Research on the Application of Artificial Intelligence Machine Learning Technology in Improving the Accuracy of Engineering Image Processing

Xiaowen Liu¹, Juncheng Lei²,*
¹School of Information Engineering, Shaoyang University
²Shaoyang University Network Information Center

*Corresponding author: ljc@hnsyu.edu.cn

Abstract. Image recognition technology mainly includes image feature extraction and classification recognition. Feature extraction is the key link, which determines whether the recognition performance is good or bad. Deep learning builds a model by building a hierarchical model structure like the human brain, extracting features layer by layer from the data. Applying deep learning to image recognition can further improve the accuracy of image recognition. Based on the idea of clustering, this article establishes a multi-mix Gaussian model for engineering image information in RGB color space through offline learning and expectation-maximization algorithms, to obtain a multi-mix cluster representation of engineering image information. Then use the sparse Gaussian machine learning model on the YCrCb color space to quickly learn the distribution of engineering images online, and design an engineering image recognizer based on multi-color space information.

Keywords: artificial intelligence; machine learning algorithm; engineering image processing; image accuracy.

1. Introduction
With the continuous development of technology, electronic devices such as smart phones and digital cameras have gradually become popular, and people have more and more ways to obtain images. How to extract useful feature information from many images, and then to distinguish the images has become very important. The proposal of deep learning algorithm in 2006, with its powerful mathematical representation ability, quickly became a research hotspot in the field of image recognition [1]. Deep learning uses deep neural network construction algorithms to extract image features hierarchically, and trains a large amount of data to automatically learn features and extract global features and contextual information, which greatly promotes the development of image recognition technology.

In recent years, deep learning algorithms for image recognition have continuously made breakthroughs, and algorithm models have continued to emerge. There are mainly early deep belief networks, denoising automatic coding machines, ternary factor Boltzmann machines, deep Boltzmann machines and convolutional neural networks. This paper takes the human-computer interaction coordinated assembly as the research background, takes the effective modeling of the engineering
image information in the hand assembly motion detection as the goal, uses the clustering idea to establish the mixed Gaussian model on the RGB color space, and realizes the project by the expectation maximization algorithm. The learning of image nonlinear distribution, and the online learning of the Gaussian machine learning model of engineering image sparseness on the YCrCb color space, which improves the model’s adaptability to light distortion in a certain work scene and improves the recognition effect, and finally achieves based on multi-color space information accuracy recognition of engineering image processing.

2. Gaussian mixture model

2.1. Hybrid model expectation maximization algorithm

As a non-parametric estimation method for modeling unknown targets or distributions based on observations, the K-mixed distribution model is expressed as a random form, describing that the labeled discrete variable \( Z \) of the observation data obeys a polynomial distribution, that is

\[
\begin{align*}
\mathbb{P}(x) &= \sum_{i=1}^{K} \phi_i p_i(x; \theta_i), \quad \phi > 0, \sum_i \phi = 1 \\
Z &\sim \text{Multinomial}(\phi_i, \phi_i, \ldots, \phi_i) \\
x | Z = p_i(x; \theta_i) \\
\sum_i \phi_i &= 1
\end{align*}
\]

(1)

In the above formula, \( p(x) \) is the probability value of variable \( x \) belonging to the target category, \( K \) parameters correspond to \( K \) single Gaussian distributions, \( Z \) is the single-mode component label corresponding to the subordinate polynomial distribution, \( \theta \) is the distribution parameter corresponding to each single Gaussian, \( \phi_i \) is the corresponding weight. Compared with the single-mode distribution, the above-mentioned mixed form has higher flexibility.

Using maximum likelihood estimation, select the likelihood value of the data to the model as the objective function when estimating the model parameters. The optimal solution of the parameter corresponds to the extreme position of the likelihood function [2]. The natural log-likelihood value \( l(\theta) \) of the data of \( K \)-mixing and \( n \)-observation and its derivative with respect to the parameters of the distribution family are

\[
l(\theta) = \sum_{i=1}^{n} \log p(x_i; \theta) =
\sum_{i=1}^{n} \log \left( \sum_{k=1}^{K} \phi_k p(x_i; \theta_k) \right)
\]

(2)

\[
[\Phi^*, \Theta^*] = \arg \max_{\phi, \theta} l(\phi, \theta, X)
\]

(3)

\[
\frac{\partial l(\theta)}{\partial \theta_j} = \sum_{i=1}^{n} \frac{\phi_i}{\sum_{k=1}^{K} \phi_k p(x_i; \theta_k)} \frac{\partial p(x_i; \theta_k)}{\partial \theta_j} = \sum_{i=1}^{n} \frac{\phi_i p(x_i; \theta_j)}{\sum_{k=1}^{K} \phi_k p(x_i; \theta_k)}
\]

(4)
The expectation maximization method respectively estimates the two factors in formula (4), alternately calculates the expectation of the weight parameters (E step), and estimates the distribution family parameters (M step).

Step E: Calculate the weights according to the current model parameters, record

\[ \omega^{(t)}_j = Q_j^{(t)} \left( z^{(t)} = j \right) = p \left( z^{(t)} = j \mid x^{(t)}, \phi^{(t)} \right) \]  

(5)

Step M: Solve the equation according to the current fixed weight parameters

\[ \frac{\partial}{\partial \theta} \sum_{i,j} \omega^{(t)}_j \log \left( \frac{p(x^{(t)}, z^{(t)}; \theta)}{Q_j^{(t)}(z^{(t)})} \right) = 0 \]  

(6)

Obtain the updated value of the clustering parameter \( \theta \), and update the polynomial parameters

\[ \phi_k = \sum_{i \in C_k} 1 \left( z^{(i)} = k \right) / N \]  

(7)

Among them is the category label indicator function according to the update results of each E step and M step. It can be proved that iterating the above process can make the target value of equation (2) monotonically increase, and the iterative process can obtain a stable optimization process.

### 2.2. Expectation maximization learning algorithm of Gaussian mixture model

The K-mixture Gaussian model of a certain multi-dimensional unknown distribution is expressed as

\[ f(x) = \sum_{i=1}^{K} \phi_i N(x \mid \mu_i, \Sigma_i) \]  

\[ \sum_{i=1}^{K} \phi_i = 1 \]  

(8)

In the formula, \( N \) is the multi-dimensional Gaussian distribution, \( \mu_i \) is the multi-dimensional mean vector, and \( \Sigma_i \) is the corresponding covariance matrix. The goal of model learning is to estimate all the parameters in equation (8) based on the training data \( X \). In the expectation step (E step), calculate the posterior expectation

\[ \omega^{(t)}_j = Q_j \left( z^{(i)} = j \right) = p \left( z^{(i)} = j \mid x^{(i)}, \phi^{(t)}, \mu^{(t)}, \Sigma^{(t)} \right) \]  

(9)

In the maximization step (M step), the weights obtained above are fixed, and the likelihood value of the training data is expressed as

\[ l(\phi, \mu, \Sigma) = \sum_{i \in C} \sum_{j \in C} \omega^{(t)}_j \log p \left( x^{(i)}, z^{(i)}; \phi, \mu, \Sigma \right) \geq \sum_{i \in C} \sum_{j \in C} \omega^{(t)}_j c(i, j) \]  

(10)
\[ C(i,j) = \log \left( \frac{1}{(2\pi)^{n/2}} \exp \left( -\frac{1}{2} (x_i^{(j)} - \mu_j)^T \Sigma_j^{-1} (x_i^{(j)} - \mu_j) \right) \phi_j^{(i)} \right) \] 

(11)

Fix \( \phi_j, \Sigma_j \), and calculate the partial derivative of the likelihood value with respect to \( \mu_j \)

\[ \nabla_{\mu_j} \sum_{i=1}^{K} \omega_j^{(i)} C(i,j) - \nabla_{\mu_j} \sum_{i=1}^{K} \omega_j^{(i)} \frac{1}{2} (x_i^{(j)} - \mu_j)^T \Sigma_j^{-1} (x_i^{(j)} - \mu_j) \]

\[ \sum_{i=1}^{n} \omega_j^{(i)} \left( \Sigma_j^{-1} x_i^{(j)} - \Sigma_j^{-1} \mu_j \right) \]

(12)

Based on this, the mean parameter update can be obtained

\[ \mu_j = \left( \sum_{i=1}^{n} \omega_j^{(i)} x_i^{(j)} \right) / \sum_{i=1}^{n} \omega_j^{(i)} \]

(13)

Similarly, the zero point of the partial derivative of the covariance can be obtained by the log likelihood function to obtain the corresponding covariance update

\[ \Sigma_j = \left( \sum_{i=1}^{n} \omega_j^{(i)} \right) \sum_{i=1}^{n} \phi_j^{(i)} (x_i^{(j)} - \mu_j)(x_i^{(j)} - \mu_j)^T \]

(14)

So far, the parameters of the Gaussian mixture model are updated.

3. Image processing information system design

3.1. The overall structure of the system

The overall structure of the system is mainly composed of five parts, as shown in Figure 1, which are CCD camera, FPGA chip, synchronous dynamic random access memory (SDRAM) chip, digital/analog converter (DAC) chip and liquid crystal display (LCD). Among them, the CCD camera uses DALSA’s Spyder3 series line scan camera; DDR2SDRAM uses the MT47H64M16 chip, which has a total storage capacity of 1GB, and the SDRAM uses IS42S16160B; the FPGA chip uses the CycloneIV series chip and the LCD uses the computer display.
Figure 1. Overall block diagram of the system structure

Engineering image acquisition is realized by SOPC on FPGA through the system bus to control the image acquisition module, and the image acquisition module controls the CCD camera. The collected data is stored in the external memory DDR2SDRAM, and then SOPC judges and classifies the image processing algorithm. The steps suitable for the hardware completion processing are sent to the corresponding hardware module through the bus with the corresponding parameters and commands, and the hardware module feeds back the end signal to the SOPC through the bus after completing the corresponding processing [3]. Algorithms that are not suitable for hardware processing are handed over to SOPC, and SOPC uses its internal CPU to perform operations. These two parts of operations are independent and independent of each other, achieving parallel acceleration between modules. In the entire engineering image processing system, SOPC has been processing in accordance with this method until the end of the processing, the processed image is displayed on the LCD, so that the image can be judged manually, and the algorithm can be improved and optimized to meet the actual application requirements.

3.2. Software and hardware collaborative engineering image processing system

The engineering image processing process can be divided into three parts: image preprocessing, feature extraction and engineering image recognition. This system combines FPGA software and hardware. The image preprocessing work is still described and executed by FPGA Verilog HDL. Feature extraction and engineering the image recognition is completed by SOPC on the FPGA, and the computer only undertakes the task of post-processing of engineering images. This system not only saves costs, but also accelerates the processing of engineering images in parallel. The specific process is shown in Figure 2.
Software and hardware coordination is to design the hardware and software parts of the system at the same time, so that the mutual influencing factors between the systems can coordinately complete the functions that need to be realized, improve the execution speed of the system, and enhance the stability of the system at the same time [4]. The design flow chart is shown in Figure 3.

4. Engineering image accuracy algorithm detection
The experimental environment is the human-machine coordinated assembly experiment scene, the lighting conditions are general lighting, and the imaging system adopts the 1.22-million-pixel industrial camera imaging system of Weishi image. Considering the two typical actions in the assembly action and the different complexity of the imaging scene, online tests were performed on three typical images. Initialization of the online test uses a mixture of Gaussian model to calculate the likelihood value, and the empirical likelihood values of 0.45 and 1.9 are respectively used as the two thresholds for the confirmed project image and the project image to be confirmed, that is, the image area with the likelihood value greater than 1.9 is regarded as the project image area. The image area with a likelihood value of less than 1.9 and greater than 0.45 is regarded as the area to be confirmed [5]. The project image area is corroded through the 3 3 template to obtain the initial confirmation area, and after 50 iterations to convergence.
Figure 4 shows the online engineering image sparse learning and detection test of different scene complexity and different typical assembly postures. It can be generally judged that the online sparse learning algorithm affects different parts of the "operator-workbench-workpiece" context. It can be seen from the probability density map that although the RGB Gaussian mixture model obtained according to learning can correspond to the overall high likelihood value of the hand region, there are still certain differences in the adaptability to different semantic complexity environments. The initialized image fails to obtain a complete segmentation in some areas with insufficient illumination. A relatively complete segmentation was obtained in the 3 test images at the same time. The relative relationship of the likelihood values on the 3 channels does not have too much repetition, indicating that there is no serious redundancy in the result information obtained by learning on the YCrCb3 channel, and the built model and the corresponding learning framework are effective [6]. The initialization results of test images in different illumination areas can be seen. There are many defects in the initialization of low-illumination areas under fixed threshold conditions. The initial model has low adaptability to the image, the initial confirmation area is small, and the model corresponding to the image area with insufficient illumination The likelihood value is low; when the illumination is sufficient, the initial likelihood value of the model corresponding to the engineering image processing accuracy area is high. In the experimental calculation, the more complete the initialization, the smaller the amount of calculation in the online sparse learning process.

Figure 4. Online engineering image sparse learning and detection.

5. Conclusion
This paper proposes a new sparse Gaussian machine learning model for engineering image detection, which uses multi-structure closed operation reconstruction to preprocess the original image. At the same time, it is ensured that the positions of the detected engineering image and the tracked engineering image are compared and synchronized every 10 frames. The sparse Gaussian machine learning model is more accurate and faster than the DFO and Det algorithms. Moreover, the sparse Gaussian machine learning model can also detect and track the target in the case of severe interference from the feature background, and can also correctly detect and track the short-term disappearance of the target.
References

[1] Liang, X. Image-based post-disaster inspection of reinforced concrete bridge systems using deep learning with Bayesian optimization. Computer-Aided Civil and Infrastructure Engineering, vol. 34, pp. 415-430, May 2019.

[2] Cruz, Y. J. Rivas, M. Quiza, R. Beruvides, G. & Haber, R. E. Computer vision system for welding inspection of liquefied petroleum gas pressure vessels based on combined digital image processing and deep learning techniques. Sensors, vol. 20, pp. 4505-4526, Sixteen 2020.

[3] Choi, E. & Kim, J. Deep learning based defect inspection using the intersection over minimum between search and abnormal regions. International Journal of Precision Engineering and Manufacturing, vol. 21, pp. 747-758, April 2020.

[4] Xie, Q. Li, D. Xu, J. Yu, Z. & Wang, J. Automatic detection and classification of sewer defects via hierarchical deep learning. IEEE Transactions on Automation Science and Engineering, vol. 16, pp. 1836-1847, April 2019.

[5] Yamane, T. & Chun, P. J. Crack detection from a concrete surface image based on semantic segmentation using deep learning. Journal of Advanced Concrete Technology, vol. 18, pp. 493-504, September 2020.

[6] Chao-Ching, H. Su, E. Li, P. C. Bolger, M. J. & Pan, H. N. Machine vision and deep learning based rubber gasket defect detection. Advances in Technology Innovation, vol. 5, pp. 76-79, February 2020.