

# New algorithm and system for measuring size distribution of blood cells

Cuiping Yao (姚翠萍), Zheng Li (李政), and Zhenxi Zhang (张镇西)

Key Laboratory of Biomedical Information Engineering of Ministry of Education, P. R. China, and  
Institute of Biomedical Engineering, School of Life Science and Technology, Xi'an Jiaotong University, Xi'an 710049

Received February 16, 2004

In optical scattering particle sizing, a numerical transform is sought so that a particle size distribution can be determined from angular measurements of near forward scattering, which has been adopted in the measurement of blood cells. In this paper a new method of counting and classification of blood cell, laser light scattering method from stationary suspensions, is presented. The genetic algorithm combined with nonnegative least squared algorithm is employed to inverse the size distribution of blood cells. Numerical tests show that these techniques can be successfully applied to measuring size distribution of blood cell with high stability.

OCIS codes: 290.5820, 170.3890, 170.1530.

Blood cells counting and classification are the foundations of medical diagnosis, and the automatic counting and classification of blood cells have a very important meaning for decreasing the labor intensity of physicians and improving the diagnosis accuracy. At present, Coulter counter is used as a basic equipment in this field. Coulter counter can be used for many tests, but it has the problem that it is easy to be obstructed; moreover, it is too expensive for most hospitals. This paper presents a new method of automatic counting and classification of blood cells through measuring the near forward scattering of stationary suspensions. In fact, this method is based on the inversion of the angular distribution of near forward scattering from a collimated beam.

It is known<sup>[1]</sup> that the ensemble-averaged scattered intensity  $I(\theta)$  of a suspension containing single-scattering particles can be modeled as an integral transform of the particle size distribution  $n(x)$  as

$$I(\theta) = \int_0^{\infty} I(\theta, x)n(x)dx. \quad (1)$$

This is Fredholm integral of the first kind, in which  $x = ka$  with the particle radius  $a$ , nondimensionalized by wave number  $k$ , and  $\theta$  is the scattering angle. The kernel  $I(\theta, x)$  represents the scattering from a single particle, and  $n(x)$  represents the number density. In the optical particle-sizing problem, an inverse transform is sought so that  $n(x)$  can be found from measurements of  $I(\theta)$ . The kernel  $I(\theta, x)$  is usually modeled by Mie-scattering theory for spheres or diffraction approximation.

Our system is shown in Fig. 1. When laser light goes through the pinhole, spatial filter, and collimating lens, it will be changed into a uniform-width parallel light. When this light meets the blood sample contained in the sample volume, the light will be scattered. The scattered light is collected and changed into an electric signal by the linear charge coupled device (CCD) that is positioned on the back focal plane of the Fourier lens; and the pinhole plate is attached on the CCD to collimate the focused light<sup>[2]</sup>. The circuit signal is transmitted into the computer through the data-collector board under the control of computer. Thus the results of blood cells' quantities

and size distribution can be obtained.

Isotropic sphere is used to simulate blood cell<sup>[3]</sup>, and small-angle light scattering of white blood cell is approximated by the Fraunhofer diffraction, whereas anomalous diffraction is found to be much better approximation for the scattering by red blood cells and platelets<sup>[2]</sup>. In the diffraction approximation,  $I(\theta)$  is modeled as the Fraunhofer diffraction from circular apertures of the same radii as the particles,

$$I(\theta) = \frac{1}{k^2\theta^2} \int_0^{\infty} J_1^2(x\theta)x^2n(x)dx, \quad (2)$$

where  $J_1$  is a first-order Bessel function of the first kind. Whereas in anomalous diffraction we can obtain<sup>[4,5]</sup>

$$I(\theta) = \frac{9}{k^7\theta^6} \int_0^{\infty} [\sin(x\theta) - (x\theta)\cos(x\theta)]^2 n(x)dx. \quad (3)$$

In our study, the size of blood cell sample to be measured is between some range  $[D_{\min}, D_{\max}]$ , so the measuring of scattered light is also in some scattered range  $[\theta_{\min}, \theta_{\max}]$ , it was determined by the radial measuring range of CCD  $[r_{\min}, r_{\max}]$ . Here  $\theta_{\min} = r_{\min}/f$ ,  $\theta_{\max} = r_{\max}/f$ ,  $f$  is the focal length of the focusing lens. In Eq. (2), substitute  $kD/2$  for  $x$ , after discretization we get

$$I_i = \sum_{j=1}^M \frac{\pi f^2 D_j^2}{4\lambda r_i^2} J_1^2\left(\frac{\pi D_j r_i}{\lambda f}\right) n_j, \quad i = 1, 2, \dots, K. \quad (4)$$

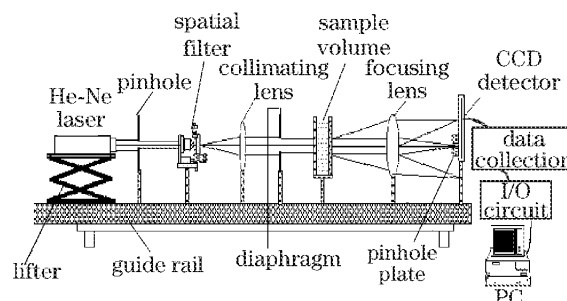


Fig. 1. Schematic of the optical blood cell-sizing instrument.

Similarly, Eq. (3) can be changed into

$$I_i = \sum_{j=1}^M \frac{9f^6 r_i^6}{64\pi^6 r_j^6} \left[ \sin\left(\frac{\pi r_i D_j}{f\lambda}\right) - \left(\frac{\pi r_i D_j}{f\lambda}\right) \cos\left(\frac{\pi r_i D_j}{f\lambda}\right) \right]^2 n_j, \quad i = 1, 2, \dots, K, \quad (5)$$

where  $n_j$  is the quantity percentage of cell whose diameter is  $D_j$  in the whole sample,  $M$  is the number of cell size range,  $K = (r_{\max} - r_{\min})/\Delta r$  is the number of sampling units of CCD, with  $\Delta r$  as the distance between each pair of neighboring CCD sampling units. Equations (4) and (5) can be written as a sample matrix form, i.e.

$$\mathbf{I} = \mathbf{T}\mathbf{N}, \quad (6)$$

where  $\mathbf{I} = (I_1, I_2, \dots, I_K)^T$  is light intensity distribution vector,  $\mathbf{N} = (n_1, n_2, \dots, n_M)^T$  is size distribution vector,

$$\mathbf{T} = \begin{bmatrix} t_{1,1} & t_{1,2} & \dots & t_{1,M} \\ t_{2,1} & t_{2,2} & \dots & t_{2,M} \\ \vdots & & & \\ t_{K,1} & t_{K,2} & & t_{K,M} \end{bmatrix} \quad (7)$$

is light intensity distribution coefficient matrix. The light intensity distribution vector  $\mathbf{I}$  is standardized. To invert the size distribution of blood cells from angular light scattering data we need to solve the following optimal problem:

$$\begin{aligned} \min_{\mathbf{N} \in R^M} \varphi(\mathbf{N}) &= \sum_{i=1}^K \left( I_{\text{calc}}^{(i)} - I_{\text{meas}}^{(i)} \right)^2, \\ \mathbf{N} &\geq 0, \quad \sum_{j=1}^M n_j = 1, \end{aligned} \quad (8)$$

where  $I_{\text{meas}}^{(i)}$  is the scattered light intensity in the position of CCD unit part from the light axis  $r_i$ ,  $\mathbf{I}_{\text{calc}} = \mathbf{T}\mathbf{N}$ .

In this study, the genetic algorithm (GA) combined with nonnegative least squared (NNLS) algorithm is employed to solve the optimal problem. In the selection of GA, the certain better individuals of the current generation will be kept so that the possibility of degeneration of the solution will be reduced. A mutation strategy with which only the ill individuals within the population are mutated is employed, and with the increase of the generation the mutated individuals are decreased.

We have done many tests to investigate the performance of the GA-based inverse algorithm for inverting size distribution of blood cells from angular light-scattering data. The measured data used in the study were collected from the apparatus shown in Fig. 1, the blood samples were randomly taken from the people who came to take blood test in our University Hospital Clinical Laboratory. The numbers of samples from males and those from females are the same (we have tested approximately 1000 persons), and the blood is fresh fingertip blood. The ages of tested persons range from 22 to 75. The blood diluents we used are the common diluents

with the parameters as follows. Red blood cell diluent: 0.9% NaCl liquid; white blood cell diluent: 3% acetum; blood platelet diluent: 1% ammonium oxalate.

We dilute the blood to solutions with different red blood cell concentrations (200×, 400×, and 800×) and different white blood cell and platelet concentrations (20×, 40×, and 60×). Each sample's blood cell quantity is counted and classified beforehand by counting platelet method (manual work) for comparing with the results by our method. It is found that when the concentration for red blood cells of 200×, for the white blood cells and platelets of 40× is used, we can get better result.

The numerical tests are performed with light-scattering data sampled from a small-angle region. We use 7 angular locations from 0.9 to 9.8 degrees. In the numerical tests the size of the population is chosen as 100<sup>[6]</sup>, the number of reserved better individuals in the current population is 30. In this study we combine GA with NNLS algorithm, which means that we set the result of GA as the initial value of NNLS to accelerate reliable convergence rate.

We measured the size distribution of red blood cells, white blood cells, and platelets. The mean diameter of red blood cells was also measured, and we classified the white blood cells to three types.

According to Ref. [7], human red blood cells have mean diameters of 7.4 μm and the platelets have the diameter range of 2 – 4 μm. The numerical tests are performed to invert the distribution of the red blood cells and platelets with GA-based inverse techniques. The stop criterion is chosen as that point where the maximum generation reaches 200. The probabilities of crossover and mutation are 0.70<sup>[8]</sup> and 0.2, respectively. We have tested about 1000 samples with every sample tested 3 times, and the average result was obtained. The distributions of the relative error of distribution of three kinds of blood cells for all samples are shown in Fig. 2. From the figure we can know, the relative error is almost within ±5%, only for some white blood cells it reaches 10%. Here is a concrete sample. The blood was taken from a male tester who is 50 years old. The relative refractive index of the dilution is 1.33. Figures 3 and 4 are, respectively, the results of the red blood cells and platelets by the GA-based inversion. In addition, we have conducted the tests with two kinds of sample volume: 10 × 2 × 45 mm<sup>3</sup> (v1) and 10 × 5 × 45 mm<sup>3</sup> (v2).

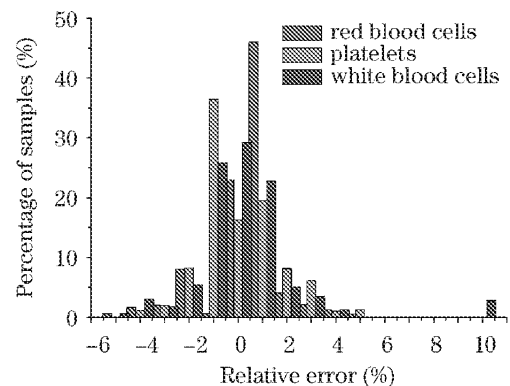


Fig. 2. The relative error distribution of the blood cells for the GA-based method.

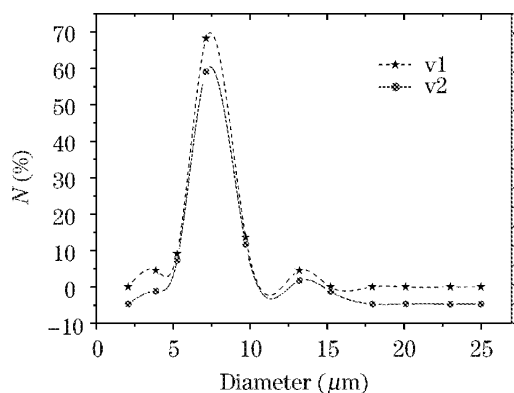


Fig. 3. Inversion results for red blood cells' distributions.

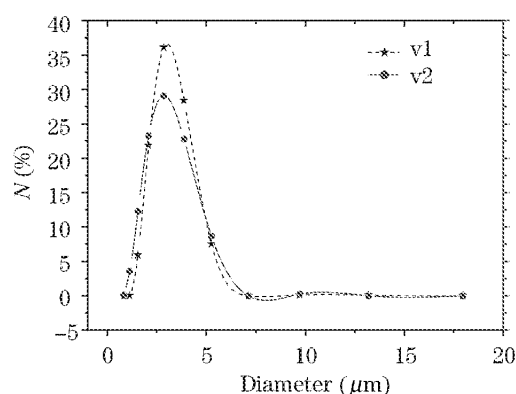


Fig. 4. Inversion results for platelets' distributions.

We classified the white blood cells by measuring the size distribution of white blood cells. According to Ref. [7], the size between  $4.67 - 7.6 \mu\text{m}$  can be counted as lymphocyte (actually, lymphocyte most between  $5.6 - 8.0 \mu\text{m}$ ). We count  $7.6 - 10.1 \mu\text{m}$  cells as middle-sized cells, which may include less frequently occurring and rare cells correlating to monocytes, eosinophils, basophils, blasts, and other precursor white cells. And we regard  $10.1 - 17.0 \mu\text{m}$  as neutrocyte. So we can classify white blood cells as lymphocyte, neutrocyte, and middle-sized cells according to their sizes. The conditions for the tests are chosen to be the same as for the former tests, including a maximum generation of 200, probability of crossover 0.70, probability of mutation 0.20, and refractive index of the dilution 1.33. Also about 1000 different samples were tested and the three test results were averaged for every sample. The distribution of relative error is shown in Fig. 2. Two typical samples were given as follows.

Tester1: 50-year-old female. The counting result of counting platelet method is lymphocyte 27%, neutrocyte 72%, and monocyte 1%. Figure 5 is the result of the white blood cells' distribution by the GA-based inverse technique (averaging the three tests).

Tester2: 68-year-old male. The counting result of counting platelet method is lymphocyte 62%, neutrocyte 30%, basophil 1%, acidophil 2%, monocyte 2%, and non-normal cells 3%. Figure 6 is the result of the white blood cell's distribution by the GA-based inverse technique (averaging the three tests). The tester is a leukaemia sufferer.

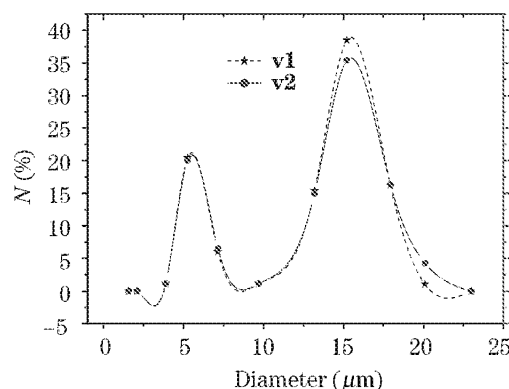


Fig. 5. Inversion results for white blood cells' distributions of the tester1.

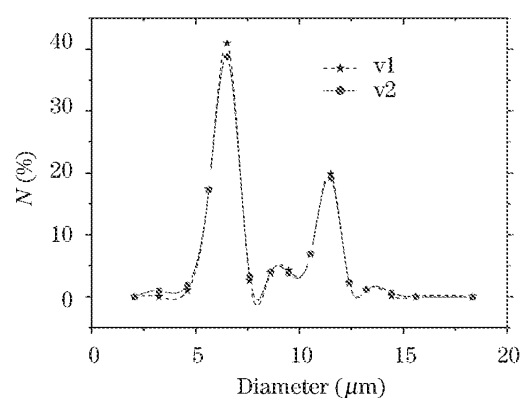


Fig. 6. Inversion results for white blood cells' distributions of the tester2.

To compare the agreement of the different sample volumes and the results of the blood cells' distributions by the GA-based inverse technique, the mean diameters of the red blood cells and platelets as well as the standard deviations are given in Tables 1 and 2.

**Table 1. Mean Diameters and Standard Deviations of the Difference between the Inversion Results and Clinic Measurements**

Parameter	Mean Diameter		Standard Deviation	
	v1	v2	v1	v2
Red Blood Cell	$7.45 \mu\text{m}$	$7.8 \mu\text{m}$	0.0025	0.16
Platelet	$3.08 \mu\text{m}$	$2.9 \mu\text{m}$	0.0064	0.01

**Table 2. Standard Deviations of the Difference between the Inversion Results and Clinic Measurements for White Blood Cells**

Parameter		Standard Deviation	
		v1	v2
Lymphocyte	Tester1	0.424	0.613
	Tester2	0.011	0.013
Middle-Sized Cells	Tester1	0.023	0.03
	Tester2	0.007	0.041
Neutrocyte	Tester1	0.645	0.916
	Tester2	0.001	0.008

Totally, the results plotted in Figs. 3 – 6 are in good agreement with the results by the counting platelet method. It can be seen from the inversion results that for the small-angle region the GA-based inverse technique is less susceptible to the volume depth the light penetrated and is capable of dividing the white blood cells into three classes with satisfactory accuracy. On the other hand, the effect of computing time of the GA-based inverse technique is evaluated. In comparison with the Ref. [9], we have shown that the GA-based inverse technique can work more efficiently than the inverse Monte Carlo (IMC) technique. However, in that study GA was employed and the stop criterion was chosen at the point where the maximum generation reaches 5000, whereas in our study we have improved the algorithm and combined it with NNLS, the stop criterion was used when the maximum generation reaches 200, which shows that the GA-based inverse technique can work more efficiently. In other studies, the different particle inversion methods were used, either for the standard particles or for the concentration of reticulocyte. Our group has used this method for the classification of all blood cells. The proposed algorithm overcomes the shortcomings of NNLS and GA, that is, for NNLS, if the initialization is not suitable, it will take more time; for GA, its reliable convergence rate is not so good.

In this study a new method of counting and classification of blood cells is presented, and an inverse technique based on GA for blood cells' sizing from angular light-scattering data is developed. In this method, we combine NNLS with the genetic method. We sampled 7 data on the CCD between 0.9 to 9.8 degrees, and each pixel of the CCD is 14  $\mu\text{m}$ , it meets the sampling criterion<sup>[1]</sup>. As a result, it may be concluded that the GA-based inverse technique can be used to classify the blood cells based on the laser light scattering method

from stationary suspensions. It is also shown that the new method not only can reach the global scope research, but also has relatively quick convergence. Finally we have shown the GA-based inverse technique is more efficient than the IMC technique in use of computing time to obtain results with satisfactory accuracy.

This work was supported by the National Natural Science Foundation of China (No. 60178034) and the Natural Science Foundation of Shaanxi Province (No. 96c17). Z. Zhang's e-mail address is zxzhang@xjtu.edu.cn, C. Yao's e-mail address is cpyao@sina.com.

## References

1. J. B. Riley and Y. C. Agrawal, *Appl. Opt.* **30**, 4800 (1991).
2. Y. Yang, *Blood Cell Counting and Classification by Stationary Suspending Laser Light Scattering Method* Ph. D. Dissertation (in Chinese) (Xi'an Jiaotong University, Xi'an, 1998).
3. Y. Yang, Z. X. Zhang, and D. Z. Jiang, *Journal of Biomedical Engineering* (in Chinese) **14**, 170 (1997).
4. H. Schnablegger and O. Glatter, *Appl. Opt.* **30**, 4889 (1991).
5. L. L. Li, *Platelet Size Distribution Measured by Static Suspending Mode Laser Light Scattering Method* Master Dissertation (in Chinese) (Xi'an Jiaotong University, Xi'an, 1999).
6. T. Dirk, *Evolutionary Computation* **7**, 331 (1999).
7. X. H. Lü, *Modern Medical Instrument and Its Application* (in Chinese) **6** (4), 12 (1994).
8. A. A. Salman, C. K. Mohan, and K. Mehrotra, *Evolutionary Computation* **8**, 341 (2000).
9. M. Ye, S. M. Wang, Y. Lu, T. Hu, Z. Zhu, and Y. Q. Xu, *Appl. Opt.* **38**, 2677 (1999).