JOURNAL OF OPTOELECTRONICS AND ADVANCED MATERIALS Vol. 8, No. 1, February 2006, p. 376 - 382

Normalization algorithm of the fiber drop fingerprint obtained from the optoelectronic drop analyzer

QING SONG[a*], GUOXIONG ZHANG[b], ZURONG QIU[b], JIAN XU[c], ZHONGPING FANG[c]

[a]*Automation School, Beijing University of Posts and Telecommunications, Beijing 100876, China*

[b]*State Key Laboratory of Precision Measuring Technology and Instruments, College of Precision Instrument and Opto-electronics Engineering, Tianjin University, Tianjin 300072, China*

[c]*Singapore Institute of Manufacturing and Technology, Singapore, 638075*

The fiber drop fingerprint (FDF) results from the light signal variation inside the liquid drop during drop formation and it can be used for liquid identification. Normalization is a necessary prior step before extracting features from FDF and constructing the liquid database. The processing method and algorithm for normalizing FDF is introduced in this paper. Start and end positions of each drop are determined through extremum method and threshold detection, based on excellent linearity of the capacitive signals. Normalized fiber and capacitive data in each normalized interval are obtained by using linear interpolation. Multiple drop signals corresponding to the same normalization position are averaged to get the final fingerprint. The mathematical deduction will be discussed in detail. Some graphic experimental results are presented.

(Received October 17, 2005; accepted January 26, 2006)

Keywords: Drop analysis, Optoelectronic, Fiber, Fingerprint, Normalization

1. Introduction

The optoelectronic drop analyzer (OEDA) is a new instrument that has been developed on the principles of fiber, capacitive and image drop analysis [1]. Fig. 1 is the schematic diagram of OEDA. The tested liquid is slowly delivered by the software-controlled micro-flow feeding pump and shaped into a satiated and uniform drop after being pumped into the drop head. The modulated infrared light is injected into the liquid drop by a source fiber positioned in the drop head and is collected by a detector fiber on the opposite side after various reflection, refraction and absorption of the optical signal inside the drop. The coupled light intensity changes along with the drop growing and produces a curve named fiber drop fingerprint (FDF) [2].

In addition, the specially designed capacitive sensor uses the drop head as one of its plate and a cylindrical ring plate, which surrounds the drop head and the space occupied by the formed drop, as another. The drop, which can be seen either as an extension of the drop head plate if the liquid is highly conductive, or as a dielectric material if it is less conductive, changes the capacitance along with drop growth. The instant drop volume can be obtained through a simplified mathematics model [3].

The CCD image processing provides another choice for drop volume measurement and drop growth monitoring, by making direct records of the instantaneous drop shape during its formation based on real-time image acquisition and image storage. The drop volume can be determined using a Soble or Laplacian edge detection method and image processing technology [4].

By merging the fiber signal and the capacitive signal or equivalent drop volume signal, a volume-based fiber drop fingerprint (VFDF) can be constructed [5], which shows the variation regularity of the light signal passing through the liquid drop during the drop growth. The VFDF uses the equivalent drop volume as the horizontal axis, instead of time series. The new representation of VFDF makes the time-based FDF independent from the speed of drop growth and the volatility of liquid, and accordingly ensures the repeatability of measurement against the variation of the feeding speed of the pump. In this paper, FDF refers to VFDF directly.

Large quantities of experiments for different samples [6] prove that the FDF provides a very fruitful source of information on the bulk properties of the tested liquids, and it is unique and definite for a certain liquid under certain conditions. So FDF is favorable for fine discrimination among different liquids.

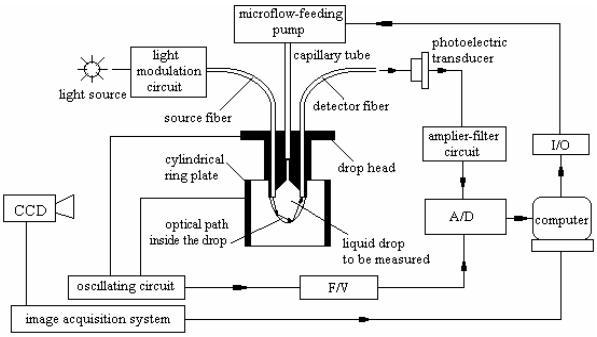

Fig. 1. The principle of the optoelectronic drop analyzer (OEDA).

Although visual features and qualitative differences can be observed in FDFs of different liquids, the method for constructing the liquid database and for extracting features from FDF should be studied, in order to quantitate fine discrimination and measurement of liquid properties. Original captured data of different tested liquids should be normalized to be with the same data format and data length, and also to represent the multiple measuring results and to reflect the characteristics of FDF clearly. FDFs of different liquids are comparable after normalization and suitable for feature extraction. It is the purpose of this paper to introduce the normalization method and algorithm of FDF in detail.

2. Normalization method and algorithm

The aim of normalization is to process the original measured data containing multiple drops into "one drop". So the fingerprints after normalization reflect the signal variation during one drop period. There are three steps. Firstly, start and end positions of each drop are determined through extremum method and threshold detection, based on excellent linearity of the capacitive signals. Secondly, normalized fiber and capacitive data in each normalized interval are obtained by using linear interpolation. Thirdly, multiple drop signals corresponding to the same normalization position are averaged to get the final fingerprint.

2.1 Detection for the drop separation points

The fiber curve $f(N)$ produced by the original light intensity signal according to the data acquisition system is shown in Fig. 2(a), which includes several continuous drops from "drop 1" to "drop 5". This is time-based fiber drop fingerprint as mentioned in introduction. The simultaneously obtained capacitor curve $c(N)$ is proved to be excellently linear, as shown in Fig. 2(b). The volume-based fiber drop fingerprint is shown in Fig. 2(c), by merging the fiber signal and the capacitor signal.

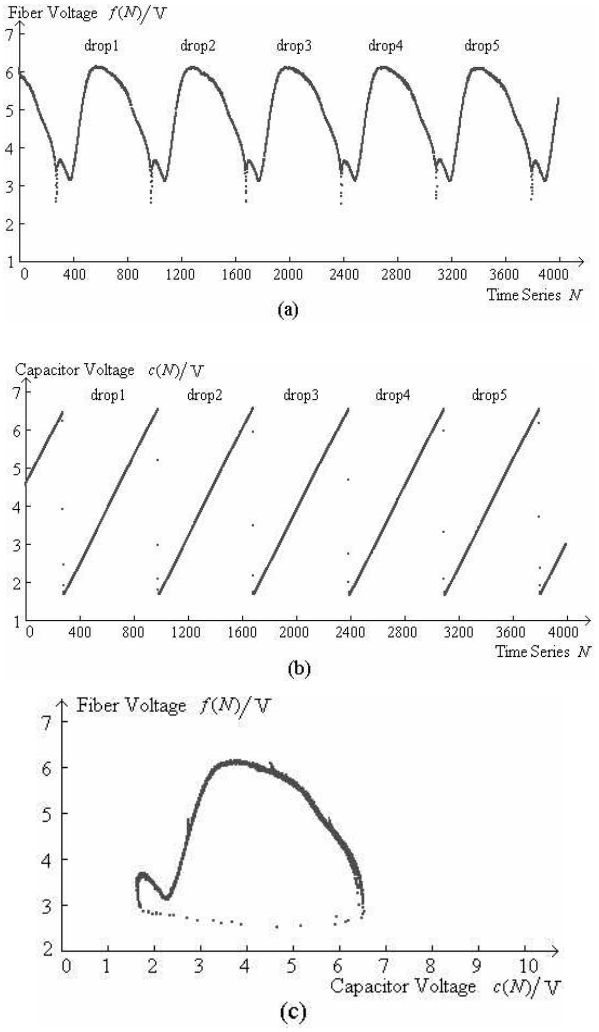

Fig. 2. The experimental graphs of pure water obtained from the drop analyzer. (a) The fiber drop sensor signal (the time-based fiber drop fingerprint); (b) The capacitor drop sensor signal; (c) The volume-based fiber drop fingerprint.

The first step of normalization is to determine separation points between multiple drops according to the linearity of the capacitor signal. In other words, the start positions ($\min Num[i]$) and end positions ($\max Num[i]$) of each drop are decided based on extremum method, and the number of the liquid drops ($nDrops$) is also got.

Extremum method is relatively simple in theory. From the second datum to the converse second datum in time series, those data bigger than their adjacent preceding and posterior data can be defined as maximum and those smaller can be defined as minimum. There are two

problems: one is that neglecting the first and the last data means possibly neglecting the first drop (if the first datum is minimum) and the last drop (if the last datum is maximum) in the measurement. Generally a certain algorithm can be used to estimate the first and last data in strict data analysis. But the liquid drop fingerprint as a special curve, is an overlapping result of multiple drops in a measurement cycle and a repetitive result of multiple measurements for a certain liquid. So it is unnecessary to put emphasis on a single drop, so as to simplify the program. Another problem is that the capacitor signal should be strictly increasing. However in fact there are some abnormal data and the capacitor curve appears to be slightly fluctuating in linearly increasing trend as a whole. Threshold detection provides a choice to solve this problem. Firstly, the biggest ($\max Data$) and the smallest voltage value ($\min Data$) are found out. Afterwards the threshold coefficient α is specified (usually $\alpha = 5\%$) and the threshold voltage for maximum ($\Delta V_{\max}$) and for minimum ($\Delta V_{\min}$) can be achieved, as shown in Fig. 3(a). Accordingly, those data falling outside the voltage threshold line cannot be decided as extremum. In brief the first decision rule is that the maximum should be enough close to the biggest and the minimum enough close to the smallest, as can be expressed in the following mathematic formula:

$$\max Data - c(\max Num[i]) < \Delta V_{\max} = \max Data \times \alpha \quad (1)$$

$$c(\min Num[i]) - \min Data < \Delta V_{\min} = \min Data \times \alpha \quad (2)$$

As for those data which falls inside the voltage threshold line but are not extremum, the time threshold line is employed. The second decision rule is that every extremum should be enough far from preceding extremum as the following mathematic representation:

$$\max Num[i] - \max Num[i-1] > \Delta N_{\max} \quad (3)$$

$$\min Num[i] - \min Num[i-1] > \Delta N_{\min} \quad (4)$$

It is worth noting that the minimum positions are searched in a sequential order among the original data array, while the maximum positions in an inverse order so that the formula (3) is effective.

Now take an example for minimum. Referring to Fig. 3(b), the first decision rule excludes the possibility of those abnormal data in area "B" to be minimum and the second rule excludes those in area "A". In order not to miss any abnormal data, area "A" and area "B" should be contiguous, as shown in instance (I) of Fig. 3(b), or should be intersectional, as shown in instance (II) of Fig. 3(b). Otherwise there will be a missing area as shown in instance (III) of Fig. 3(b). In this case, abnormal data in the missing area meet both the first and second decision rules but they are not minimum. So the threshold value of time series for minimum ($\Delta N_{\min}$) should be at lest greater than $\Delta N_{\min T}$ in instance (I). The same principle is suitable for maximum, as shown in Fig. 3(c). The above process can be expressed as:

$$\Delta N_{\max} \geq \Delta N_{\max T} = \frac{\max Data \times \alpha}{\max Data - \min Data} \times L_{TS} \quad (5)$$

$$\Delta N_{\min} \geq \Delta N_{\min T} = \frac{\min Data \times \alpha}{\max Data - \min Data} \times L_{TS} \quad (6)$$

where L_{TS} is the time series length of a single drop.

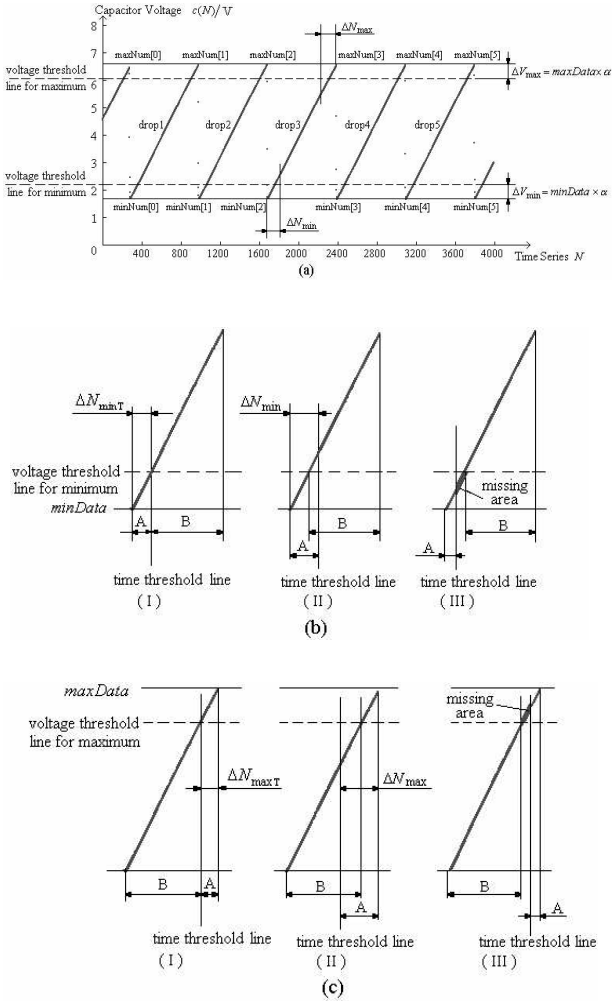

Fig. 3. Detection for the drop separation points between multiple drops according to the capacitor signal. (a) Extremum seeking and voltage threshold method; (b) Localization of time threshold line for minimum seeking; (c) Localization of time threshold line for maximum seeking.

2.2 Calculation of the normalized fiber and capacitor values in each drop

After extremum-seeking process is finished, the normalization region for each drop can be got according to various forms of original capacitor signal. As has mentioned above, the maximum positions are searched in an inverse order among original data, which means $\max Num[5]$ shown in Fig. 3(a) is actually $maNum[0]$ in extremum seeking process, and $maNum[0]$ shown in Fig. 3(a) is actually $\max Num[5]$. Now the elements in maximum data array are inverted as the same order in Fig. 3(a), so as to be convenient for representation in the following formulae. Suppose that the number of elements in maximum data array is $nMax$, and then as for instance (I) and (II) of Fig. 4(a), the number of the liquid drops $nDrops = nMax - 1$ and the normalization region is $(\min Num[i], \max Num[i+1])$, $i = 0,1,...nDrops-1$. As for instance (III) and (IV) of Fig. 4(a), the number of the liquid drops $nDrops = nMax$ and the normalization region is $(\min Num[i], \max Num[i])$, $i = 0,1,...nDrops-1$. The following formulae is directly applicable for instance (I) and (II). Instance (III) and (IV) are in the same principle.

The second step of normalization is to preset the number of normalization points (NP) and to evenly divide each drop into NP parts in time series, as shown in Fig. 4(b). Of course NP is usually not equal to the number of original measurement points in each drop, but maybe the positions of some normalization points are coincident with the positions of measurement points. After that, a linear interpolation is used to calculate the normalized fiber values and capacitor values at all positions of normalization points for each drop, according to the original data at the measurement positions adjacent to each normalization position. Adjacent points here means the measurement points preceding and posterior to the normalization point after comparing a normalization point with measurement points in time series.

In every normalization region of $(\min Num[i], \max Num[i+1])$, the normalization unit ($unit[i]$) is firstly calculated:

$$unit[i] = \frac{\max Num[i+1] - \min Num[i]}{NP-1}, \; i = 0,1,...nDrops-1. \tag{7}$$

Consequently the positions of normalization points ($NorPos[i][j]$) can be got:

$$NorPos[i][j] = \min Num[i] + unit[i] \times j, \quad j = 0,1,...NP-1; \quad i = 0,1,...nDrops-1. \tag{8}$$

Then the linear interpolation is used in the measurement region where the normalization points fall, as shown in Fig. 4(c). When $M \le NorPos[i][j] \le M+1$, $M \in (\min Num[i], \max Num[i+1])$, then there will be the following relation in the region of $(M, M+1)$:

$$\frac{f(M+1) - f(NorPos[i][j])}{f(M+1) - f(M)} = \frac{(M+1) - (NorPos[i][j])}{(M+1) - (M)} = \frac{c(M+1) - c(NorPos[i][j])}{c(M+1) - c(M)} \tag{9}$$

So the normalized fiber values and capacitor values at the position $NorPos[i][j]$ are expressed as:

$$f(NorPos[i][j] = f(M+1) - \frac{M+1-NorPos[i][j]}{f(M+1) - f(M)} \tag{10}$$

$$c(NorPos[i][j] = c(M+1) - \frac{M+1-NorPos[i][j]}{c(M+1) - c(M)} \tag{11}$$

where $j = 0,1,...NP-1$, $i = 0,1,...nDrops-1$, M and $M+1$ are adjacent points mentioned above.

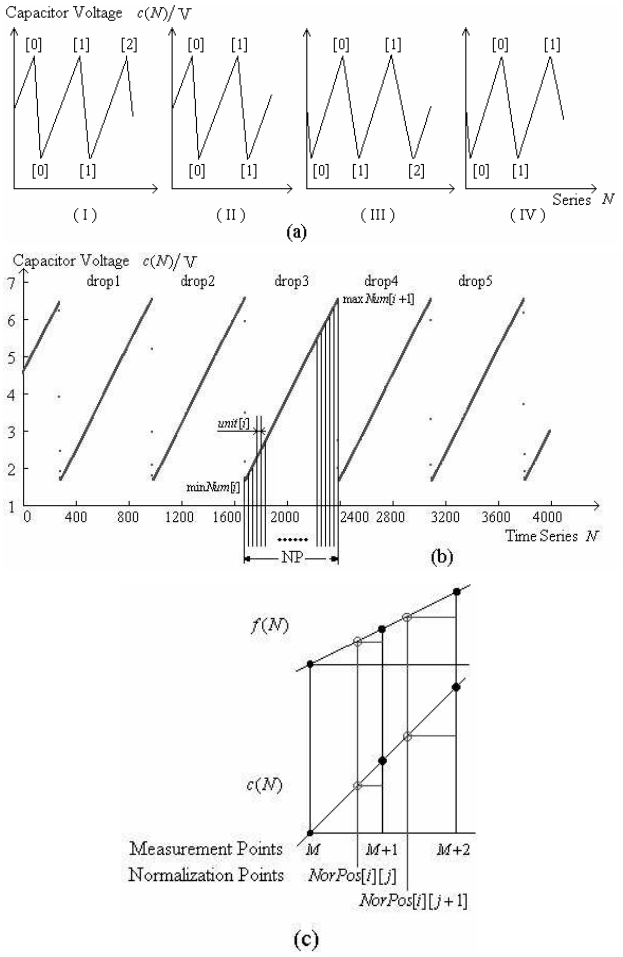

Fig. 4. Calculation of the normalized fiber and capacitor values. (a) Four forms of the original capacitor signal; (b) Divide each drop into NP parts; (c) Linear interpolation.

2.3 Fiber drop fingerprint constructed from the averaged values of multiple drops

The third and the last step of normalization is to average the fiber values and capacitor values of multiple drops correspondingly in the same position of normalization points, to get the final normalized fiber values ($fibNorMean[j]$) and capacitor values ($capNorMean[j]$), as represented in the following formula:

$$fibNorMean[j] = \frac{\sum_{i=0}^{nDrops-1} f(NorPos[i][j])}{nDrops}, \quad j = 0,1,...NP-1 \tag{12}$$

$$capNorMean[j] = \frac{\sum_{i=0}^{nDrops-1} c(NorPos[i][j])}{nDrops}, \quad j = 0,1,...NP-1 \tag{13}$$

Fig. 5(a) and 5(b) show the fiber signal and the capacitor signal of pure water after normalization, when $NP = 1000$.

The normalized fiber drop fingerprint (FDF) can be plotted from the normalized fiber and capacitor data, as shown in Fig. 5(c). In comparison with the original FDF shown in Fig. 2(c), constructed by raw fiber and capacitor signals, normalized FDF can represent the multiple measuring results. Normalization makes the characteristics of FDF more clear and apparent, on the basis of obedience to the original data. What's more, normalization improves the smoothness of FDF curve greatly.

It is worth noting that there is a significant difference between Fig. 5(a) and Fig. 5(c). Fig. 5(a) is the light signal variation with time series, viz. time-based fiber drop fingerprint. It is just the processing result by normalizing the original multiple drops shown in Fig. 2(a) into "one drop". Fig. 5(c) is the normalized volume-based fiber drop fingerprint, which uses the capacitor signal as the horizontal axis. This graph will be used directly for liquid identification and property study in the future.

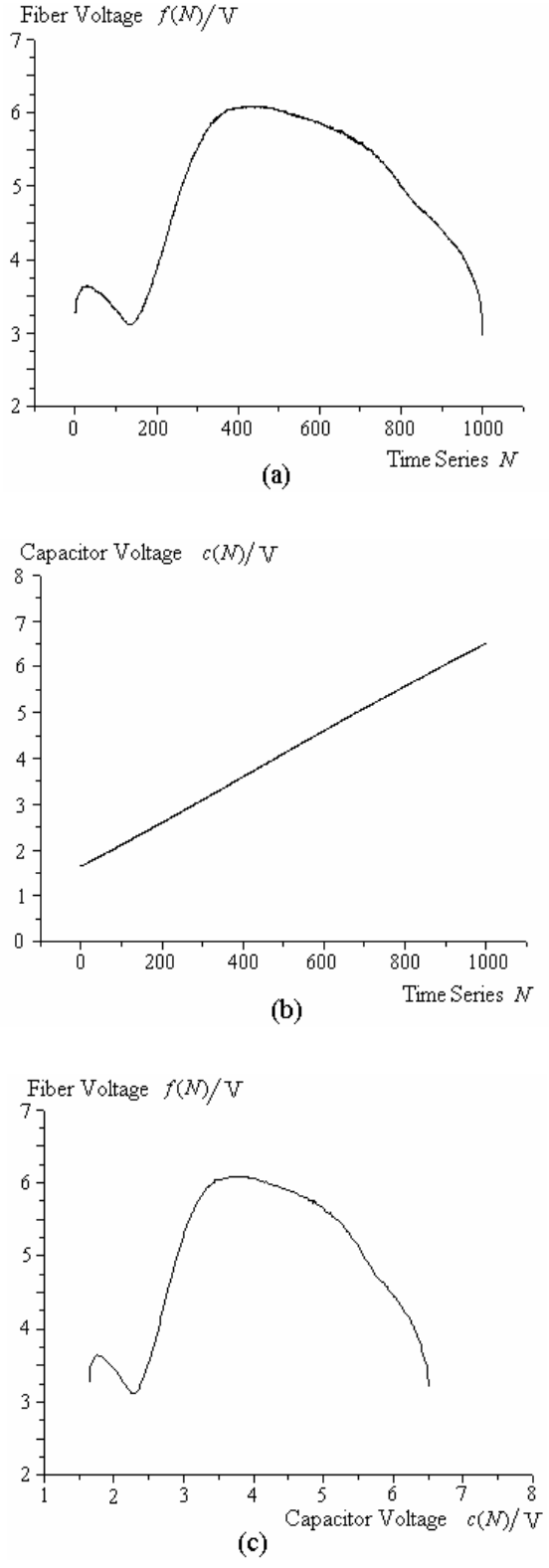

Fig. 5. Signals of pure water after normalization. (a) The fiber signal; (b) The capacitor signal; (c) Normalized fiber drop fingerprint.

3. Experiments and discussions

Fig. 6 are fiber drop fingerprints of different kinds of liquids, including KangShiFu pure water, Coca Cola, Beijing ErGuoTou white spirits, LaoChou soy and ShanXi mature vinegar. Fig. 6(a) shows their original FDFs and Fig. 6(b) shows their normalized FDFs, with the same normalization parameter $NP = 1000$.

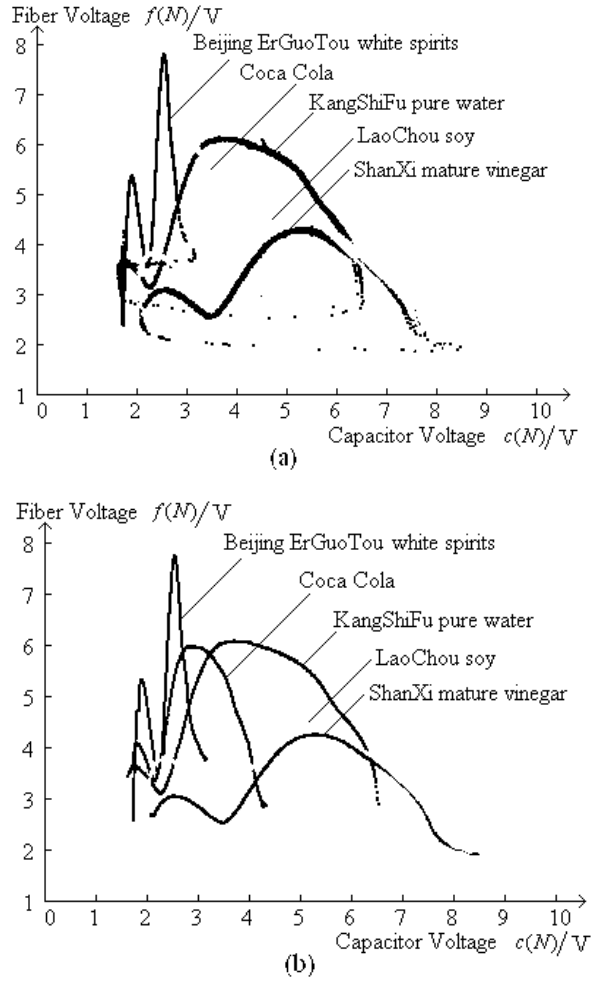

Fig. 6. Comparison of the fiber drop fingerprints of different liquids. (a) Original FDFs; (b) Normalized FDFs.

Although normalization improves the smoothness of FDF curve, it should not be considered as a process of noise filtering or data smoothing. Normalization is necessary, because:

As we know, all processing on the FDF such as feature extraction, are essentially operated on the captured data, which compose the fingerprints. In fact, the representation of volume-based FDF constructs a fingerprint in one drop period overlapped by multiple drops, only from the aspect of visible graph. But the data itself still involves multiple drops.

Suppose that the total feeding flow flux is the same and the data length for acquisition is the same, and then the number of measured drops of different liquids are different, because the drop volume of different liquids are different. There are only 5 drops for "Kangshifu pure water" as shown in Fig. 2(a), while there are 12 drops for "Beijing Erguotou white spirits", as shown in Fig. 7. It is obvious that in this case their fingerprint data (including fiber and capacitor data) are incomparable, because their data corresponding to the same time series reflect different hour or position during drop growth. In addition, on condition that the total feeding flow flux is different or the data length for acquisition is different when measuring different liquids, their fingerprint data are also incomparable. Therefore it is infeasible to extract features from the original captured data and it is quite necessary to normalize the fingerprints.

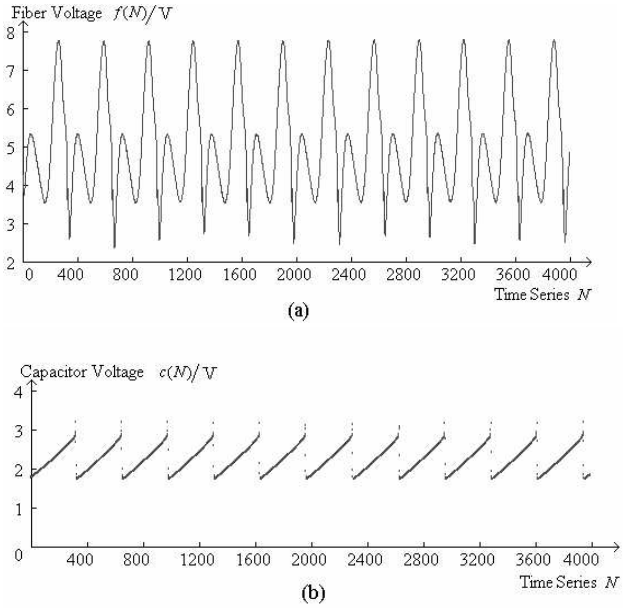

Fig. 7. The experimental graphs of Beijing Erguotou white spirits. (a) The fiber signal; (b) The capacitor signal.

Fig. 8(a) and 8(b) show the fiber signal and the capacitor signal of Beijing Erguotou white spirits after normalization, when $NP = 1000$. These data is now comparable with those of pure water shown in Fig. 5.

It should be emphasized that NP must be set the same value when compare FDFs of different liquids, because only in this case the signal corresponding to the same time series is comparable. Its meaning is to evenly divide one drop of different liquids into NP parts in drop volume, or in time axis if the speed of drop growth is uniform, and then we study their signals at the corresponding same position.

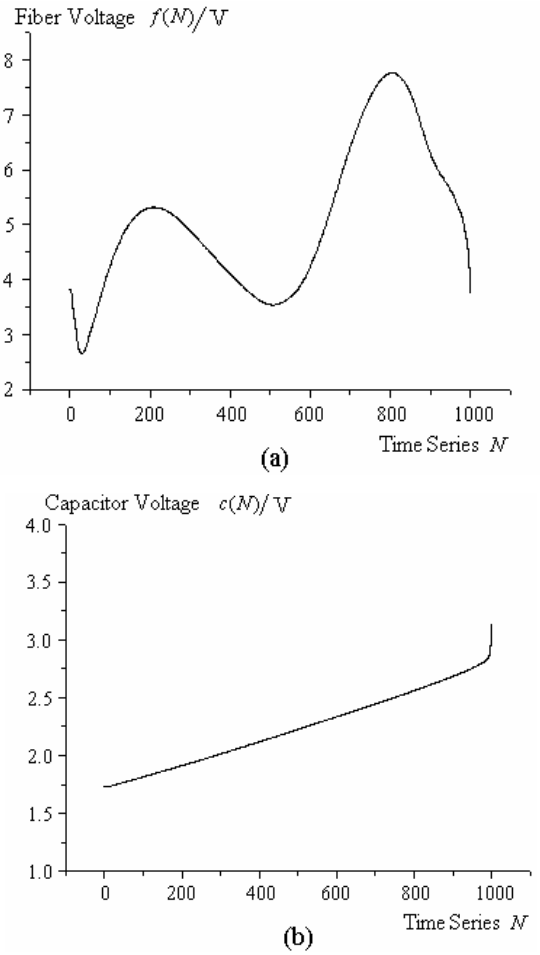

Fig. 8. Signals of Beijing Erguotou white spirits after normalization. (a) The fiber signal; (b) The capacitor signal.

4. Conclusions

The fiber drop fingerprint (FDF) can be obtained by fiber drop analysis, capacitive drop analysis or image drop analysis. FDF reveals the variation regularity of the light signal passing through the liquid drop during the drop growth, and reflects some physical and chemical properties of the tested liquids. It is unique and definite for a certain liquid under certain testing conditions, just like the fingerprint of a certain person. Therefore FDF can be used to discriminate different liquids, for example, to distinguish quality goods from counterfeits, such as fake beverage, fake medicine and fake wine.

Normalization is a necessary prior step before constructing the liquid database and for extracting features from FDF. Its main purpose is to make FDFs of different liquids comparable. The processing method introduced in this paper is firstly to determine start and end positions of each drop based on excellent linearity of the capacitive signals. Secondly normalized fiber and capacitive data in each normalized interval are obtained by using linear interpolation. Multiple drop signals corresponding to the same normalization position are averaged to get the final fingerprint. The fingerprints after normalization reflect the signal variation during one drop period. The anticipative aims are achieved through experimental verifying.

References

[1] Song Qing, Zhang Guoxiong, Qiu Zurong, Zhang Aiping, Shi Qingwei, Instrumentation Science and Technology **32**, 43 (2004).
[2] N. D. McMillan, O. Finlayson, F. Fortune, M. Fingleton, D. Dely, D. Townsend, D. D. G. McMillan, M. J. Dalton, Measure Science and Technology **3**, 746 (1992).
[3] C. H. Wang, A. T. Augousti, J. Mason, N. D. McMillan, Measure Science and Technology **10**, 19 **(**1999).
[4] Song Qing, Zhang Guoxiong, Qiu Zurong, Instrumentation Science and Technology **31**, 1 (2003).
[5] Song Qing, Qiu Zurong, Zhang Guoxiong, J. Optoelectron. Adv. Mater. **6**, 393 (2004).
[6] Song Qing, Zhang Guoxioing, Qiu Zurong, Chen Haixiu, Chinese Journal of Scientific Instrument **24**, 348 (2003).

* Corresponding author: sunny512@eyou.com