A Rapid Quantitative Determination Method of AFP Concentration with Gold Immunochromatographic Strip

Haiyan Jiang
College of Electrical Engineering and Automation, Fuzhou University, Fuzhou, China
Fujian Key Laboratory of Medical Instrumentation & Pharmaceutical Technology, Fuzhou, China
Email: jianghaiyan@fzu.edu.cn

Min Du and Dongzhong Ke
College of Physics and Information Engineering, Fuzhou University, Fuzhou, China
Fujian Key Laboratory of Medical Instrumentation & Pharmaceutical Technology, Fuzhou, China
Email: kedongz@fzu.edu.cn

Abstract—alpha-fetoprotein (AFP) is a useful diagnostic marker for the detection of a great number of infantile diseases, especially for some forms of malignant tumors and liver disorders. Although there are several methods for quantitative determination of AFP, compared with these methods, the gold immunochromatographic assay (GICA) has the advantages of easy to handle, low costs, quick, no requirement for skilled technicians and most important, well suitable for point-of-care measurements. However, GICA mostly has been used for qualitative or semi-quantitative detection visually. This paper presents a rapid quantitative determination method of AFP with GICA strip based on the reflective optical detection. Under the driving of the micro-stepper motor, we can get a curve of GICA strip signal distribution. Then after segmentation of the test line and control line using fuzzy c-means clustering algorithm, the features were extracted to be as the input features of Support Vector Regression (SVR) model. SVR was used for prediction of the AFP concentration. To evaluate the reliability and accuracy of GICA for fast quantitatively determination of AFP and its clinical value, the AFP concentration of specimen was tested by GICA and chemiluminescent immunoassay analyzer, and analyzed the coincident rate and correlation of these two methods.

The experiment results demonstrated that the correlation coefficient results between the two methods was 0.961. The test results showed that the SVR model yields a good result and is proved to be appropriate in quantitative determination of AFP with GICA strip.

Index Terms—Gold immunochromatographic assay strip; Support Vector regression; alpha-fetoprotein (AFP)

I. INTRODUCTION

Alpha-fetoprotein (AFP) is a glycoprotein that is a prominent component of serum proteins in early embryonic life. Shortly before birth, serum AFP levels begin to fall, and AFP is replaced by albumin as the major serum protein. Serum AFP levels thereafter remain at very low levels throughout life, typically value <10 ng/mL. Serum AFP levels can become elevated again in adults with hepatocellular carcinoma (HCC), germ-cell tumors and liver disease [1, 2]. Measurement of AFP is used widely to assist in the diagnosis of HCC and increasingly is used as a screening tool for HCC in patients with chronic liver disease. It is a useful diagnostic marker for the detection and differentiation of a great number of infantile diseases, especially for some forms of malignant tumors and liver disorders [3, 4].

Gold immunochromatographic assay (GICA) strip is a kind of rapid immunolabelling analysis technology boomed in the early 1990s, which has been gaining increasing attentions for rapid and direct analyses of proteins, nucleic acids, hormones, pesticides and drugs [5-10]. Although a diverse range of laboratory methods is now available for detection AFP. These include radioimmunoassay, fluorescence immunoassay, real time RT-PCR methods and enzyme immunoassay methods. All these methods generally involve relatively sophisticated laboratory. The GICA strip has many advantages such as easy to handle, low costs, quick, no requirement for skilled technicians or expensive equipments, and perhaps most important, it is well suitable for point-of-care measurements [11-13]. Therefore, in recent years GICA strip has been rapidly developed and widely applied in biomedical field.

However, unlike traditional quantitative immunoassay methods, GICA mostly has been used for qualitative or semi-quantitative detection visually. For example, the positive or negative qualitative test results mainly by comparing with the color standard card [14]. The study on GICA quantitative determination not only enhances the sensitivity of detection and the objectivity of determination result, enlarges the detecting rang, but also has a sound application value. For example, human chorionic gonadotropin (hCG) is a heterogeneous molecule produced by trophoblastic cells in pregnancy and in gestational trophoblastic diseases. hCG can only qualitatively indicate the pregnancy status with its positive
or negative result, however with the quantitative determination results, hCG can provide more useful information in ectopic pregnancy differentiation and in fetal Down Syndrome screening test.

The GICA method can be designed as rapid, qualitative methods or as standard quantitative laboratory procedures and it can also be used as quantitative methods to identify samples that need to be analyzed further by analytical method. Therefore, research about how to use GICA strip for a rapid quantitative detection has been paid a great deal of research attention. Some research work focus on the quantitative determination method of GICA strip based on image processing techniques. In these works, the image of the strip was acquired by a commercial optical scanner such as the array Charge Coupled Device (CCD) sensor or Complementary Metal-Oxide Semiconductor (CMOS) sensor, and usually the distribution curve of image horizontal grayscale was obtained in one-dimensional space. What they importantly take into account is how to use a kind of artificial neural network image analysis system to realize the quantitative detection [15]. However, these work neglected GICA strip congenital characteristic, such as absorption spectra of gold nanoparticles on the strip, therefore they can only improve their sensitivities by algorithm compensation. Other research work based on the reflective optical detection using the spectra characters of the strip. In these work, the important procedures involve the design of the reflective optical detection module and the building of the regression algorithm to get quantitative result by extracting proper features from the control line and test line to establish a quantitative regression relationship between the features and the concentration of strip.

The traditional quantitative regression method is Back-Propagation Neural Network (BPNN). The BPNN method has some congenital insufficient, such as large calculation, falling into local extremum easily etc. In addition, it also suffers from difficulty in selecting a large number of controlling parameters such as hidden layer size, learning rate, and momentum term [16, 17].

Support vector regression (SVR) is an extension of the well developed theories of Support Vector Machine to regression problems with introduction of ε-Insensitivity loss function by Vapnik [18]. Unlike traditional learning algorithm for function estimation such as Neural Network that minimizes the error on the training data based on the principle of Empirical risk minimization, SVR embodies the principle of Structure Risk minimization which minimizes an upper bound on the expected risk. Hence, it is characterized by better ability to generalize, and at the same time less prone to the problems of overfitting and local minimal [19]. It is a powerful technique for solving the nonlinear function approximation problems, such as in power systems, computer networks, leeway prediction, Fuel cells, environment, etc [20-22].

In this paper, a rapid quantitative determination method of AFP immunochromatographic assay strip based on the reflective optical detection was proposed. We use fuzzy c-means clustering algorithm to segment the test line and control line from the obtained GICA strip signal curve, then extracted the features as the input features of SVR regression model to solve the nonlinear function regression problems in the determination of AFP concentration.

II. GOLD IMMUNOCROMATOGRAPHIC STRIP AND SYSTEM ARCHITECTURE

A. Gold Immunochromatographic Strip

Gold immunochromatographic strip is based on colloidal gold-labeled and chromatographic technology for the rapid detection. A typical immunochromatographic strip was constructed in the form of sandwich by using monoclonal antibodies with two distinct specificities. One antibody is immobilized in a defined detection zone on porous nitrocellulose membranes, while the other antibody was conjugated with colloidal gold nanoparticles which served as a detection reagent [8]. The schematic view of a typical GICA strip is shown in Fig. 1[7, 23]. The sample pad on top of the conjugate pad to prefilter the sample and the absorbent pad to take up the sample solution on the opposite side of the membrane complete the assay setup.

![Image](https://example.com/image1.png)

**Figure 1.** The schematic description of the GICA strip

Colloidal gold particles is a suspension (or colloid) of gold nanoparticles in a fluid. Colloidal gold particles have strong absorption on some optical wave with certain wave length. Absorption spectra of gold nanoparticles in diameters of 25nm are shown in figure 2. As seen from the figure, the maximum of the absorption spectrum is about 520 nm.

![Image](https://example.com/image2.png)

**Figure 2.** Absorption spectrum of gold nanoparticles

Once the sample is added, the sample solution through wicking migrated onto the strip by capillary action. As the sample flowed sequentially through the detection antibody
(conjugate pad) and the capture antibody, the gold conjugates get captured on test line and a red coloured band was visible. A second red coloured line was also observed on the control line of the membrane, generated by surplus gold conjugates, indicating the proper test performance. The coloured band must be visualized on the control line, so the test could be considered as invalid if there was no colour line present in the control region. The intensity of gold conjugates in test line on GICA strip correlates with concentration of the test target in a certain range. Therefore by detecting the content of colloidal gold particles on the strip after immunoreactions, we can obtain the concentration level of the target analyze in sample. This provides the basis for the quantitative determination of the AFP concentration with GICA strips.

B. Determining System Architecture

The general hardware architecture of the proposed determination system is illustrated in Fig. 3. The system mainly consists of digital signal processing unit (DSP), reflective optical detection module, data acquisition module, Scanning driving module, data analysis module, storage memory, liquid crystal display (LCD), and interface display module.

The system with a core of Texas Instrument TMS320F2812 digital signal processor (DSP), which in charge of driving the micro stepper motor and processing of data, setting the displays of liquid crystal and the printout of printer, and so on. TMS320F2812 is a 32-bit DSP of 150MHz maximum frequency which is highly integrated, high-performance solutions for demanding control applications. It can be applied in Adaptive control, Kalman filter, and other advanced control technologies efficiently and reliably.

The function of data acquisition module is acquiring the voltage signal from the reflective optical detection module, the voltage signal after amplifying was transferred from the analog signal data into digital signal data through the Analog-to-Digital Converter (ADC) peripherals. The ADC module of TMS320F2812 is a 12-bit pipelined analog-to-digital converter, including a converter, two programmable conversion sequencers and 16 sampling channels. The acquired digital signal was processed with a series of filtering processing and smoothness by data analysis module, such as median average filtering algorithm.

The driving module was composed of the micro stepper motor, motor driver circuit, bell wheel, etc. The module driving the strip holder shift to ensure acquire whole signal on both the test line and the control line. The test data also could be transported to PC through the communication module such as USB port for further analysis.

The most important component is the reflective optical detection module, which acquire the reflective optical signal of the GICA strip and transfers into electrical signal. The module consists of LED light source, photoelectric conversion module, lens, optical fiber, strip holder, etc. (see Fig.4)

According to the maximum wavelength of the colloidal gold absorption spectra, colloidal gold particles absorb green light strongly, therefore two green LED was used as the light source. The green LED was irradiate to the GICA strip and a cylindrical lens focused the LED output beam to a rectangular spot on the strip. The reflective optical was focused by a focusing lens and entered an optical fiber after optical filter, and finally sampled by the low noise, the high precision photodiode conversion module.

Under the driving of the motor, the green LED spot shift from the test line to control line for signal sample. Therefore the signal on the whole detection zone of the strip was acquired. These signals after pre-filter, was processed by the DSP to establish SVR nonlinear function regression, the first step includes some preprocesses of the digital signal data such as filter, smoothness. In order to predict the GICA concentration, we must extract the features of the test line and control line. In this paper Fuzzy C-means algorithm was used to obtain the position of the test line and control line.

The following flowchart (Fig. 5.) shows the overall steps in the determination method for the AFP with GICA strip based on the reflective optical detection.
Under the driving of the stepper motor, the signal of the test line and control line of these strip were acquired by the reflective optical detection module in turn, and was transferred from the analog signal data into digital signal data through the ADC peripherals after amplify. Finally we can get a curve of GICA strip signal distribution.

The top image of Fig.6 is the image of the GICA strip with AFP concentration 200ng/mL and 50ng/mL. The bottom image is the curve of GICA strip signal corresponding. The left peak of the strip is the test line; the other is the control lines. From the figure we can see that the high AFP concentration the higher peak of the test line. The acquired digital signal after pre-filter was transported to the DSP. In order to realize the quantitative determine of the GICA strip, the curve of the strip signal was first segment by Fuzzy C-means (FCM) algorithm to get the location of the test line and control line.

III. FUZZY C-MEANS (FCM) SEGMENTATION ALGORITHM AND SUPPORT VECTOR REGRESSION

A. Fuzzy C-means (FCM) Segmentation Algorithm

In order to get the location of the test line and control line, the signal data was segmented by Fuzzy C-means (FCM) algorithm. FCM is an unsupervised clustering algorithm that has been applied successfully to a number of problems such as feature analysis, clustering and classifier design [24]. The objective function of FCM cluster algorithm $E(U, V)$ is defined as (1).

$$E(U, V) = \sum_{i=1}^{K} \sum_{j=1}^{n} (\mu_{ij})^\alpha (d_{ij})^2 \tag{1}$$

Where $\sum_{j=1}^{K} \mu_{ij} = 1; 0 \leq \mu_{ij} \leq 1$  

$U$ —matrix of Membership degree,

$U = \{\mu_{ij}\} \ i = 1,2,\cdots,K ; j = 1,2,\cdots,n$ ;  

$\alpha$ — fuzzy coefficient;

$\mu_{ij}$—Membership degree of the object $x_j'$ belonging to the $i$th cluster center in the data collection;

$d_{ij}$ — refers to the distance between $x_j'$ and the cluster center $V_i$, $d_{ij} = \left\| x_j' - V_i \right\|$;

$V_i$ — the $i$th clustering center vector, $V = (V_1, V_2, \cdots, V_i, \cdots, V_K)^T$.

$K$ — clustering category, $2 \leq K \leq n$.

According to the Lagrange multiplier optimization algorithm and the Matrix of Clustering center $V$, the matrix of membership degree $U$ can be iterated as (2):

$$\mu^{(q+1)}_{ij} = \frac{1}{\sum_{j=1}^{K} \left( \sum_{k=1}^{n} (x_k - V^{(q)}_k) \right)^{2(\alpha-1)} \left( \sum_{k=1}^{n} (x_k - V^{(q)}_k)^2 \right)^{2(\alpha-1)}}$$

The iterative clustering center is calculated as (3):

$$V^{(q+1)}_i = \left( \sum_{j=1}^{n} \left( \sum_{j=1}^{n} u^{(q+1)}_{ij} \right)^{2/(\alpha-1)} \right)^{1/(2/(\alpha-1))} \tag{3}$$

The iteration continues according to (3)-(4), until $v^{(q+1)}_i - v^{q}_i \leq \epsilon_1$ and $u^{(q+1)} - u^{(q)} \leq \epsilon_2$ ($\epsilon_1, \epsilon_2 > 0$), then the clustering center and the membership degree are corresponding to the cluster category K.
In this experiment, set \( m = 2 \) and \( K = 3 \). The input features of FCM are the curve of the strip signal. An example of the segmentation result was shown in Fig.7.

![Figure 7. (a) the curve of the strip signal; (b) The location of control and test line segmented by FCM.](image)

After segment by the above method, the location of both the control and test line were obtained. In this study, appropriate and crucial features are required to achieve an accurate SVR prediction model. As mentioned previously, the peak value and the integration of the signal in test line provide the basis for the quantitative determination of the AFP. Therefore we compute the maximum and the integration value of the signal in the test lines, and the average of the control lines as the 3 input features of the SVR model. And finally SVR model was used to find a mapping function between the features and the AFP strip concentration.

B. Support Vector Regression

The basic concept of the SVR is to map the input data into a high-dimensional feature space by nonlinear mapping, to yield and solve a linear regression problem in this feature space. Given a set of data \((x_j, y_j), i = 1, \cdots, n, x \in \mathbb{R}^n, y \in \mathbb{R}\), the objective is to get a regression function \( y = f(x) \), which could accurately predict the output corresponding to a new set of input-output examples. To nonlinear regression, SVR using a \( \varepsilon \)-insensitive loss function and the kernel function \( k(x_i, y_j) = \langle \phi(x_i) \cdot \phi(x_j) \rangle \), which can map non-linear learning problem into linear one in a high-dimensional feature space. A best regression function \( f(x) = \langle \omega \cdot \phi(x) \rangle + b \) is estimated in that feature space.

The convex optimal regression function is given by the minimum of the following function:

\[
\min J = \frac{1}{2} \| \omega \|^2 + C \sum_{i=1}^{l} (\xi_i + \xi_i^*) \quad \begin{cases} y_i - (\omega \cdot \phi(x_i)) - b \leq \varepsilon + \xi_i \\ (\omega \cdot \phi(x_i)) + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, i = 1, \cdots, l \end{cases}
\]

\( C \) is the tradeoff constant between the smoothness of the SVR function and the total training error. We can write the Lagrangian and express the dual optimization problem:

\[
L(\omega, b, \xi, \xi^*, a, a^*, \eta, \eta^*) = \frac{1}{2} \| \omega \|^2 + C \sum_{i=1}^{l} (\xi_i + \xi_i^*) \\
- \sum_{i=1}^{l} a_i (\varepsilon + \xi_i - y_i + (\omega \cdot \phi(x_i)) + b) \\
- \sum_{i=1}^{l} a_i^* (\varepsilon + \xi_i^* + y_i - (\omega \cdot \phi(x_i)) - b) - \sum_{i=1}^{l} (\eta_i \xi_i + \eta_i^* \xi_i^*)
\]

\[
\begin{align*}
\frac{\partial L}{\partial \omega} & = 0 \Rightarrow \omega = \sum_{i=1}^{l} (a_i - a_i^*) \phi(x_i) \\
\frac{\partial L}{\partial b} & = 0 \Rightarrow \sum_{i=1}^{l} (a_i - a_i^*) = 0 \\
\frac{\partial L}{\partial \xi_i} & = 0 \Rightarrow \eta_i = C - a_i - \eta_i \\
\frac{\partial L}{\partial \xi_i^*} & = 0 \Rightarrow \eta_i^* = C - a_i^* - \eta_i^*
\end{align*}
\]

where \( a_i, a_i^* \) are the nonnegative Lagrange multipliers. After solve the above dual quadratic programming problem, we obtain the Lagrange multipliers \( a_i \) and \( a_i^* \). The regression function estimated by SVR can be expressed as following:

\[
f(x) = \omega \cdot \phi(x) + b = \sum_{i=1}^{l} (a_i - a_i^*) k(x_i \cdot x) + b
\]

Finally, we can predict the output for the input data by applying the above regression function.

C. Support Vector Regression Architecture

We designed a 3-dimensional vectors by the above 3 input features as the SVR model input vectors. In order to obtain appropriate SVR model, the diagram of the SVR model is shown in Fig.8.

Compare to other kernel functions, the following cubic spline function has the best performance in this experiment.

\[
k(x_i, y_j) = 1 + x_i \times y_j + \frac{1}{2} x_i \times y_j \times \min(x_i, y_j) - \frac{1}{6} (\min(x_i, y_j))^3
\]
In the SVR model, the $\varepsilon$-insensitive loss constant $\varepsilon = 0.001$, punishment factor $C = 1000$ and strength factor $\theta = 12$.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

In this experiment, the results of the proposed GIGA Quantitative method were compared with that of quantitative standard laboratory method, chemiluminescent Immunoassay analyzer (CLIA), which was considered as gold standard. A total of 115 clinical human serum samples were collected by Fujian Provincial Hospital, of those 43 were found to be AFP positive and 72 were found to be AFP negative by CLIA. All samples were tested using ROCHE Modular Analytics E170 according to the recommended procedure.

The positive and negative result obtained by the two methods was showed as Tab.1.

<table>
<thead>
<tr>
<th>the results of GICA method</th>
<th>the results of CLIA method(gold standard)</th>
</tr>
</thead>
<tbody>
<tr>
<td>positive</td>
<td>positive 43(tr) negative 0(fn) total 43</td>
</tr>
<tr>
<td>negative</td>
<td>positive 43(tr) negative 71(fn) total 77</td>
</tr>
</tbody>
</table>

In the table, (tr) both two method classified the test sample as positive (tn) both two method classified the test sample as negative (fn) CLIA classified the test sample as positive but GICA negative (fp) CLIA classified the test sample as positive but GICA negative.

Fourfold table and Chi-square test ($\chi^2$-test) were used to evaluate the test’s sensitivity and specificity. The true positive rate was 100% and a true negative rate was 98.61%. The results showed the determination method was accurate and reliable with $\chi^2 < 0.001$ and $P > 0.005$.

The coefficient of correlation of proposed GICA method with CLIA is defined by following [25]:

$$ corr(X, Y) = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} = \frac{S_{xy}}{n\sigma_x \sigma_y} $$

where $S_{xx} = \sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2$,

$$ S_{yy} = \sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2 $$

and $S_{xy} = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}$

The coefficient of correlation measures the strength of a possible linear relationship between two methods. Necessarily, $-1 \leq corr(x, y) \leq 1$ and a coefficient close to 1 indicates that the sequences vary in a similar way.

The coefficient of correlation obtained with the GICA and ASCCESS chemiluminescent analyzer $corr(x, y)$ was 0.961. (showed in Fig.9.)

In this study, qualitative determination based on chemiluminescent analyzer and the GICA strip was perfect agreement while AFP concentration <800 ng/mL. However, GICA strip determination get low performance when the AFP concentration > 800 ng/mL. This might be explained by human operation error since the AFP need to be diluted when concentration > 800 ng/mL.

V. CONCLUSION

This study proposed a rapid quantitative determination method for AFP with GICA strip based on the reflective optical detection. Under the driving of the stepper motor, we can get a curve of GICA strip signal distribution. Then after segmentation of the test line and control line using fuzzy c-means clustering algorithm, the features were extracted to be as the input features of SVR regression model. Finally, Support vector regression was used to solve the nonlinear function regression problems in the determination of the AFP concentration.

The comparison study of the GICA method and the chemiluminescent analyzer revealed a strong correlation coefficient of 0.961 between these two methods. The results demonstrate that the developed method is applicable to the quantitative determine the AFP.
concentration. It should be noted however, the error introduced by the human operation of dilution would account for low sensitivities while AFP values >800ng/mL.

The test results show that the SVR model yields a good result and is proved to be appropriate in quantitative determination of the AFP with GICA strip. The study can not only enhance the detection sensitivity and the objectivity of test result, also can be considered as suitable for point-of-care measurements, such as other sorts of antigens and antibodies determination.

ACKNOWLEDGMENT

This work was supported by the International Science and Technology Cooperation Project of China (Grant: No.2009DFA32050), Fujian Natural Science Foundation of China (Grant: No.2009J0016, 2010J01252).

REFERENCES


[22] Yuan Ren,Guang-Yi Cao, Xin-Jian Zhu. Predictive control of proton exchange membrane fuel cell (PEMFC) based on support vector regression machine, Proceedings of the


Haiyan Jiang was born in Shandong Province, China, in 1976. She received her master degree in control theory and control engineering from Fuzhou University, Fuzhou, China, in 2005. Now she is a lecture at Fuzhou University, where she is currently working toward the Ph.D. degree in electrical engineering. Her research interests include biomedical instrument and photoelectrical detection.

Min Du was born in Fujian Province, China, in 1955. She received her PhD in Electrical Engineering from Fuzhou University, Fuzhou, China, in 2005. Now she is a professor and doctoral supervisor at Fuzhou University. And since 2007, she is the associate director of Fujian Key Laboratory of Medical Instrumentation & Pharmaceutical Technology. Her research interests include smart instrument and photoelectrical system.

Dongzhong Ke was born in Fujian Province, China, in 1960. He received his Master Degree in Computer Science from Fuzhou University, Fuzhou, China, in 1982. Now he is a senior engineer at Fuzhou University. His research interests include smart instrument and computer application.