method of underwater gravity-aided inertial navigation." Information Engineering University (2010).

[2] Yuwei, Zhu. "Rout Planning Research Based on Gravity Information." Master, Bei**g Institute of Technology (2016).

[3] Wu, Lin, et al. "Performance evaluation and analysis for gravity matching aided navigation." Sensors 17.4 (2017): 769.

[4] Melo, José, and Aníbal Matos. "Survey on advances on terrain based navigation for autonomous underwater vehicles." Ocean Engineering 139 (2017): 250-264.

[5] LI, ZhaoWei, et al. "Optimizing suitability area of underwater gravity matching navigation based on a new principal component weighted average normalization method." Chinese Journal of Geophysics 62.9 (2019): 3269-3278.

[6] Kumar, Munish, R. K. Sharma, and M. K. **dal. "Efficient feature extraction techniques for offline handwritten Gurmukhi character recognition." National Academy Science Letters 37 (2014): 381-391.