

Curve fitting e compensazione per segnali di citometria ad impedenza

Mastrofini Alessandro

alessandro.mastrofini@alumni.uniroma2.eu

Abstract

Nella citometria ad impedenza sono state sviluppate diverse tecniche, di progettazione e di post processing, per migliorare l'accuratezza nella misura. Uno dei parametri che più affligge le misurazioni di citometria è la dipendenza del segnale dalla posizione della cellula all'interno del canale per cui cellule identiche, passanti in posizioni diverse, danno luogo a segnali differenti.

Nel seguente report viene analizzata una configurazione ad elettrodi planari e la dipendenza dei segnali dalla posizione all'interno del canale microfluidico. Viene quindi analizzata e applicata una tecnica di compensazione su un dataset derivante da misurazioni su famiglie con tre differenti diametri elettrici.

Una mappatura secondo la retta interpolante il parametro di forma e il diametro normalizzato consente di rimuovere il contributo di posizione e separare i segnali di famiglie con differenti diametri, quindi risalire alle singole distribuzioni eliminando il contributo di posizione.

1 Introduzione

I sistemi per l'analisi microfluidica implementano soluzioni tecnologiche per contare, intrappolare, focalizzare, separare e identificare le proprietà della cellula.

Possono essere identificate le proprietà di singole cellule basandosi su differenze nelle dimensioni e nelle proprietà dielettriche, sfruttando tecniche non invasive e senza richiedere l'ausilio di labelling cellulare.

Tra le differenti tecniche risalta l'analisi di impedenza in microfluidica. Si utilizzano più coppie di elettrodi, poste intorno al canale microfluidico, per applicare una tensione a frequenze discrete e misurare le fluttuazioni della corrente elettrica. Tali fluttuazioni sono dovute al passaggio della cellula all'interno del canale e quindi all'interno del campo elettrico non uniforme generato dagli elettrodi.

Tali sistemi possono essere più o meno accurati a seconda della configurazione utilizzata.

In questo report viene presentata l'analisi di una tecnica di compensazione per eliminare la dipendenza dalla posizione nel canale che affligge i segnali misurati e non permette una corretta stima del diametro cellulare.

Tale tecnica di compensazione sfrutta la dipendenza del diametro elettrico dal parametro di forma, legato alle proprietà del segnale misurato. Nonostante le diverse famiglie analizzate (differenti diametri) appaiono come segnali non distinguibili il mapping mediante tale procedura di compensazione permette di produrre dei segnali dove le diverse famiglie sono chiaramente distinguibili.

2 Background

Il citometro ad impedenza è un dispositivo microfluidico utilizzato per misurare le perturbazioni del campo elettrico all'interno di un microcanale dove passa una cellula. La cellula attraversa uno

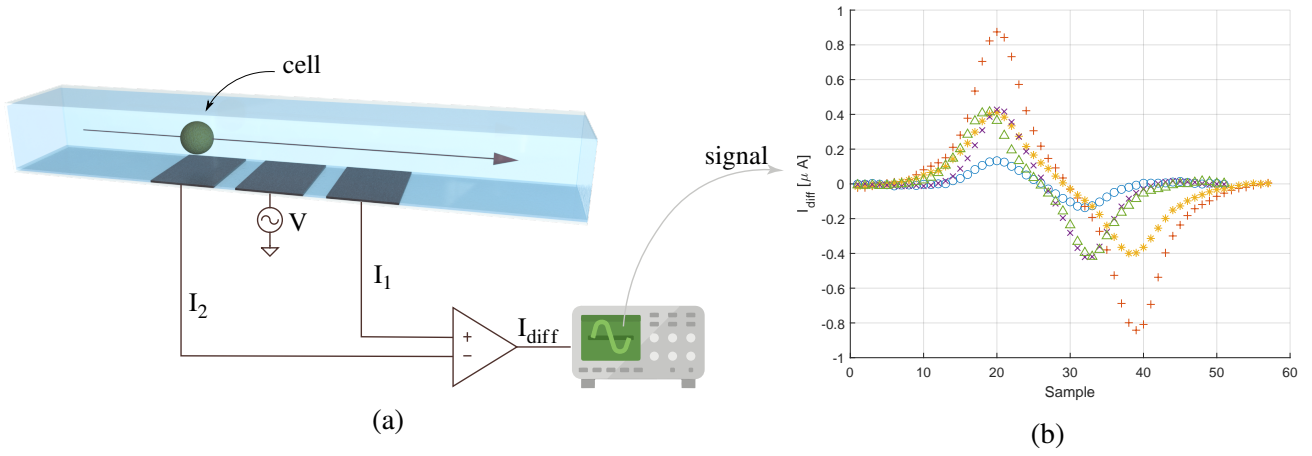


FIG. 1: Schema rappresentativo del citometro ad impedenza con la configurazione di elettrodi selezionata e la cellula passante (a); segnali della corrente differenziale misurata per 5 diversi campioni (b)

specifico pattern di elettrodi sul quale viene applicata un tensione alternata e questo porta ad una variazione della corrente misurata.

Ovvero si considera un microcanale riempito di un buffer conduttivo al cui interno passano delle correnti elettriche. Nel dispositivo in questione si applica un potenziale all'elettrodo centrale e si misura una corrente differenziale tra i due elettrodi laterali.

Si utilizzando basse frequenze per determinare le proprietà sulla dimensione della cellula in quanto il segnale è tipicamente proporzionale al volume cellulare. Alte frequenze, invece, vengono utilizzare per avere informazioni sulla conduttività della membrana cellulare.

Nel seguente report viene considerata una configurazione ad elettrodi coplanari rappresentata in fig. 1.

Tramite la misura di corrente differenziale è possibile stimare alcune proprietà della cellula che passa nel canale. In particolare, al passaggio delle cellula, si misura un segnale con una forma d'onda di tipo gaussiana bipolare.

Tramite l'ampiezza di picco è possibile stimare il diametro elettrico. Il segnale è proporzionale al volume della cellula per cui il diametro sarà legato all'ampiezza del segnale (a) come:

$$D = Ga^{1/3} \quad (1)$$

Dove G è un guadagno che risente delle proprietà elettriche specifiche del dispositivo utilizzato.

2.1 Gaussiana bipolare

La gaussiana bipolare è una forma d'onda caratteristica composta da due gaussiane identificate dalla generica equazione:

$$g(t) = a [e^{g_+(t)} - e^{g_-(t)}] \quad (2)$$

Ovvero, considerata un'ampiezza di riferimento a (i.e. il valore massimo di picco), la forma d'onda complessiva è data dalla somma di due gaussiane nel tempo di cui la seconda ribaltata.

La distanza picco-picco è pari a δ e si introduce un parametro di centratura t_c .

Le due gaussiane condividono la medesima deviazione standard σ e sono identificate dalle eq. (3) e eq. (4).

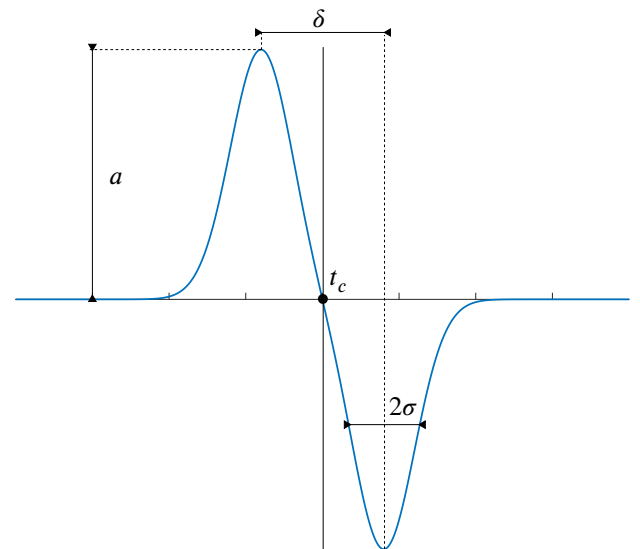


FIG. 2: Forma d'onda di tipo gaussiana bipolare identificata dall'ampiezza a , punto centrare t_c , larghezza a metà altezza 2σ e distanza picco-picco δ .

$$g_+(t) = \frac{-(t - (t_c + (\delta/2)))^2}{2\sigma^2} \quad (3)$$

$$g_-(t) = \frac{-(t - (t_c - (\delta/2)))^2}{2\sigma^2} \quad (4)$$

2.2 Procedura di fitting

Partendo dai dati sperimentali è necessario introdurre una procedura di fitting numerico per identificare la gaussiana, e quindi i suoi quattro parametri descrittivi, tale da rappresentare il segnale analizzato [1].

Tale procedura di fitting viene implementata secondo un algoritmo di ottimizzazione. Ovvero, si cerca di ridurre la differenza tra il dato misurato $[d]_i$ e il template di fitting $g(t_i)$ allo stesso istante temporale.

Definita la funzione di errore come tale differenza:

$$\underline{e} = [d]_i - g_i(t_i, a, t_c, \delta, \sigma) \quad (5)$$

Si cerca di minimizzare la funzione obiettivo definita proprio come la norma dell'errore:

$$E(a, t_c, \delta, \sigma) = \frac{1}{2} \sum_i \|d_i - g(t_i, a, t_c, \delta, \sigma)\|^2 \quad (6)$$

2.3 Accuratezza

I dispositivi microfluidici possono essere più o meno accurati. Questo è legato alla configurazione degli elettrodi, alla geometria, al mezzo di sospensione e velocità del flusso [2].

In particolare, per il caso di riferimento si osserva una dipendenza del segnale dalla posizione all'interno del canale. Nonostante l'ampiezza del picco della gaussiana sia correlata con il diametro cellulare si osserva una dipendenza anche dalla posizione.

Ovvero, a parità di diametro, una cellula passante vicino agli elettrodi darà un picco di ampiezza maggiore alla stessa cellula passante ad una distanza maggiore [3].

Nonostante nella configurazione utilizzata non si può avere una misura diretta della posizione all'interno del canale è possibile ottenerne una stima effettuando una compensazione del segnale

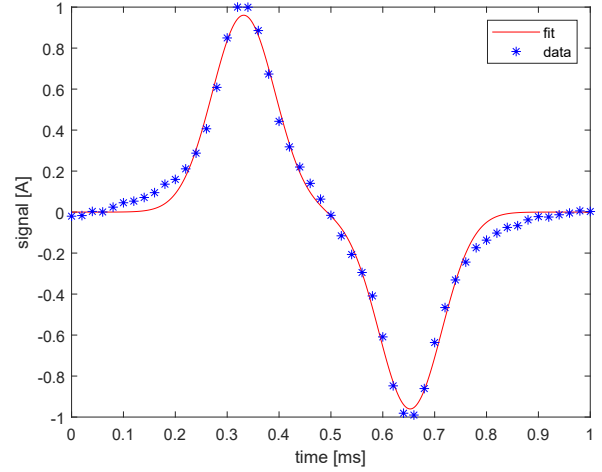


FIG. 3: Grafico dell'andamento temporale del segnale n. 100. In blu i campioni e in rosso il segnale fittato con una gaussiana bipolare

tenendo conto della correlazione tra la posizione e lo shape parameters.

In particolare, risulta valida la relazione tra il parametro di forma shape e il diametro normalizzato che segue una legge lineare:

$$\frac{D}{d} = c_1 + c_2 \left(\frac{\sigma}{\delta} \right) \quad (7)$$

Dove:

$$\frac{\sigma}{\delta} = \text{shape} \quad (8)$$

Si possono quindi usare tali coefficienti della retta, calcolati per ogni segnale, per correggere il diametro elettrico rimuovendo l'effetto di posizione [4].

2.4 Dataset di riferimento

Il dataset di riferimento è un insieme di dati grezzi ottenuti da misurazioni di citometria ad impedenza su un campione di biglie di test.

Sono presenti i segnali di tre famiglie di biglie di test con diametro nominale di 5.2, 6 e 7 μm .

Per il dispositivo in questione il guadagno è pari a 10.5 $\mu\text{m} / \text{A}^{1/3}$ e la distanza inter-elettrodo è pari a $L = 40 \mu\text{m}$.

I segnali sono campionati con una frequenza $f_s = 115 \text{ kHz}$ con un totale di oltre 50.000 segnali.

3 Risultati

Inizialmente la procedura viene applicata su un dataset ristretto considerando solamente 400 seg-

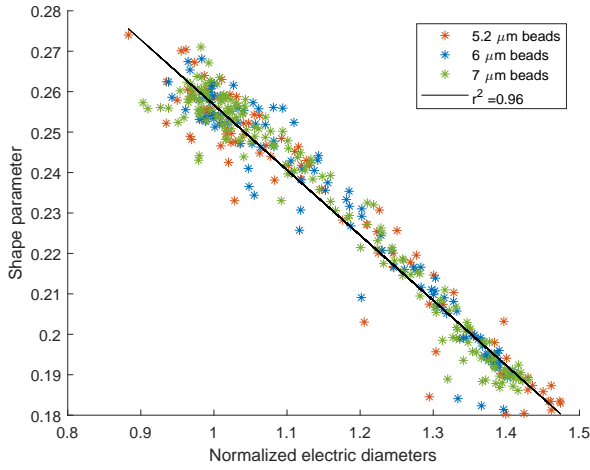


FIG. 4: Andamento dello shape parameter rispetto il diametro normalizzato. Sono raffigurati i campioni delle tre famiglie distinte con colori differenti e la retta ottenuta con una procedura di fitting lineare.

nali. Vengono selezionati quindi i segnali di indice da 400 a 800.

Successivamente la procedura verrà applicata anche su un riferimento più esteso.

3.1 Fitting numerico

A partire dal singolo segnale all'interno dell'intero dataset è possibile fittare la gaussiana bipolare andando a stimarne i quattro coefficienti descrittivi.

Considerando il template in eq. (2) è possibile utilizzare il comando Matlab `fit()` fornendo il template stesso, i dati e i valori iniziali per la stima ai minimi quadrati.

Inoltre, essendo il segnale molto piccolo $\propto 10^{-6}$ è utile normalizzare il segnale di ampiezza e scalare l'asse dei tempi riportandolo in secondi nel range $[0; \frac{1}{f_s}]$ con un numero di campioni equivalente, dove f_s è la frequenza di campionamento il cui inverso indica l'ultimo campione temporale.

A questo deve seguire poi necessariamente una riscalatura nel dominio originario per poter confrontare i segnali ed effettuare la procedura di compensazione correttamente.

Un esempio di fitting è presente in fig. 3 ed è evidente come la gaussiana bipolare sia descrittiva del segnale misurato.

3.2 Compensazione

Per la procedura di compensazione è necessario stimare i coefficienti della retta derivante dal fitting

lineare.

Tale fitting deve essere fatto nello spazio $[\frac{\delta}{\sigma}; \frac{D}{d}]$ ed è quindi necessario normalizzare i diametri elettrici rispetto i singoli diametri nominali delle tre famiglie.

Bensì in questo dataset non sia possibile identificare a priori, con una procedura automatica, la famiglia di appartenenza della misura si può sfruttare lo scatter plot del diametro elettrico vs shape parameters per distinguere visivamente tre famiglie (fig. 6a). Quindi, mediante la funzione `inpolygon()` è possibile selezionare manualmente i valori corrispondenti alle tre famiglie, separarle e quindi normalizzarle per il rispettivo diametro nominale.

Quindi è possibile calcolare la retta che meglio approssima l'andamento dello shape parameter in funzione del diametro elettrico normalizzato (fig. 4). Dunque è possibile estrarre i coefficienti:

- $c1 = 2.59$
- $c2 = -6.20$

All'interno dell'ambiente Matlab è possibile utilizzare il comando `fitlm()` così da ottenere direttamente i due coefficienti della retta e anche una stima del coefficiente di determinazione $r^2 = 0.96$.

Tramite tali coefficienti è possibile correggere i valori tramite l'eq. (7).

Si ottengono quindi le famiglie separate sia nello scatter plot dello shape parameter che della velocità, in fig. 6 e fig. 7.

3.3 Distribuzioni

Chiarmente tale procedura, separando le famiglie, permette il calcolo della loro distribuzione.

Analizzando la distribuzione dei segnali, prima della compensazione, è evidente come non sia

	Deviazione standard	Media	Coefficiente di variazione	Diametro nominale
# 1	0.187	5.196	0.035	5.2 μm
# 2	0.184	6.016	0.031	6 μm
# 3	0.205	7.005	0.029	7 μm

TAB. 1: Coefficienti della distribuzione gaussiana per le tre famiglie ottenuti dal fitting dei valori dell'istogramma

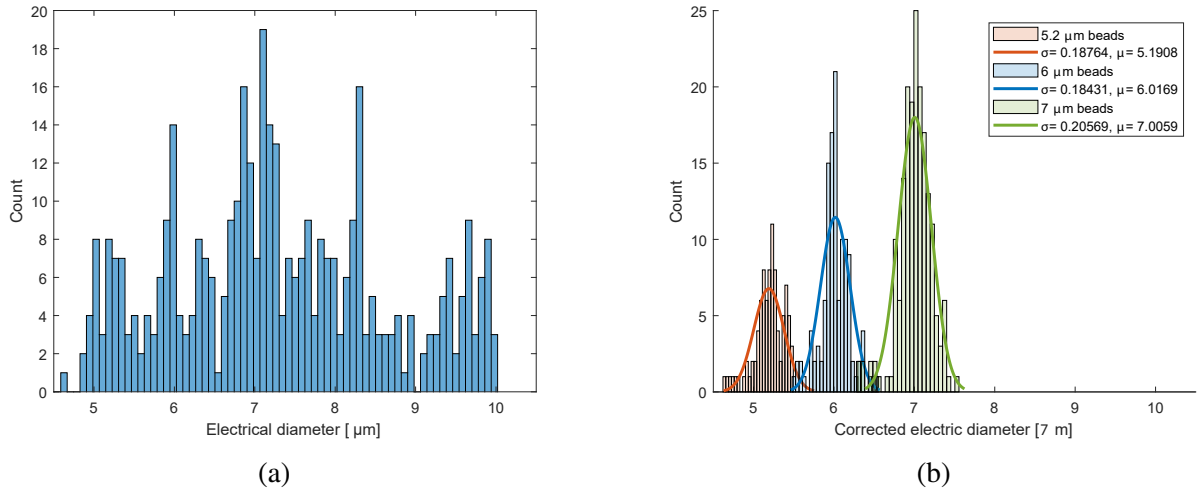


FIG. 5: Distribuzione dei diametri elettrici rappresentata con un istogramma prima della correzione (a) e dopo la procedura di compensazione (b). Sono evidenti le tre famiglie distinte dopo la procedura di compensazione.

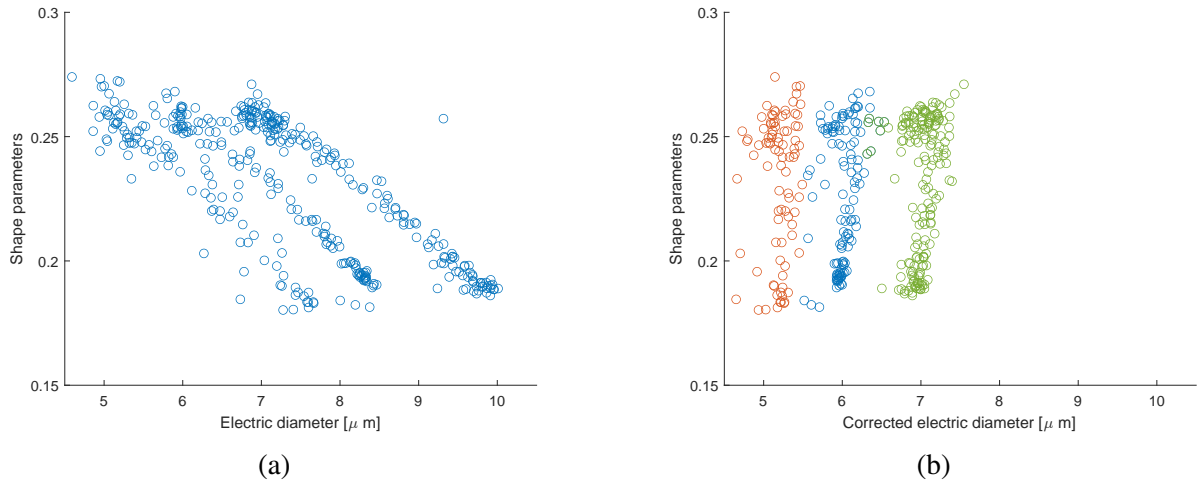


FIG. 6: Scatter plot dello shape parameter rispetto al diametro elettrico prima (a) e dopo (b) la compensazione. La procedura di compensazione porta ad una netta distinzione delle tre famiglie. Si passa da tre bande diagonali, dove è impossibile stabilire il diametro senza portare in conto l'incertezza del contributo posizionale, a tre bande verticali che chiaramente separano le tre differenti famiglie.

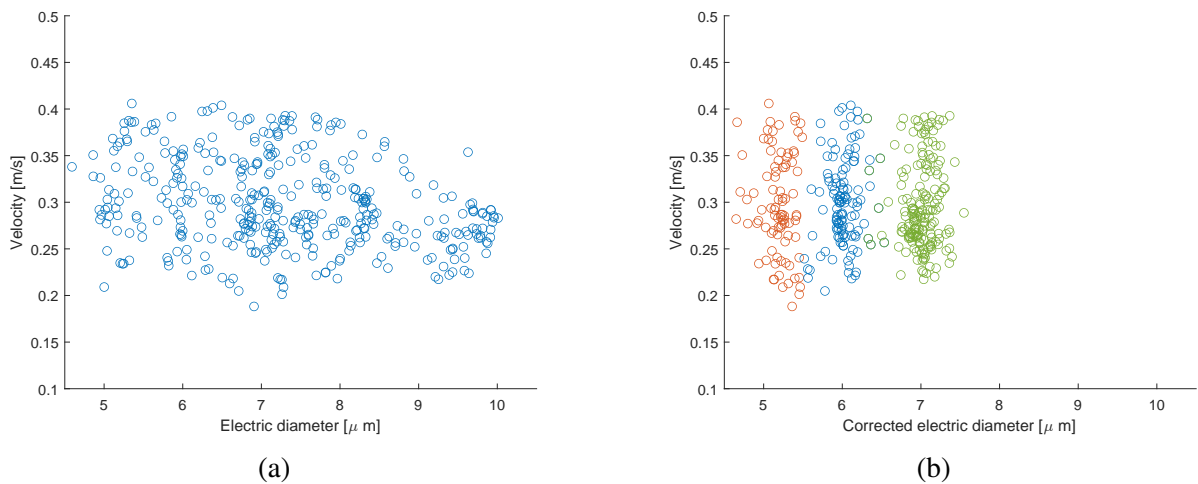


FIG. 7: Scatter plot della velocità rispetto al diametro elettrico prima (a) e dopo (b) la compensazione. Si vede come la procedura di compensazione porta ad una netta distinzione delle tre famiglie che nel caso pre-compensazione non era presente.

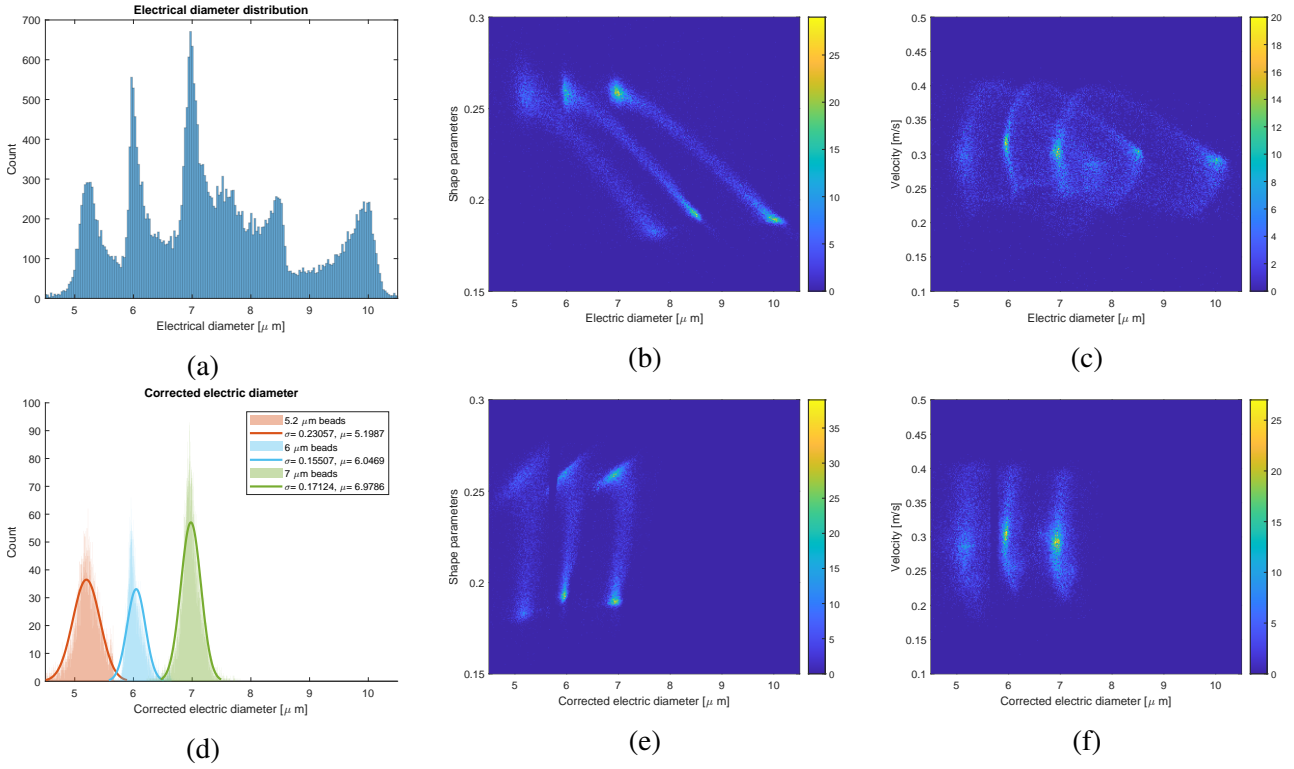


FIG. 8: Analisi su 35000 campioni. Viene nuovamente applicata la procedura di test ma su un numero di segnali maggiori la cui distribuzione è presentata tramite istogramma (a). Considerando molti segnali risultano statisticamente evidenti i picchi relativi alle tre famiglie ma non è possibile distinguere nettamente i segnali. La distinzione invece appare netta ed evidente dopo la compensazione (d). Sono riportati anche i plot di densità per lo shape parameters rispetto al diametro elettrico prima (b) e dopo (e) la compensazione e per la velocità rispetto al diametro elettrico prima (c) e dopo (f) la compensazione. Dai diagrammi di densità risulta evidente la netta separazione che viene seguita dalla procedura di compensazione.

possibile identificare alcun picco nel grafico dell'istogramma in fig. 5a. Questo grafico è pienamente rappresentativo di come i segnali siano effettivamente mischiati a causa del contributo di posizione.

Dopo la compensazione, invece, è chiaramente possibile distinguere le diverse famiglie come è evidente dai tre picchi nettamente distinguibili in fig. 5.

Vengono riportate quindi le distribuzioni delle tre famiglie ciascuna con un andamento a gaussiana in cui i parametri sono presenti in Tab. 1. Risulta evidente come la maggioranza delle biglie rientrano nella terza famiglia, a diametro maggiore, ma tale valore rispecchia anche una maggiore dispersività. Le due famiglie a dietro più piccolo presentano invece una dispersione confrontabile.

3.4 Estensione del dominio di interesse

La stessa procedura può essere applicata anche su una quantità di segnali notevolmente maggiore. Vengono quindi considerati, nello stesso dataset di

riferimento, i segnali di indice a 200 a 35200.

La procedura può essere applicata allo stesso modo ma gli scatter plot vengono sostituiti da plot di densità che permettono di osservare meglio la

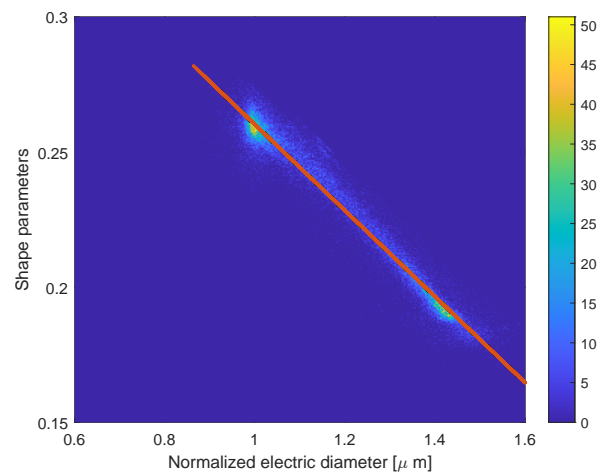


FIG. 9: Grafico di densità della distribuzione dei segnali del parametro di forma rispetto al diametro normalizzato. Sono presenti i dati di tutte e tre le famiglie con la retta interpolante unica.

distribuzione dei valori (fig. 8).

Nuovamente si ottengono i due parametri di fitting $c_1 = 2.63$, $c_2 = -6.28$ tramite i quali applicare la procedura di compensazione.

In modo analogo segue l'istogramma dove si separano le tre famiglie (fig. 8d). In questo caso la maggioranza delle biglie rientrano comunque nella famiglia con diametro maggiore ma la famiglia con diametro più piccolo presenta una maggiore dispersione.

4 Conclusioni

La procedura di compensazione permette di separare le differenti famiglie rispetto al diametro elettrico e risulta affidabile sia su pochi segnali che su campioni più vasti.

Tale procedura richiede di conoscere preventivamente i diametri delle biglie.

Nel caso in cui i singoli segnali non siano associabili al relativo diametro nominale con certezza viene richiesta una procedura manuale di identificazione che, per sua natura, indurrà un certo errore nella classificazione e nel fitting lineare, quindi sulla procedura di compensazione.

Tale errore si riflette principalmente in una dispersione nelle distribuzioni delle biglie. La deviazione standard delle distribuzioni delle famiglie rimane sotto lo 0.23 e macroscopicamente rimangono fortemente distinguibili le tre famiglie e le loro distribuzioni.

Chiaramente, a causa dell'errore numerico presente e della grande quantità di outliers, non è facile classificare il singolo segnale quanto classificare la distribuzione macroscopica.

Disponibilità dei dati

Il materiale è disponibile alla repository online del progetto: <https://github.com/mastroalex/curve-fitting>

Riferimenti

- [1] Federica Caselli and Paolo Bisegna. “A Simple and Robust Event-Detection Algorithm for Single-Cell Impedance Cytometry”. In: *IEEE Transactions on Biomedical Engineering* 63.2 (Feb. 2016), pp. 415–422. ISSN: 0018-9294, 1558-2531. DOI: 10 . 1109 / TBME . 2015 . 2462292. URL: <http://ieeexplore.ieee.org/document/7169531/> (visited on 06/04/2022).
- [2] Tao Sun and Hywel Morgan. “Single-cell microfluidic impedance cytometry: a review”. en. In: *Microfluidics and Nanofluidics* 8.4 (Apr. 2010), pp. 423–443. ISSN: 1613-4982, 1613-4990. DOI: 10 . 1007 / s10404 - 010 - 0580 - 9. URL: <http://link.springer.com/10.1007/s10404-010-0580-9> (visited on 06/04/2022).
- [3] Daniel Spencer et al. “High accuracy particle analysis using sheathless microfluidic impedance cytometry”. en. In: *Lab on a Chip* 16.13 (2016), pp. 2467–2473. ISSN: 1473-0197, 1473-0189. DOI: 10 . 1039 / C6LC00339G. URL: <http://xlink.rsc.org/?DOI=C6LC00339G> (visited on 06/04/2022).
- [4] Vito Errico et al. “Mitigating positional dependence in coplanar electrode Coulter-type microfluidic devices”. en. In: *Sensors and Actuators B: Chemical* 247 (Aug. 2017), pp. 580–586. ISSN: 09254005. DOI: 10 . 1016 / j . snb . 2017 . 03 . 035. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0925400517304525> (visited on 03/19/2022).