

## Research Article

# Singing and Nervous System Regulation Based on Wireless Sensor Network Perception

Feng Yu 

*School of Music, Guizhou Normal University, Guiyang 550001, China*

Correspondence should be addressed to Feng Yu; 460023011@gznu.edu.cn

Received 12 August 2021; Revised 15 September 2021; Accepted 7 October 2021; Published 25 October 2021

Academic Editor: Mu Zhou

Copyright © 2021 Feng Yu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to build an intelligent platform that can be applied to singing and nervous system adjustment, this paper optimizes the positioning and information processing algorithms for wireless sensor network perception. Moreover, this article combines binocular vision to realize the singer's real-time positioning, combines the singer's emotion recognition with the intelligent sensor system, and combines the emotion recognition with the adjustment of the nervous system, so that the singer can better control the intelligent platform. In addition, in order to solve the problem of multisensor information fusion, this paper improves the sensor fusion algorithm to make it suitable for the information fusion of vision sensors and information sensors. Finally, this paper designs the functional structure of the system, transmits data through wireless sensor networks, simulates human emotion models, and studies the process of singing and the adjustment of the nervous system. It can be seen from the experimental research results that the method proposed in this paper has a certain effect.

## 1. Introduction

The process of singing is not only a vocal process, but also an emotional transmission process, and people's emotions are mainly controlled by the nervous system. Therefore, in order to improve the effect of singing, it is necessary to improve the regulation of singing and nervous system while singing, so as to improve the effective regulation and control between singing and emotion [1].

In vocal music learning and singing, only by correctly playing the musical instrument "voice" can singers make a correct sound. However, due to the special attribute of voice, singers cannot intuitively operate the "voice" instrument. They can only experience their own singing movement state through various internal feelings and adjust their singing behavior activities in time, so as to make a good and high-quality voice [2]. This makes singing complex and difficult to control. When a singer sings, on the one hand, the brain and nervous system direct the coordinated movement of the singer's physiological organs; on the other hand, they also dispatch internal psychological activities to make a correct response, so as to interact with the singer's physiological

body and jointly assist the voice. Singing behavior activity is the combination of singer's body physiology and psychology. This process not only makes each organ of singing physiology in a positive movement state, but also correctly guides singer's psychology, which makes singer's body physiology in an effective movement and cooperates with singing psychology to carry out artistic practice. It can be seen that in vocal music learning and singing, only by effectively grasping the combination of singer's body physiology and psychology can we effectively carry out vocal music artistic activities. Furthermore, we can reproduce the connotation and artistic plot of vocal music works and send out the voice that moves the audience to make the internal psychology of the audience consistent with the emotional expression of the song, so as to infect the audience and make them affirm the singer's artistic behavior [3].

The singer's body physiology is not only the basis of singing, but also the basis of the emergence and development of singing psychological activities. In the study and singing of vocal music, the comprehensive training of singer's physiological function is also the basis of singer's vocal music practice. The training of singers' physiological

skills is the basis of singing. In the early stage of vocal music learning, because the physiological organs of singing body have not been systematically and scientifically trained, there will be some problems in singing, such as broken sound, hoarseness, tension of laryngeal muscle tissue, shallow breath, and inability to reasonably open and apply the resonance chamber. These problems are the result that the singer's psychology cannot reasonably regulate the coordinated movement of the body's physiological organs and muscle tissues without training.

Based on the above analysis, this paper applies wireless sensor network to the regulation of singing and nervous system, transmits data through wireless sensor network, and simulates human emotion model, so as to study the regulation process of singing and nervous system.

## 2. Related Work

The purpose of WSN is to perceive, collect, and transmit data information related to monitoring objects in the monitoring area, process the information, and finally provide the processed data to users [4]. A typical WSN network system consists of sensor node, sink node, transportation network, and monitoring center. A large number of sensor nodes are randomly deployed in or near the monitoring area, which can form a network through self-organization. The data monitored by the sensor node is transmitted through the transmission network. In the transmission process, the data information may be processed by multiple nodes and routed to the sink node after multiple hops [5].

Data aggregation is an important research problem in wireless sensor networks [6]. For the study of data aggregation, either the algorithm based on energy balance consideration or the algorithm in the case of multiple base stations or the algorithm in the case of single base station or the real-time algorithm in the routing layer, literature [7] proves that the data aggregation problem of reducing delay is NP difficult. They also designed an algorithm for data aggregation with an approximate coefficient of  $\Delta-1$  ( $\Delta$  is the degree of the largest node). Literature [8] studies the optimal forwarding moment problem as a policy node for the decision process model to determine the optimal decision of the sensor. In [9], an algorithm is designed which has a delayed binding  $O(\log(N))$ . It is believed that each node can learn the nearest neighbors and has a special collision detection function, although this situation cannot always guarantee that the network is normal. The data aggregation model studied in literature [10] is based on a tree structure. Literature [11] proposed a scheduling algorithm based on the maximum independent set, which has a delayed binding of  $23R + \Delta - 18$ . They focused their main energy on a special scene.

Literature [12] proposed the Ken algorithm, which uses two dynamic probability models; one runs on the sink node (sink node), and the other runs on each node in the network, passing these two models under the premise of a small loss of precision, a model to reduce the amount of data transmission as much as possible. Literature [13] proposed the ALVQ algorithm, which uses historical data to construct a cipher to explore the inherent characteristics of the data

and then uses this cipher to piecewise linearly compress other data to reduce the amount of data transmission. Literature [14] proposed a piecewise linear approximation algorithm, which uses a straight line to approximate the current but uncompressed data point under the condition of a given error limit, until a new data point makes this limit be broken. Then similarly, starting from this new data point, use a new straight line to approximate the subsequent arrival point. Literature [15] proposed the EAQ algorithm, which first converts the original time series into a special time series description MVA (multiversion array). Using this MVA prefix, an approximate version of the original time series with a certain error can be recovered. As the prefix increases, the error gradually decreases. The time series approximation with variable error is realized. In addition, researchers have also proposed algorithms for data compression using Discrete Fourier Transform [16], Discrete Cosine Transform [17], and Discrete Wavelet Transform [18], which are all based on time series time correlation algorithms. But its computational complexity is high, so it is not suitable for WSN.

## 3. General Framework for Multisensor Fusion Positioning of Singing Voice Based on Graph Optimization

Graph optimization is a method for processing optimization problems. The variables and constraint equations between variables that are specifically solved in the optimization problem are expressed as nodes and edges in an undirected graph, and the variables to be optimized are calculated iteratively through methods such as gradient descent. The positioning algorithm framework proposed in this section uses the idea of graph optimization to convert the amount of motion and sensor observations on a certain segment of the singer's trajectory into constraints and variables in the graph optimization and iteratively calculate the optimal singer's pose.

The singer's motion model and viewing model are usually modeled as nonlinear equations with Gaussian noise [19]:

$$\begin{aligned} x_i &= g_i(x_{i-1}, u_i) + w_i, w_i \sim N(0, \Delta_{w_i}^{-1}), \\ z_i &= h_i(x_i) + v_i, v_i \sim N(0, \Delta_{v_i}^{-1}). \end{aligned} \quad (1)$$

$x_i$  is the system state vector to be calculated at the main time,  $z_i$  is the observation vector of the sensor at time  $i$ , and  $z_i$  is the system motion equation that maps the system state vector from time  $i$  to time  $i+1$ .  $u_i$  is the system control variable input at time  $i$ , and  $h_i$  is the system observation equation that maps the state of the system to the observation space at time  $i$ .  $w_i$  and  $v_i$  are the motion noise and observation noise of the system, which are Gaussian white noise with variances  $\Delta_{w_i}^{-1}$  and  $\Delta_{v_i}^{-1}$ , respectively, and  $\Delta_{w_i}$  and  $\Delta_{v_i}$  are the signal matrix of motion noise and the information matrix of observation noise, respectively. This state vector can be calculated by solving the following least squares optimization problem:

$$\begin{aligned}
x^* &= \min_x \sum_l \left( \|x_l - g_l(x_{l-1}, u_l)\|_{\Delta_{w_l}}^2 + \|z_l - h_l(x)\|_{\Delta_{w_l}}^2 \right) \\
&= \min_x \sum_l \left( \|r_{g_l}(x)\|_{\Delta_{w_l}}^2 + \|r_{h_l}(x)\|_{\Delta_{w_l}}^2 \right) \\
&= \min_x \sum_l \|r_l(x)\|_{\Delta_{w_l}}^2.
\end{aligned} \tag{2}$$

$r_{g_l}$  and  $r_{h_l}$  are the residual expressions of the motion equation and the observation equation, respectively, which can usually be approximated by the Gauss-Newton method. Since it is a nonlinear equation and needs to be linearized by Taylor expansion, the expression of the update amount for each iteration should be

$$\delta x = \min_{\delta x} \sum_l \|r_l(\hat{x} \boxplus \delta x)\|_{\Delta_{w_l}}^2 \sim \min_{\delta x} \sum_l \|r_l(\hat{x}) + J_l \delta x\|_{\Delta_{w_l}}^2. \tag{3}$$

Among them,  $J_l := \partial r_l(\hat{x} \boxplus \delta x) / \partial \delta x|_{\delta x=0}$  is the covariance matrix of the error state.  $\boxplus$  is the lie algebra operators. By expanding the above formula, the expression of  $\delta x$  is obtained [20].

$$\delta x = -(J^T \wedge J)^{-1} J^T \wedge r. \tag{4}$$

By superimposing this amount of change on the state amount at the previous moment, the updated state amount at the next moment can be obtained.

$$x = \hat{x} \boxplus \delta x. \tag{5}$$

In the positioning problem, the state equation of the above optimization problem may include pose, geosphere point, and parameters. The pose is usually represented by SE3 manifold, and the map point is represented by its Euclidean coordinates. The motion model is related to the singer's motion sensor, and the observation model is related to the singer's sensory sensor. The motion model is to constrain the states of two singers at different times, and the observation model is to constrain the states of the singers and the map points. Therefore, the above optimization model can be expressed as an undirected graph shown in Figure 1.

The general positioning framework is used to describe specific positioning problems, which will greatly help the subsequent system observability analysis and system matrix decomposition. The reason is that the undirected graph here is often similar to the sparse matrix of the system in terms of spatial structure, which facilitates the analysis of the system. At the same time, graph representation is also helpful to the realization of algorithms, because many algorithms in the computer field are implemented based on graph representation.

Aiming at the binocular camera and IMU, combined with the feature-based visual positioning algorithm, a visual positioning system based on binocular-IMU fusion is designed and implemented. The visual mileage estimation is performed by extracting feature points from the binocular image, combined with IMU preintegration. Carry out tight

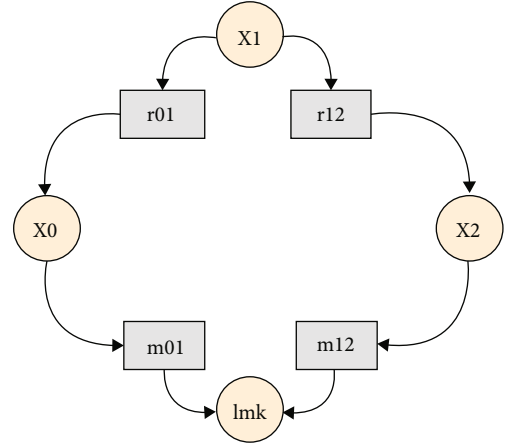


FIGURE 1: Schematic diagram of the general framework of multisensor fusion positioning.

coupling optimization to achieve high-precision and high-stability visual positioning. This method uses the high frequency and high dynamics of the IMU to improve the accuracy and robustness of the feature-based visual positioning algorithm and to ensure the computational efficiency. Experiments show that even in high-speed motion and undertextured scenes, the method can still estimate the singer's motion stably and accurately.

The structural block diagram of the system is shown as in Figure 2. First, the visual inertial registration algorithm is introduced, and then, the IMU preintegration and binocular visual information are tightly coupled and optimized using the fusion algorithm framework based on graph optimization. Finally, the performance of the algorithm is verified through experiments.

With IMU preintegration, the relative constraints between poses can be obtained for a certain period of time. However, the IMU preintegration is calculated in the IMU local coordinate system and needs to be fused with the result of the calculation in the visual coordinate system. In the visual positioning algorithm, the world coordinate system usually selects the first frame camera coordinate system, and the acceleration of gravity measured by the IMU is in the IMU coordinate system. For the multisensor fusion algorithm, only one coordinate system can be selected as the navigation coordinate system. In this paper, the camera coordinate system is selected as the navigation coordinate system, and the world coordinate system is the camera coordinate system of the first frame. Under this assumption, within a period of time from the beginning of the binocular vision and IMU fusion positioning algorithm proposed in this paper, it is necessary to initialize the visual positioning algorithm to obtain the camera trajectory and the constructed environment map during this period, and corresponding. The IMU preintegration is calculated between adjacent key frames to realign the gravity in the IMU coordinate system to the visual coordinate system. This process is called visual inertial registration.

Unlike a monocular camera, a binocular camera can obtain the left and right images at the same time every time

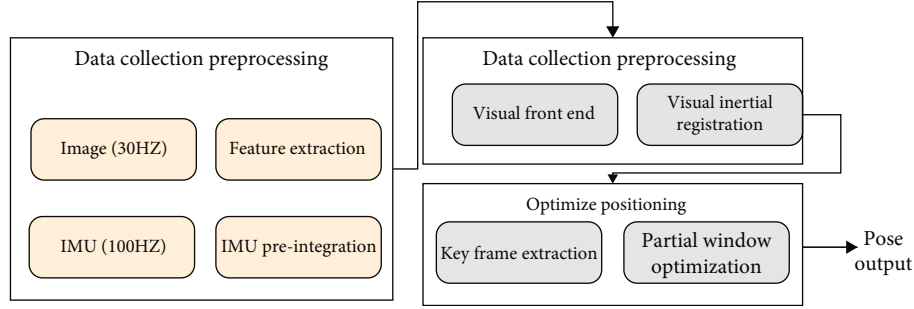


FIGURE 2: IMU binocular vision fusion algorithm frame.

it samples and can directly restore the feature point depth through the binocular matching algorithm. After the first frame of the binocular image is acquired, the initial visual feature map can be constructed through binocular matching, and through this map and the subsequent binocular images, the PnP algorithm is performed to calculate the pose tracking result. The specific steps are shown in Figure 3.

**3.1. Feature Extraction.** On binocular images, ORB feature points are extracted to ensure that 1000-1500 feature points can be extracted from each image. In order to make the extracted feature points evenly distributed on the image, this paper divides the image into grids of the same size. The number of feature points extracted in each grid is roughly the same to ensure that enough feature points can be extracted in each image area.

**3.2. Three-Dimensional Reconstruction.** The binocular vision method can reconstruct the visual feature map from the left and right camera images at the same time. The most commonly used is the module matching algorithm, referred to as the BM algorithm. The binocular image needs to be pre-processed by the binocular alignment algorithm. After this step, the same pixel blocks in the left and right images are the same in height. Then, when performing the BM algorithm, searching for a pixel in the left image to the corresponding pixel block in the right image can simply search on a pixel of the same height, which greatly improves the search efficiency. However, because the binocular alignment algorithm cannot perfectly align the left and right images, when searching, we can not only search for a certain pixel in the right image, but should also search for a certain pixel and its surrounding area to improve the accuracy of the search. After the correct matching point pair is obtained by the BM algorithm, the depth recovery of the feature points can be carried out through the binocular stereo. A schematic diagram of binocular depth recovery is shown in Figure 4. The calculated depth is the Z-direction coordinate of point X, which can be solved by solving the triangle constructed in the figure.

**3.3. Pose Solution.** Through binocular depth recovery, a feature point map can be reconstructed. Some points in this map can be observed by the current image and associated through feature matching. The method of calculating the pose of the current image through the association between

the map points in the three-dimensional space and the pixels on the image is called the PnP algorithm (Perspective n Point). As shown in Figure 5, the map point obtained by binocular recovery at time  $k-1$  is a small blue circle in the figure. Some of these map points are observed by the image at time  $k$ , that is, the map points connected to the image at time  $k$  with a purple line. These map points are projected back to the image at time  $k$  and are associated with the feature points in the image through feature matching. In order to calculate  $T_k^{k-1}$ , we set the projection matrix at time  $k$  to  $P_k$  [21].

$$P_k = K \begin{bmatrix} R_k^{k-1} & P_k^{k-1} \\ 0 & 1 \end{bmatrix}. \quad (6)$$

$K$  is the camera internal parameter matrix. The coordinates of a certain map point in space in the  $k-1$  camera coordinate system are denoted as  $X$ ; then, the pixel coordinates  $x$  of  $X$  in the camera at time  $k$  can be calculated by  $P_k$  as follows:

$$x = P_k X. \quad (7)$$

The corresponding  $T_k^{k-1}$  can be solved by solving the above projection equation. There are a total of 6 degrees of freedom for  $T_k^{k-1}$ , and each projection point can provide two linear equations, that is to say theoretically, there are 3 projection point pairs to calculate the relative motion  $T_k^{k-1}$ . However, due to factors such as mismatches in the matching process of map points and feature points, we cannot correctly calculate the accurate relative motion  $T_k^{k-1}$  by selecting only three projection point pairs in practice, but a large amount of sampling is required. Therefore, this paper uses an optimized method to calculate a result that is closest to the true value.

In addition to the above three modules, key frames are also selected in the visual front end. This paper proposes the following two criteria for selecting key frames. The first is to calculate the average disparity between the current frame and the previous frame. If the average disparity exceeds a certain threshold, it is considered that the current frame should be selected as a key frame. The disparity is obtained by calculating the coordinate pixel difference between the matching feature points of the current frame

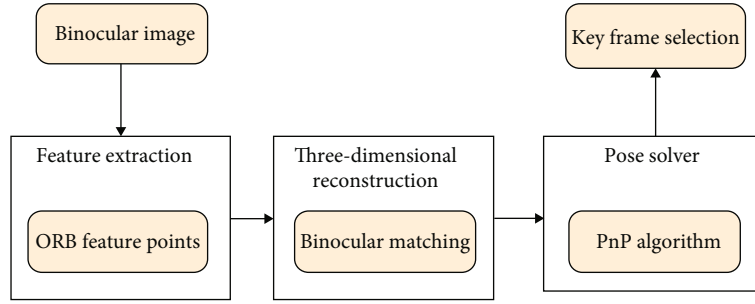


FIGURE 3: Flow chart of binocular vision front end.

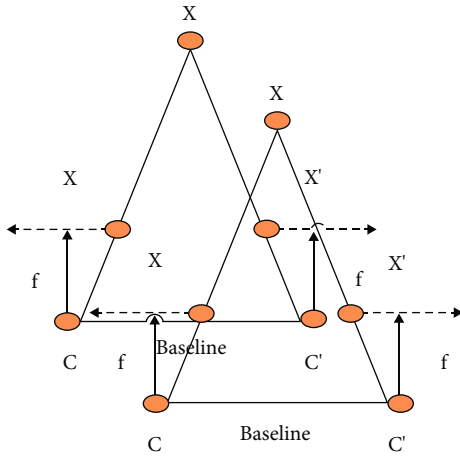


FIGURE 4: Schematic diagram of binocular depth recovery.

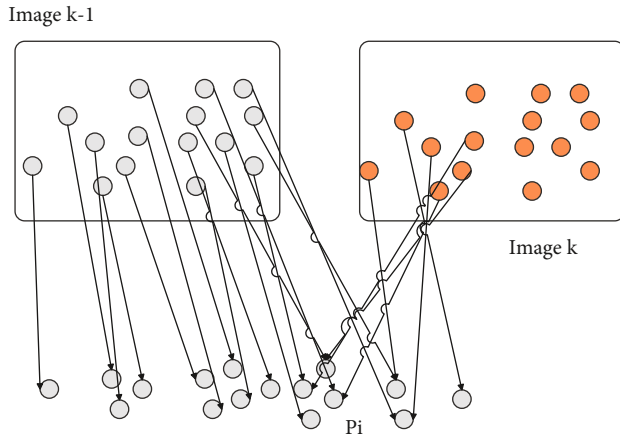


FIGURE 5: Schematic diagram of PnP algorithm.

and the previous frame. It is worth noting that not only translation will produce parallax, but also rotation. However, it is difficult to reconstruct visual features under pure rotation. In order to avoid this situation, this paper integrates the gyroscope signal in a short period of time and separates it during the calculation of the parallax, so as to ignore the influence of rotation on the visual reconstruction. The second is to consider the quality of pose tracking. If the number of points obtained by matching the current frame with the previous frame is less than a certain threshold, the current frame will be selected as a key frame. This criterion can greatly reduce the probability of pose tracking failure.

The relative pose between the key frames is obtained through IMU preintegration, and the relative pose of the key frames is obtained through the visual front end. Because the coordinate systems are different, the registration needs to be performed in a loose coupling manner, which is mainly to convert the gravity in the IMU coordinate system to the visual coordinate system. Figure 6 is a schematic diagram of a visual inertial registration.

Through the visual front end and IMU preintegration, it is possible to independently calculate the movement trajectory of the camera and the movement of the IMU coordinate system within a period of time. Since the rotation obtained by the preintegration of the gyroscope has nothing to do with gravity, we can first preintegrate the gyroscope to obtain a measurement with offset and compare it with the rotation obtained by the camera motion estimation to estimate the gyroscope offset. The rotations of two adjacent key frames estimated by vision are  $R_{C_k}^{C_0}$  and  $R_{C_{k+1}}^{C_0}$ , and the rotations of the IMU coordinate system can be obtained by the external parameters of the camera and IMU as  $R_{B_k}^{C_0}$  and  $R_{B_{k+1}}^{C_0}$ . Compared with the result of IMU preintegration, there is a relation [22]:

$$R_{B_k}^{C_0 T} R_{B_{k+1}}^{C_0} \approx \Delta R_{k,k+1}. \quad (8)$$

The gyroscope bias  $b_g$  is included in  $\Delta R_{k,k+1}$ , so the optimization objective equation can be obtained.

$$\min_{\delta b_g} \sum \left\| \log \left( R_{B_{k+1}}^{C_0 T} R_{B_{k+1}}^{C_0} \Delta R_{k,k+1} \right) \right\|^2, \quad (9)$$

$$\Delta R_{k,k+1} \approx \tilde{\Delta R}_{k,k+1} \exp(\delta g).$$

By iteratively optimizing this objective equation, the gyroscope bias  $b_g$  in the registration phase can be solved.

The gravitational acceleration is coupled in the acceleration measurement of the IMU. In order to register the gravitational acceleration with the visual coordinate system, it is considered to associate the translation component estimated by the visual motion with the velocity and position measurement in the IMU preintegration. The relationship between the position of the visual coordinate system at time  $k$  and

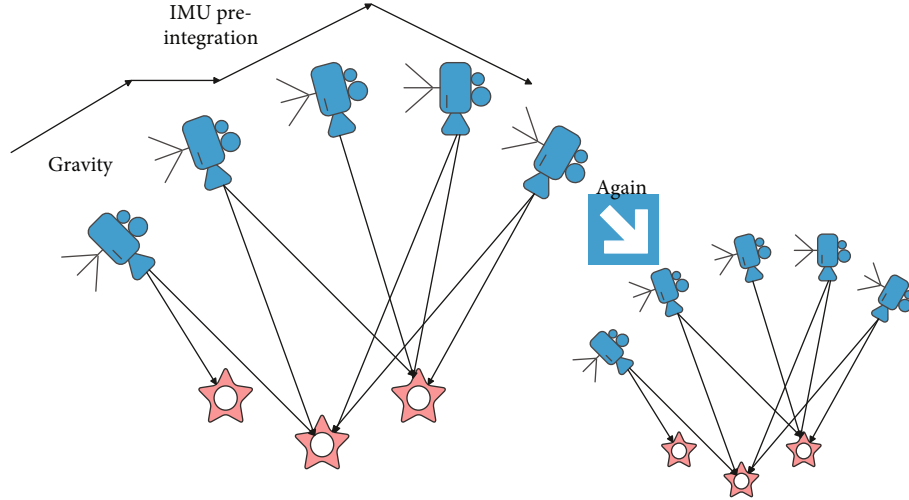


FIGURE 6: Schematic diagram of visual inertial registration.

the position of the IMU coordinate system is as follows:

$$p_{B_k}^{C_0} \approx p_{C_k}^{C_0} - R_{B_k}^{C_0} p_C^B. \quad (10)$$

In order to convert the gravitational acceleration to the visual coordinate system  $C_0$ , the IMU preintegration from time  $k$  to  $k+1$  is rewritten as follows:

$$\begin{aligned} \Delta p_{k,k+1} &= R_{B_k}^{C_0 T} \left( p_{C_{k+1}}^{C_0} - p_{C_k}^{C_0} + \frac{1}{2} g \Delta t^2 - R_{B_k}^{C_0} v_{B_k} \Delta t \right), \\ \Delta v_{k,k+1} &= R_{B_k}^{C_0 T} \left( R_{B_{k+1}}^{C_0} v_{B_{k+1}} + C_0 g \Delta t - R_{B_k}^{C_0} v_{B_k} \right). \end{aligned} \quad (11)$$

Visual motion estimation and IMU preintegration are integrated to obtain the linear observation equation:

$$z_{B_{k+1}}^{B_k} = \begin{bmatrix} \Delta p_{k,k+1} - p_C^B + R_{B_k}^{C_0 T} R_{B_{k+1}}^{C_0} p_C^B \\ \Delta v_{k,k+1} \end{bmatrix} = H_{B_{k+1}}^{B_k} \chi + n_{B_{k+1}}^{B_k}. \quad (12)$$

Among them,  $H_{B_{k+1}}^{B_k} = \begin{bmatrix} -I \Delta t & 0 & 1/2 R_{B_k}^{C_0 T} \Delta t^2 \\ -I \end{bmatrix}$  is the

variable to be optimized, including  $H_{B_{k+1}}^{B_k} =$

$$\begin{bmatrix} -I \Delta t & 0 & 1/2 R_{B_k}^{C_0 T} \Delta t^2 \\ -I \end{bmatrix}, v_{B_k}, \text{ and } v_{B_{k+1}}, \text{ and } H_{B_{k+1}}^{B_k} = \begin{bmatrix} -I \Delta t & 0 & 1/2 R_{B_k}^{C_0 T} \Delta t^2 \\ -I \end{bmatrix} \text{ is as follows:}$$

$$H_{B_{k+1}}^{B_k} = \begin{bmatrix} -I \Delta t & 0 & \frac{1}{2} R_{B_k}^{C_0 T} \Delta t^2 \\ -I & R_{B_k}^{C_0 T} R_{B_{k+1}}^{C_0} & R_{B_k}^{C_0 T} \Delta t \end{bmatrix}. \quad (13)$$

Just by solving the linear equations,  $H_{B_{k+1}}^{B_k} =$

$\begin{bmatrix} -I \Delta t & 0 & 1/2 R_{B_k}^{C_0 T} \Delta t^2 \\ -I \end{bmatrix}$  can be calculated and obtained.

$$\min_{\chi} \sum \left\| Z \wedge_{B_{k+1}}^{B_k} - H_{B_{k+1}}^{B_k} \chi \right\|^2. \quad (14)$$

In the multisensor fusion problem, there are usually two fusion methods, namely, loose coupling and tight coupling. Loose coupling means that each sensor in a multisensor system uses its own information to perform calculations and then fuse the results calculated by each sensor. Tight coupling means that the algorithm directly integrates the information obtained by each sensor for processing without preprocessing the information. It can be seen that under the loose coupling optimization method, each sensor is independent of each other and cannot really restrict each other. The tight coupling is to directly unify the raw data of each sensor for processing, and the sensor information is coupled with each other to obtain a better estimation effect. In most applications of multisensor fusion algorithms, tight coupling optimization methods are adopted.

The so-called tight coupling optimization is actually adding the IMU preintegration and visual observation and the camera pose to be optimized into a partial window for optimization. In this partial window, the variables to be optimized are as follows:

$$\begin{aligned} \chi &= [x_0, x_1, \dots, x_n, l_0, l_1, \dots, l_m], \\ x_k &= [p_{B_k}^W, v_{B_k}, R_{B_k}^W, b_a, b_g], k \in [0, n] \times, \\ l_j &= [x, y, z], j \in [0, m]. \end{aligned} \quad (15)$$

Among them,  $x_k$  is the system state of the  $k$ -th captured camera, including the position  $p_{B_k}^W$  of the IMU coordinate system relative to the world coordinate system, the speed  $v_{B_k}$  and the rotation  $R_{B_k}^W$  relative to the world coordinate system, and the offset  $b_a, b_g$  in the IMU measurement. Here,

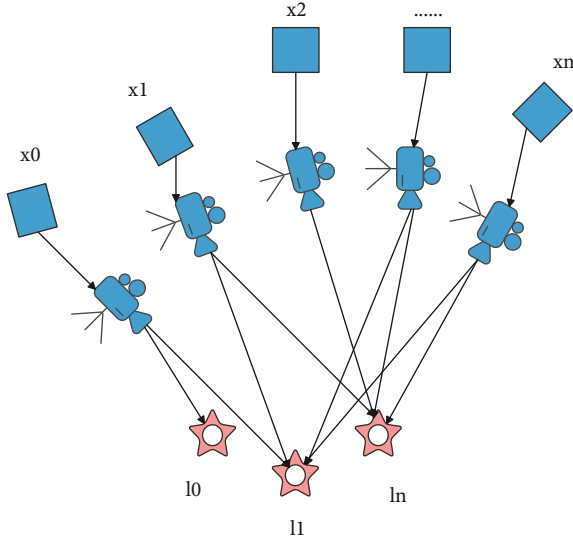


FIGURE 7: Schematic diagram of partial window.

the IMU coordinate system is selected as the navigation coordinate system, and the calculated pose is updated to the camera coordinate system through the external parameters between the IMU and the camera.

Figure 7 is a schematic diagram of partial window optimization in a certain period of time. The combination of the blue rectangle and the blue trapezoid is the camera, the blue rectangle is the IMU, and the black line connecting them is the external parameter. The red five-pointed star is the visual map point, and the dotted line between the camera and the visual map point indicates that the camera observes the map point. Figure 7 is a schematic diagram of map optimization corresponding to the partial window in Figure 8, and the map points are all summarized in blue ellipses. The yellow circle represents the state quantity to be estimated,  $P$  is the pose of the IMU,  $V$  is the speed of the IMU, and  $B$  is the partial mass measured by the IMU. The squares represent constraints, and the blue squares are the constraints of the reprojection error between the visual feature map points and the camera. The green square is the IMU preintegration constraint, and the relationship between the state to be estimated and the constraint is represented by a thin black line. For example, the blue square connects  $P$  and the map point.

The optimization problem represented by the above graph optimization is as follows:

$$\chi^* = \min_{\chi} \left( \sum_k E_{\text{proj}}(k, j) + E_{\text{IMU}}(i, j) \right). \quad (16)$$

The green square in the figure is the preintegration constraint of  $E_{\text{IMU}}(i, j)$ , and the blue square is the reprojection error constraint  $E_{\text{proj}}(k, j)$ . When the  $k$ -th map point and the  $j$ -th camera are given, the projection error  $E_{\text{proj}}(k, j)$  between them is defined as follows:

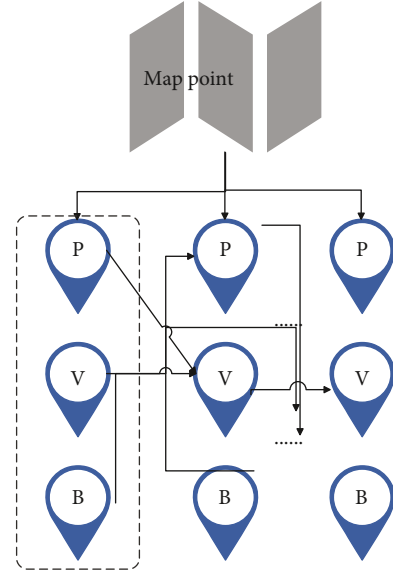


FIGURE 8: Schematic diagram of graph optimization.

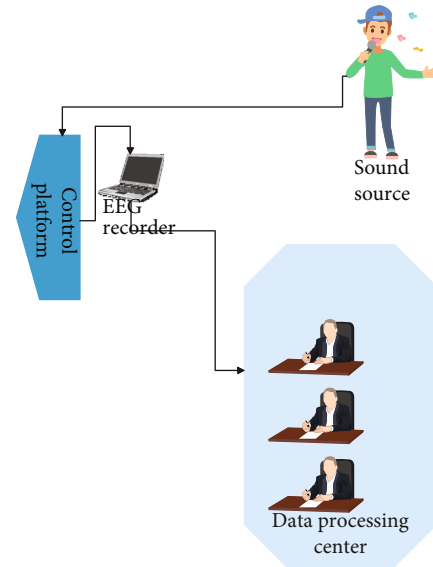


FIGURE 9: Singing and nervous system regulation platform based on wireless sensor perception network.

Among them,  $p_k$  represents the pixel coordinates of the feature point matching the  $k$ -th map point to the image, and  $\pi(\cdot)$  represents the projection transformation, which projects a point  $l_c^k$  in the visual coordinate system onto the image.  $\rho(\cdot)$  is a function of error mode detection. When the internal error exceeds a certain threshold, the weight of this error in the optimization is reduced to reduce the impact on the final estimation result.  $\sum_k$  is the information matrix of the reprojection error, which is related to the error model of the projection. The larger the error, the smaller the information value.  $l_c^k$  is converted from the pose in the IMU coordinate system and the camera IMU external parameter  $l_k$ . For the IMU preintegration constraint  $E_{\text{IMU}}(i, j)$ , there are

TABLE 1: Experimental research results of singing and nervous system regulation platform based on wireless sensor network perception.

Number	System assessment	Number	System assessment	Number	System assessment	Number	System assessment
1	93.37	14	95.44	27	96.31	40	94.58
2	93.99	15	85.41	28	84.21	41	85.68
3	86.76	16	95.52	29	93.39	42	94.53
4	92.27	17	88.17	30	91.51	43	84.71
5	96.50	18	86.79	31	91.04	44	91.44
6	85.98	19	84.70	32	86.38	45	85.68
7	94.75	20	85.05	33	92.13	46	89.26
8	85.93	21	90.95	34	92.56	47	85.79
9	89.06	22	94.82	35	93.93	48	84.47
10	92.73	23	91.69	36	89.98	49	90.47
11	85.60	24	95.35	37	95.54	50	95.64
12	94.33	25	84.12	38	85.40		
13	87.07	26	93.79	39	94.34		

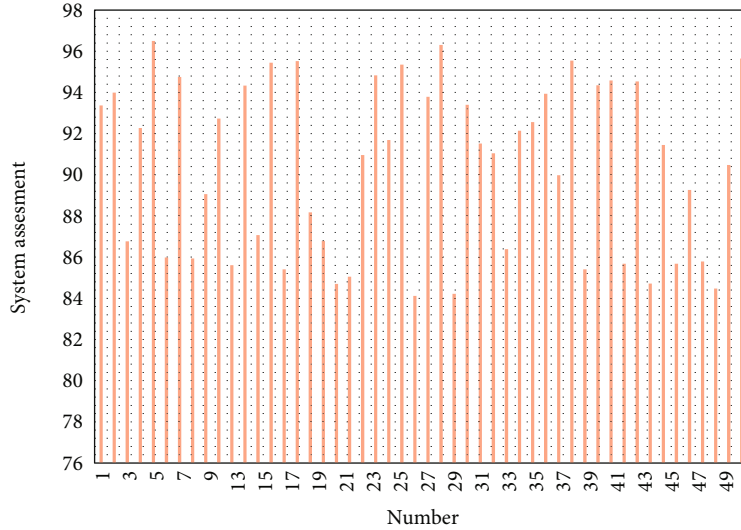


FIGURE 10: Histogram of experimental research results of singing and nervous system regulation platform based on wireless sensor network perception.

$$\begin{aligned}
E_{\text{IMU}}(i, j) &= \rho \left( \left[ e_R^T, e_v^T, e_p^T \right] \sum_l \left[ e_R^T, e_v^T, e_p^T \right]^T \right) + \rho \left( e_b^T \sum_b e_b \right), \\
e_R &= \log \left( (\Delta R_{ij} \exp(\delta b_{gi}))^T R_W^{B_i} R_{B_i}^W \right), \\
e_v &= R_W^{B_i} \left( v_{B_j} - v_{B_i} - w g \Delta t \right) - \Delta v_{ij}, \\
e_p &= R_W^{B_i} \left( p_{B_j}^W - p_{B_i}^W - v_{B_i} \Delta t - \frac{1}{2} g \Delta t^2 \right) - \Delta p_{ij}, \\
e_b &= b_j - b_i.
\end{aligned} \tag{17}$$

Among them,  $\sum_l$  is the information matrix of preinte-

gration, and  $\sum_b$  is the information matrix of partial quality. For the above graph optimization structure, an optimization solver is constructed, such as a Gauss-Newton solver or an LM solver. Therefore, the number of poses in the partial window designed in this section is limited to 10, which takes into account the calculation efficiency and accuracy of the algorithm.

#### 4. Singing and Nervous System Regulation Based on Wireless Sensor Perception Network

This paper constructs a mental emotion recognition system when singing, in which emotion recognition is directly linked to the nervous system. The data collection and



transmission of the system in this paper are realized through the wireless sensor network perception system. The system constructed in this paper is shown in Figure 9.

After constructing the system platform as shown in Figure 9, this paper evaluates the performance of this system through experimental research and studies the reliability of this system. This paper selects 50 volunteers from college vocal music majors with normal hearing and good health as the experimental subjects. Before the experiment, the subjects did not take any drugs that could affect the EEG and did not drink alcohol. At the same time, in order to ensure good electrical conductivity between the scalp and the electrode cap and improve the signal-to-noise ratio, all subjects are asked to wash their hair and dry their hair before the experiment started. Moreover, this paper uses a sensor network to simultaneously identify and locate 50 experimental subjects at the same time and calculate the effectiveness of this platform. The experimental results are shown in Table 1 and Figure 10.

From the above research results, it can be seen that the singing and nervous system regulation platform based on the wireless sensor network perception built in this paper has certain effects, and the system built in this paper can continue to be experimentally studied in the follow-up practice to explore its practical effects.

## 5. Conclusion

Singing is a manifestation of human beings' perception of the outside world or their own emotions. The content and emotions of singing are expressed through voice and body language. This is inseparable from the singer's own emotional mobilization and the coordination of the body's physiology. Moreover, whether the physical movement of the singer's body is correct or not directly affects the quality of the singer's voice and the expression of singing psychology. Therefore, the correct physiological movement of singing is inseparable from the guidance and regulation of correct singing psychological activity. Singing physiology is mainly composed of the physiological structure of the voice, respiratory system, resonance system, language system, and other elements. They affect and restrict the learning, teaching, and practical activities of vocal music. Only the coordinated movement of various singing physiological organs and parts can produce a beautiful singing voice. However, the production of the voice is inseparable from the scheduling and control of the singer's psychological activities, and the correct guidance of the singing psychological activities can lead to the correct physiological movements of the body. This article uses wireless sensor network applications to regulate singing and the nervous system, uses wireless sensor networks for data transmission, and simulates the human emotion model to study the regulation of singing and the nervous system. Through experimental research, it can be known that the singing and nervous system regulation platform based on wireless sensor network perception built in this paper has certain effects.

## Data Availability

The labeled dataset used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The author declares no competing interests.

## Acknowledgments

This study is sponsored by the Guizhou Normal University.

## References

- [1] G. Bedi, G. K. Venayagamoorthy, R. Singh, R. R. Brooks, and K. C. Wang, "Review of Internet of Things (IoT) in electric power and energy systems," *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 847–870, 2018.
- [2] I. Bisio, A. Delfino, A. Grattarola, F. Lavagetto, and A. Sciarrone, "Ultrasounds-based context sensing method and applications over the Internet of Things," *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 3876–3890, 2018.
- [3] A. Chamberlain, M. Bødker, A. Hazzard et al., "Audio technology and mobile human computer interaction," *International Journal of Mobile Human Computer Interaction*, vol. 9, no. 4, pp. 25–40, 2017.
- [4] D. B. Ç. Kiliç, "Pre-service music teachers' metaphorical perceptions of the concept of a music teaching program," *Journal of Education and Learning*, vol. 6, no. 3, pp. 273–286, 2017.
- [5] D. L. Hoffman and T. P. Novak, "Consumer and object experience in the internet of things: an assemblage theory approach," *Journal of Consumer Research*, vol. 44, no. 6, pp. 1178–1204, 2018.
- [6] B. Jia, L. Hao, C. Zhang, H. Zhao, and M. Khan, "An IoT service aggregation method based on dynamic planning for QoE restraints," *Mobile Networks and Applications*, vol. 24, no. 1, pp. 25–33, 2019.
- [7] J. Waldron, R. Mantie, H. Partti, and E. S. Tobias, "A brave new world: theory to practice in participatory culture and music learning and teaching," *Music Education Research*, vol. 20, no. 3, pp. 289–304, 2018.
- [8] J. Zhang and D. Tao, "Empowering things with intelligence: a survey of the progress, challenges, and opportunities in artificial intelligence of things," *IEEE Internet of Things Journal*, vol. 8, no. 10, pp. 7789–7817, 2021.
- [9] E. Gun, "The opinions of the preservice music teachers regarding the teaching of orchestra and chamber music courses during distance education process," *Cypriot Journal of Educational Sciences*, vol. 16, no. 3, pp. 1088–1096, 2021.
- [10] G. Muhammad, S. K. M. M. Rahman, A. Alelaiwi, and A. Alamri, "Smart health solution integrating IoT and cloud: a case study of voice pathology monitoring," *IEEE Communications Magazine*, vol. 55, no. 1, pp. 69–73, 2017.
- [11] X. Shengmin, "Analysis on the innovative strategy of national music teaching in colleges from the perspective of visual communication," *Studies in Sociology of Science*, vol. 7, no. 6, pp. 52–55, 2017.
- [12] Z. Lian, "Research on aesthetic education in instrumental music teaching," *Journal of Literature and Art Studies*, vol. 10, no. 5, pp. 435–439, 2020.

- [13] S. K. Kim, N. Sahu, and M. Preda, "Beginning of a new standard: Internet of Media Things," *KSII Transactions on internet and information systems*, vol. 11, no. 11, pp. 5182–5199, 2017.
- [14] A. Kaplan and M. Haenlein, "Siri, Siri, in my hand: who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence," *Business Horizons*, vol. 62, no. 1, pp. 15–25, 2019.
- [15] F. L. Reyes, "A community music approach to popular music teaching in formal music education," *The Canadian Music Educator*, vol. 59, no. 1, pp. 23–29, 2017.
- [16] V. K. Jones, "Voice-activated change: marketing in the age of artificial intelligence and virtual assistants," *Journal of Brand Strategy*, vol. 7, no. 3, pp. 233–245, 2018.
- [17] P. S. Aithal and S. Aithal, "Management of ICCT underlying technologies used for digital service innovation," *International Journal of Management, Technology, and Social Sciences (IJMTS)*, vol. 4, no. 2, pp. 110–136, 2019.
- [18] C. Johnson, "Teaching music online: changing pedagogical approach when moving to the online environment," *London Review of Education*, vol. 15, no. 3, pp. 439–456, 2017.
- [19] P. L. Lin, "Trends of internationalization in China's higher education: opportunities and challenges," *US-China Education Review B*, vol. 9, no. 1, pp. 1–12, 2019.
- [20] S. Y. Hong and Y. H. Hwang, "Design and implementation for iort based remote control robot using block-based programming," *Issues in Information Systems*, vol. 21, no. 4, pp. 317–330, 2020.
- [21] J. G. Bayley and J. Waldron, "'It's never too late': adult students and music learning in one online and offline convergent community music school," *International Journal of Music Education*, vol. 38, no. 1, pp. 36–51, 2020.
- [22] L. F. Luque Vega, E. Lopez Neri, A. Santoyo, J. Ruíz Duarte, and N. Farrera Vázquez, "Educational methodology based on active learning for mechatronics engineering students: towards educational mechatronics," *Computación y Sistemas*, vol. 23, no. 2, pp. 325–333, 2019.