

# UC Irvine

## UC Irvine Previously Published Works

### Title

Phasor-based image segmentation: machine learning clustering techniques.

### Permalink

<https://escholarship.org/uc/item/1qj2074g>

### Journal

Biomedical Optics Express, 12(6)

### ISSN

2156-7085

### Authors

Vallmitjana, Alex  
Torrado, Belén  
Gratton, Enrico

### Publication Date

2021-06-01

### DOI

10.1364/boe.422766

Peer reviewed



# Phasor-based image segmentation: machine learning clustering techniques

ALEX VALLMITJANA, BELÉN TORRADO, AND ENRICO GRATTON\* 

Laboratory for Fluorescence Dynamics, Biomedical Engineering, University of California, Irvine, CA 92697, USA

\*egratton22@gmail.com

**Abstract:** The phasor approach is a well-established method for data visualization and image analysis in spectral and lifetime fluorescence microscopy. Nevertheless, it is typically applied in a user-dependent manner by manually selecting regions of interest on the phasor space to find distinct regions in the fluorescence images. In this paper we present our work on using machine learning clustering techniques to establish an unsupervised and automatic method that can be used for identifying populations of fluorescent species in spectral and lifetime imaging. We demonstrate our method using both synthetic data, created by sampling photon arrival times and plotting the distributions on the phasor plot, and real live cells samples, by staining cellular organelles with a selection of commercial probes.

© 2021 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

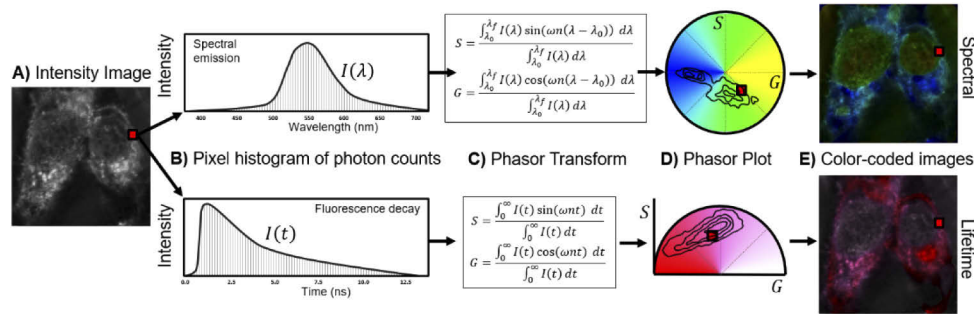
## 1. Introduction

Phasor plots applied to fluorescence arise from the frequency domain measurements of fluorescence lifetime in the 1970s. The earliest hints to a two-dimensional polar decomposition of lifetime measurements can be seen in the mathematical formalization developed by Weber in 1981 [1] and soon after the first graphical representation in the work by Jameson et.al. [2]. Some 20 years later we see a resurgence in the use of this graphical method [3,4] and the actual term is coined in 2008 [5]. The first realization of the analogy between lifetime and spectral measurements and the application of the phasor transform to spectral data instead of lifetime measurements can be found in the works by Fereidouni and Gerritsen in 2012 [6].

The phasor transform applied to the photon histogram computes two quantities, namely  $S$  and  $G$ , which are mapped to a two-dimensional space, the phasor space. The two quantities are the normalized Fourier sine and cosine transforms at a particular frequency of the photon histogram curve [5]. Since these quantities are normalized to the total number of photon counts, they are independent of the intensity and carry information only on the shape and position of such a curve. Additionally, due to the orthonormality of the two basis functions, sine and cosine, there are a series of useful vectorial algebra properties that can be used in this phasor space. Here arises one of the main reasons for using the phasor approach; if there are two or more species contributing to a photon histogram curve, the phasor position of this combination is found at a geometrical linear combination of the phasor position of the pure species, i.e. they follow the vector addition rule [7]. This property alone has allowed many applications such as metabolism characterization [8] or FRET quantification [9,10]. Another very important characteristic of the phasor transform is that it is a fit-free technique which does not require any a priori knowledge on the nature of the photon distribution curve one has acquired.

Whichever the application, lifetime or spectra, the phasor plot consists in transforming an acquired distribution of photon measurements to a new space that enhances certain properties and facilitates handling and quantification. In general, the measurement device provides a sequence of photon detection counts together with their arrival times (for the lifetime case) or wavelengths (for the spectral case). This data can then be plotted as a histogram of photon counts, as a function of time or wavelength. In the first case, it is the lifetime decay curve, that is, the number

of photons that are detected at a particular time after the latest excitation pulse or the number of photons that arrive at a particular phase shift with respect to the modulated excitation source. In the second case, the histogram depicts the number of photons collected in a particular spectral band i.e. the emission spectra. See Fig. 1. Note that, in general, these are two completely independent magnitudes; two fluorophores can have the same spectrum but different lifetimes (for example by quenching) or different spectra but the same lifetime.



**Fig. 1.** Spectral and lifetime phasor transform analogy. A) The intensity image depicts the number of photons collected at each pixel. If we can in addition measure the wavelength and/or the lifetime we add extra dimensions to our data. At each pixel, symbolized by the red square, we can construct a photon histogram distribution (B), which by means of the phasor transform (C), maps to a position on the phasor plot (D). The phasor space representation of the image can be used for instance to color-code the original image (E). The two transforms are analogous to one another, top row corresponds to spectral and bottom row corresponds to lifetime.

Since measurements usually involve whole images, each spatial location in the sample (each pixel) has its own photon histogram and therefore maps to a different location on the phasor space. This results in a reciprocal relationship (meaning one-to-one) between the pixels in the image space and the points in the phasor space. In the phasor space, the points appear scattered about some distribution, which may present clustered populations, corresponding to regions of distinct lifetime or spectra in the sample. The most straight-forward approach to connect the two spaces is to assign a color surface to the underlying phasor space which then maps to individual pixels in the image [11]. This allows using color to visualize the relationship between the image space and the phasor space but it is not image segmentation per se, it is simply a qualitative approach in which one color-codes the intensity image (see Fig. 1(E)). The additional segmentation step requires discretizing the image space into  $N$  groups by assigning a label to each pixel to associate it into one of the  $N$  groups. In this paper we concentrate on exactly this problem, how to segment the content of the image space, based on identifying clusters in the phasor space.

Up until now, identifying these regions in the phasor space in order to segment the corresponding regions on the image space has been done manually and very little attempts have been made towards an automated, unsupervised and reproducible segmentation method. In the current biological applications of phasor plots, the user manually selects points or regions on the phasor plot and then color-codes them back on the original image space [12,13]. Other techniques characterize concentration of species by means of trajectories on the phasor plot, by either drawing lines or areas on the phasor space and mapping color gradients to depict transitions which symbolize the variations in concentration of particular fluorescent species [14,15].

Enter our clustering problem; there is a need for a unified method that exploits the current available algorithms in machine learning to automatically identify and divide phasor plot data into regions in order to assign each data point to a particular cluster or assign a probability of belonging to a particular cluster.

Pattern recognition and machine learning has previously been applied to lifetime data but very little on phasor analysis. One can find examples of feature extraction and dimensionality reduction from the decay curves to compose a so-called lifetime image or for classification and clustering, but again based on the time domain decay curve [16,17]. A recent review on the different uses of machine learning in lifetime imaging can be found in the paper by Mannam et. al. [18], where one can corroborate what we previously mentioned, that the use of machine learning to the fit-free phasor analysis is currently very limited. Automated characterization of shape and size of phasor distributions has previously been done by extracting a set of features that are then used to train a classifier in order to distinguish between classes of cells (e.g. healthy vs unhealthy) based on the phasor transform [19,20]. Nevertheless, blind clustering of phasor plot data, to the best of our knowledge, has only been attempted using K-means based algorithms [21–23]. K-means, as we will show, is suboptimal because it tends to divide the space into Voronoi cells around the means and therefore is only useful in clearly separable cases that do not overlap.

Clustering is a family of machine learning and methods that are used in situations in which one disposes of a collection of N-dimensional points but has no labels assigned to them. The goal in this type of problems is to discover groups within the data that have similar characteristics in terms of the N-dimensional coordinates, usually named the set of features for each point. The ultimate output of such algorithms is a single label assigned to each of the points although in some situations one can obtain, for each point, a probability value of belonging to each group. There exist several broad families of clustering algorithms, each of which with many sub-families within them, with their own advantages and disadvantages. In an attempt to explore the range of possible clustering methods we simulated lifetime phasor data and compared the performance at clustering this data with a subset of some of the most established (k-means, hierarchical clustering, dbscan etc.).

In this paper we propose using Gaussian Mixture Models (GMM) [24]. The initial intuition behind using GMM comes from the experience of observing phasor plots in a variety of experimental conditions and realizing that the populations tend to spread in a normal fashion, but there is a rational justification on top of that, which is that phasor distributions do in fact have a normal component in the noise; the presence of noise in the measurement can have a variety of sources, but under the phasor transform this noise tends to translate into a normally distributed spreading of the phasor coordinates [25]. In this work we show the results using simulations and real data of lifetime fluorescence measurements but the same set of clustering techniques can be equally applied on data in the spectral phasor transform.

## 2. Materials and methods

### 2.1. Simulated phasor data

There are many benchmark sets available for testing clustering algorithms, but they tend to include extreme situations, e.g. the smiley face or the concentric circles [26], which are cases specifically designed to force some of the classical methods to fail, but have no relevance for us in the lifetime/spectral phasor transform data. For this reason, we have created our own set of synthetic test data attempting to emulate the different situations in which the clustering problem in the phasor space may become hard. These are: situations in which clusters have different shape (maybe in FRET or STED experiments where the data is pulled), situations in which the clusters have different size or density (noise, low counts or presence of autofluorescence) and situations in which the clusters have a different total number of points (experiments where different probes have different relative brightness or simply are tagging structures of different size).

We wrote a script that draws photon arrival times out of a distribution obtained as the convolution of an exponential decay by an impulse response function. The fluorescence decay of a single fluorophore is modelled as a single exponential decay with a single parameter being the characteristic lifetime. The impulse response function is modelled as a normal distribution

of standard deviation  $1\text{ns}$ . Once the photon arrival histogram curve is calculated from the convolution of the exponential by the gaussian, the photon arrival times are drawn from that distribution using a Montecarlo approach. This is done by dropping random coordinates in the cartesian plane and taking those that fall below the desired curve.

The simulations were done assuming a pulsed stimulation frequency of  $80\text{MHz}$  and in every case we generated pixels belonging to one of three different lifetime populations. The first cluster is composed of pixels containing photons with a lifetime of  $2.5\text{ns}$ , the second cluster a decay lifetime of  $3.5\text{ns}$  and the third cluster is composed of the sum of two distributions;  $80\%$  of the photons belong to a distribution with a lifetime of  $3.5\text{ns}$  and the remaining  $20\%$  belong to a distribution with a lifetime of  $0.5\text{ns}$ . This allowed us to create three clusters forming a triangle with which we can tune the size and shape applying additional constraints (see the results section for details on the ground truth generation).

## 2.2. Gaussian mixture models

Gaussian mixture models [24] is a clustering technique that is based on fitting a linear combination of Gaussian distributions to the data points. The probability density of the model is therefore a sum of terms with the number of terms,  $K$ , being the number of clusters:  $p(x) = \sum_{i=1}^K f_i N(\mu_i, \Sigma_i)$ . Each cluster is defined by its fraction  $f_i$  and the multivariate normal distribution  $N(\mu_i, \Sigma_i)$  with mean coordinates  $\mu_i$  and covariance matrix  $\Sigma_i$  which characterizes the data belonging to the cluster. The typical implementation for computing the set of parameters uses the Expectation-Maximization algorithm [27]. It is an iterative procedure in which one starts with some values for the parameters and, at each iteration, the data points are evaluated under the current model to obtain the likelihood of belonging to each cluster (expectation). Then this likelihood is used to weigh the points to obtain the mean position  $\bar{\mu}_i$  of each cluster –and subsequently its covariance matrix- towards the next iteration (maximization). This method always converges to some solution and the only assumption that one must make is that the data can be modeled by a (quasi-) normal distribution.

We have omitted an important detail which is how to estimate the number of clusters  $K$  with which to model the data. This is, in general, one of the main drawbacks of using GMM, together with many of the other clustering methods, and there are several approaches around this issue [28–31]. In this paper we intentionally omit this step because we assume that we always know beforehand the number of components that exist in our data. This is in general true because the components are fluorescent probes with which we have labelled our samples or because we can make certain assumptions on the nature of the sample. Additionally, we have to assume that our probes are spatially resolved, *i.e.* that the different species in the sample do not share a pixel, otherwise the distributions will lose their normality.

## 2.3. Cell culture and staining

NIH3T3 (ATCC) cells were cultured in Dulbecco's Modified Eagle Medium (DMEM, Gibco) supplemented with  $10\%$  Fetal Bovine Serum (FBS, Sigma-Aldrich), streptomycin ( $100\mu\text{g/ml}$ ) and penicillin ( $100\text{U/ml}$ ), at  $37^\circ\text{C}$  in a  $5\%$   $\text{CO}_2$  incubator. Cells were seeded on fibronectin (Sigma-Aldrich) covered glass bottom dishes (Mattek). Multiple staining of the cells was done sequentially. Dilutions of each dye were prepared in DEMEM phenol-free, high glucose, HEPES medium (Gibco).

In the final experiment shown in the results section, we used a total of six dyes (Table 1). Cells were incubated with LysoTracker Deep Red ( $1\mu\text{M}$ , Invitrogen), ER Tracker Red (BODIPY TR Glibenclamide;  $500\text{nM}$ , Invitrogen), BODIPY FL C5-Ceramide ( $500\text{nM}$ , Invitrogen), MitoTracker Green FM ( $167\mu\text{M}$ , Invitrogen) and UltraPure Ethidium Bromide ( $10\mu\text{g/mL}$ , Invitrogen), for  $40\text{min}$  at  $37^\circ\text{C}$ . Cells were then washed three times with  $1\text{X}$  Dulbecco's Phosphate Buffer Solution

(DPBS). In the final step, cells were incubated MemBrite Fix Cell Surface Staining Kit 640/660 (Biotium) following the manufacturer's instructions.

**Table 1. Commercially available dyes used to test the system.** <sup>a</sup>

|  | Lifetime(ns) | Spectrum(nm)    |                | Organelle    | Manufacturer Catalogue Number |
|--|--------------|-----------------|----------------|--------------|-------------------------------|
|  |              | $\lambda_{exc}$ | $\lambda_{em}$ |              |                               |
| LysoTracker deep Red                     | 3,0          | 647             | 668            | Lysosomes    | Invitrogen L12492             |
| MembriteRed Fix 640/660                  | 2,4          | 640             | 660            | Membrane     | Biotium 30097-T               |
| ER tracker red (BODIPY TR Glibenclamide) | 5,0          | 587             | 615            | ER           | Invitrogen E34250             |
| UltraPure Ethidium Bromide               | 12,0         | 521             | 605            | Nucleus      | Invitrogen 15585011           |
| MitoTracker Green FM                     | 0,7          | 492             | 514            | Mitochondria | Invitrogen M7514              |
| BODIPY FL C5                             | 4,8          | 502             | 510            | Golgi        | Invitrogen D7540              |

<sup>a</sup>Live cells were stained individually with each dye and its lifetime was measured using the phasor approach. The reported peak excitation ( $\lambda_{exc}$ ) and emission ( $\lambda_{em}$ ) is provided by the manufacturer.

Preliminary rounds of imaging were done with cells incubated with a single dye in order to measure its lifetime in vivo (list of final used dyes in Table 1, and an extended version with the complete list of tested dyes in supplementary table1). Subsequent rounds of incubation were performed each gradually increasing the number of dyes in the sample. This allowed to adjust the incubation time and dye concentration such that the fluorescence of each of the probes was in the same order of magnitude.

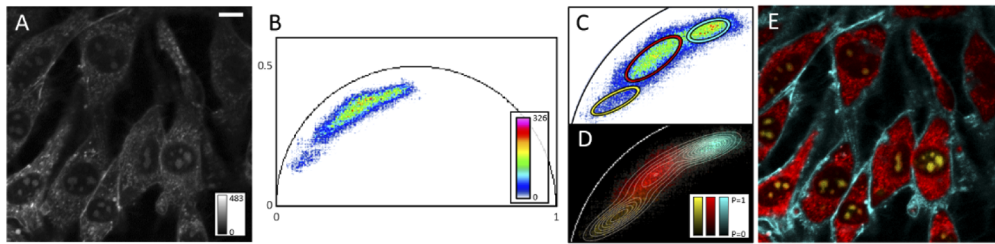
#### 2.4. Imaging

Images were acquired using a two-channel ISS-ALBA5 confocal microscope equipped with a white laser source and acusto-optic tunable filter (NKT SuperK EXTREME, SuperK SELECT), avalanche photo-diodes (Excelitas Technologies) and an ISS A320 FastFLIM acquisition unit for lifetime measurement. The repetition frequency of the laser is of 78MHz. We used this system because it is a highly tunable confocal microscope but our experiments can be reproduced with any microscope with lifetime capabilities.

The probes were excited with three laser lines selected from the white laser (490, 560 and 640nm) and the emission of the six probes was collected in three spectral bands using band-pass filters 520/35, 605/55 and 679/41. The images were acquired in three sequential rounds of imaging in order to minimize the bleed through of the spectral channels, but in principle with a three-channel detector, the experiment could be reproduced in a single shot. The example shown in Fig. 2 was taken in a single channel, collecting the emission of three probes using the 679/41 filter and exciting simultaneously with two laser lines (560nm and 640nm).

There exists somewhat of a misconception around the concept of a lifetime image. In each pixel of an image, there are always many fluorescence processes contributing to its intensity, never a single species with a single lifetime decay. When taking lifetime measurements, there are many approaches to dealing with this; usually involving reporting an ensemble or average value. If there is a clear predominant fluorescent species, the value reported will be close to that of the predominant species, but that is not necessarily the case. For this reason, the lifetimes for the individual probes we report (Table 1) are phasor-based and they are the mean two values, tau modulation and tau phase [32]. These two values are the corresponding lifetimes of the two points on the universal circle where one projects the data point that is inside the universal circle. One projection is finding the point on the universal circle that has the same phase as the data point and the other is finding the one that has the same modulation.





**Fig. 2.** Example of phasor space GMM clustering. A) Intensity image of a live cells sample, scale bar is 10µm. B) Lifetime phasor plot of the image. Gaussian mixture model for 3 clusters is run on the phasor coordinates and for each of them the posterior probabilities of belonging to each cluster are obtained. C) The probabilities are used to divide the data with hard borders. D) The probabilities ( $P$ ) are used to assign a color as a linear combination of the three chosen colors: yellow, red and cyan. E) The image is colored using the colors drawn from the probabilities.

## 2.5. Image processing

Image processing scripts were written in MATLAB. As a brief itemization of the processing pipeline; lifetime files were decoded, individual photons assigned to each pixel, pixel arrival time photon distributions were phasor-transformed and GMM was applied to phasor coordinates to obtain the pixel probability of belonging to each cluster. We wrote our own GMM routine fitting function using the expectation-maximization algorithm, although MATLAB already implements its own version *fitgmdist.m*. All functions written for the simulations and for the processing of the live cell data (also included) are available in a public repository ([Dataset 1](#) [33]).

One of the advantages of using the phasor plot is that one can obtain the lifetime characterization at each pixel with a relatively low number of counts. In our case the images contained at most 500 photons per pixel, but the bulk of the points in the images contained much fewer than 100. For the clustering step we set a threshold of 40 photons. Although in terms of single photon counting 40 may seem a lot, it is actually very little in terms of modeling a lifetime decay curve in the time domain. For the phasor approach 40 is probably double the minimum amount [25]. The thresholding step was done for two main reasons: first to reduce the computational load by feeding less points to the model, but most importantly to reduce the spreading in the phasor space and therefore ease the performance of the GMM fitting. The outcome is that the GMM routine only uses the pixels with more than 40 photon counts to find the best mixture of Gaussians, but once the model is estimated it is used to obtain the likelihood of all the pixels in the image, including those with lower than 40 counts. This provides a probability estimate of each pixel to belong to each cluster and is the quantification of the segmentation.

Figure 2 shows the basis of the method, in which analysis is performed in an example image of a large field of view of a sample tagged with the six probes in Table 1. The particular example (Fig. 2(A)) shows the emission in the far red [660 700]nm, exciting with two laser lines (642nm and 561nm). The expected emission collected in these conditions is mainly from three of the probes (LysoTracker, Membrite and Ethidium Bromide), the rest are spectrally cropped by the filter. The phasor plot is shown for the pixels above 40 counts (Fig. 2(B)). The standard way of depicting a phasor plot is as a 2D histogram where color codes for density, that is, each phasor-space bin is colored according to the number of pixels of the original image that fall in the particular bin (in this case 256 bins to the phasor space unity distance). The other way one can depict the phasor plot is as a scatter plot, i.e. a cloud of points, which although it carries the real individual coordinate information of all pixels, the size of the points rapidly saturates the

density and is therefore never used. This list of pixel phasor coordinates is the input data that the GMM algorithm deals to assign a probability of each pixel to belong to each cluster.

With this list of probabilities one option is to simply cut hard edges, i.e. at each coordinate choose the cluster with the higher probability and assign that cluster to the pixel. Figure 2(C) shows the ellipses that contain 68% of the points that belong to each cluster using this hard-edge method. The alternative is to plot the scatter plot of the points by color-coding the points using the probabilities. Figure 2(D) shows this approach, overlaid with a set of 10 contour lines of the resulting 3 component GMM fit to the data. In this method, a color is chosen for each cluster, and the linear combination of the RGB coordinates of the colors is computed using the probabilities as coefficients. This results in a color assignment for each pixel, and one can produce a color image matching the original. The product of the normalized version of the original intensity image with the probability obtained color image is shown in Fig. 2(E).

### 3. Results

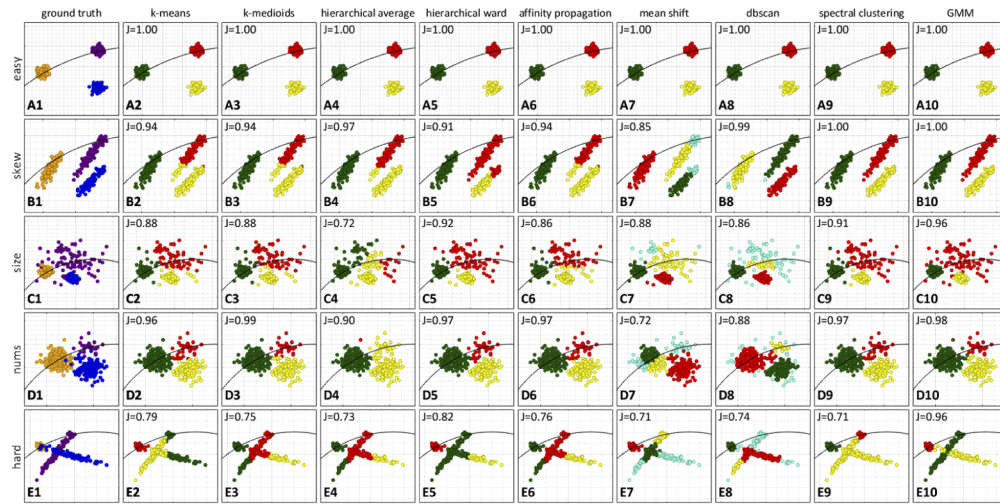
#### 3.1. Simulations

A total of 5 data sets were simulated to attempt to cover different possible situations, namely: easy case, skewed distribution case, different sizes case, different number of elements case and hard case. The first easy case consists of three compact distributions, 104 photons per pixel, 300 pixels, and exactly one third obtained from each of the three characteristic lifetimes described above. The second case, skewed case, uses the same distributions but adds a number of photons with uniformly sampled lifetimes, the number obtained from a normal distribution of mean 103 and standard deviation 800. This simulates the presence of unmodulated light which pulls the data to the origin of the phasor plot. In the third case, size, the pixels from the first distribution contain only 100 photons each, the other two distributions 3000 photons each, so that the first is much more spread on the phasor plot. The fourth case, number, is again like the first, but the number of photons is reduced to 103, and the number of pixels of the first cluster is 30 as opposed to 150 each of the other two. The final hard case consists of a mixture of the others; the first distribution is skewed using the unmodulated light, the second is compact as in the easy case and the third is skewed using a similar approach but by drawing different fractions of the photons belonging to the 3.5 and 0.5ns lifetimes, thus pulling the distribution along the line joining these two lifetimes. This produces a hard case where the two elongated distributions overlap and one of them ends where the compact appears.

We used our simulated data described above against some of the best-known algorithms in the literature. Figure 3 shows the results of the simulations. Each row corresponds to a different data set, with the first column being the known ground truth and successive columns are the output of each of the tested methods. The reported value  $J$  is the Jaccard index [34], which is computed as a cluster-wise ratio between the intersection and the union of the ground truth and the result. The one method that consistently stands out in our battery of tests is the Gaussian Mixture Model (GMM) which systematically improves the Jaccard index compared to the others that may be stronger in some situations and weaker in others.

K-means [35], and its many variants like k-medoids [36], are amongst the most well-known of all the clustering techniques. They are elegant in terms of the simplicity by which they are formalized; find the  $K$  mean locations that define clusters such that the sum of squared distances of all the members within the  $K$  groups is minimal. These methods divide the space into areas such that each point is classified according to the nearest center of a cluster, so they generally fail when the clusters have irregular sizes and do not handle high covariance very well (examples in Fig. 3, panels B2, B3, C2, C3, D2, D3). Other well-known methods are based on the connectivity between the points by defining some kind of a metric. In the specific case of phasor data, the points live in a two dimensional space, the two dimensions being  $S$  and  $G$  which obey the vector addition rules of an Euclidean space, therefore it makes sense to establish an Euclidean distance





**Fig. 3.** Example of clustering methods applied to simulated phasor data. Each row corresponds to a situation where a different condition is tested, first column depicts the known ground truth and successive columns show the results of the different algorithms. The reported goodness-of-fit value is the Jaccard index.

metric between points. Hierarchical clustering [37] and affinity propagation [38] are examples of such connectivity-based methods and there are many variations depending on the linking criteria [35] within the connectivity map. Due to the fact that they use a cutoff in the metric, they tend to fail when clusters present non-zero covariance and, like k-means, cannot overcome overlapping (Fig. 3, B4, B5, B6, C4, D4). Mean shift [39] was included also in the examples, it is a geometrical approach where a grid of points is laid on the space and each is shifted until convergence using the data points in some close neighborhood. Its performance is best when the data is similar to a smooth surface without any big gaps or regions of irregular density in the clusters (Fig. 3(B7)) and will simply omit clusters with low number of members (Fig. 3(D7)). Similarly, a more recent method, density-based spatial clustering (dbscan) [40], is also highly dependent on the density of the points and tends to fail when there are conditions of sparse data points (Fig. 3(C8 and D8)). These last two are powerful because they automatically find outliers in the data and can handle situations with arbitrarily shaped distributions. Spectral clustering [41,42], a generalized version of dbscan, equally overcomes arbitrarily shaped distributions, but at the cost of finding somewhat arbitrary connections in the data (Fig. 3(E9)). Spectral clustering is the only true contender to our proposed method GMM. We note here that the term ‘spectral’ in the name ‘spectral clustering’ refers to the eigenvector decomposition of the data points, and not to any sort of spectral imaging or spectral phasor transform.

As can be seen in Fig. 3, looking at the colors of the clustering results and at the quantification using the Jaccard index, GMM systematically performs better than the rest of clustering algorithms. The main reason for this is that phasor data tends to spread normally [25] and therefore GMM is the best choice to model the scattered coordinates. On top of that, it is one of the few methods that can robustly handle overlapping data (Fig. 3(C10 and 3E10)). Related to this, the other great feature is that it provides a posterior likelihood of belonging to the clusters, so one can deal with outliers or dubious cases.

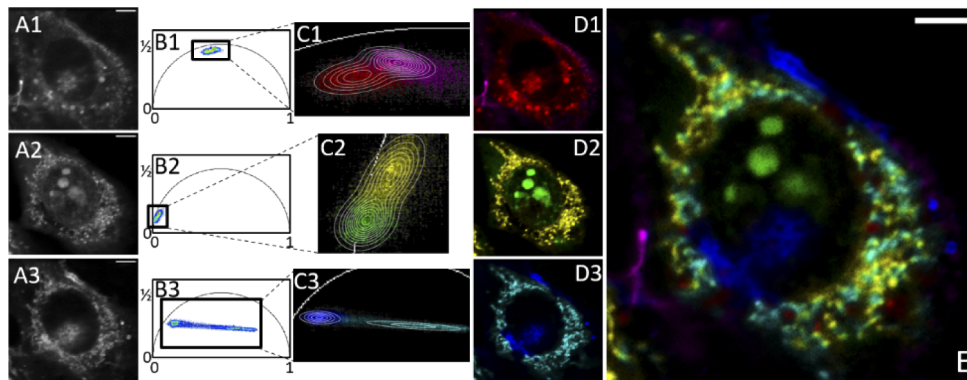
All the cases shown in Fig. 3, that is every one of the 9 algorithms tested against every one of the 5 datasets, was run 100 times and the best performance was kept. We are aware that there are many possible implementations of each and that the parameters need to be tuned for each case, we objectively attempted to obtain the best possible outcome in each algorithm. For each of the

algorithms we then produced the mean Jaccard index over the 5 datasets. GMM outperformed the rest obtaining a mean  $J=0.98$ , with the runners-up being hierarchical with Ward's criteria, spectral clustering and k-means (all around  $J=0.92$ ). The loser in our implementation was mean shift at  $J=0.83$ .

### 3.2. Live cell imaging

We present here as a proof of concept an example application in which GMM successfully clusters data obtained with a fluorescence lifetime microscope. The sample consisted of live cells stained with a collection of organelle-targeting dyes chosen to be spread out in the spectral and lifetime dimensions. It is proven that the lifetime of a molecule changes with modifications of the surrounding microenvironment [43], and therefore, the measured lifetime of dyes in solution is different to when they are inside the cells. For this reason, we first measured the lifetimes of the dyes individually in live cells (see Table 1, for a broader selection of tested dyes see Supp Materials) and we adjusted the concentration of the dyes such that the fluorescence intensity was approximately the same.

In the final experiment, we stained live cells with a total of six probes targeting different cell organelles: lysosomes, plasma membrane, endoplasmic reticulum, nuclei, Golgi apparatus and mitochondria (Fig. 4). We acquired 3 consecutive lifetime images in separate spectral bands in order to resolve the dyes according to their spectra, as detailed in the methods section. Roughly, the emission of two of the six dyes fall in each of the three spectral bands. The expected per-channel organelle fluorescent emission was; lysosomes and plasma membrane in the first channel (Fig. 4(A1)), endoplasmic reticulum and nucleus the second channel (Fig. 4(A2)), and Golgi apparatus and mitochondria in the third channel (Fig. 4(A3)). Phasor plots containing the lifetime representation for each channel present two populations as expected (figures in 4B). GMM was run on the list of phasor coordinates, demanding 2 clusters in each channel, and the outcome of the GMM clustering correctly clustered these populations (Figs. 4(C1-3)). As described in the previous section, the original images are color-coded according to posterior probabilities of the model (Figs. 4(D1-3)) and a final 6-plex image is shown as an overlay of the three channels (Fig. 4(E)).



**Fig. 4.** Multichannel example of image segmentation using phasor-GMM clustering. A live cell is stained using 6 probes targeting different cellular organelles. A) Three intensity images of the same field of view are taken in three different spectral bands. B) The phasor plot of each of the spectral bands is shown where we expect to find two populations in each. C) GMM successfully models two Gaussians in each of the phasor plots. D) Original images are color-coded according to posterior probabilities of the model. E) A final 6 colored image is shown as an overlay with a threshold to prevent overlapping of background. Scale bars are  $5\mu\text{m}$ .

We are fully aware that loading live samples with so many fluorophores may be interfering with cellular function and even hindering the detection of the individual fluorophores due to interactions amongst them. We consider that this handicap only strengthens the validity of the method because it still succeeds in clustering the six individual cellular organelles despite the overcrowding. We would like to point out that this is only a proof-of-concept experiment rather than an attempt to solve any biological question. On top of that, during the short time of the experiments (1h including incubation), we did not observe any cytotoxic effects in the cells such as formation of vacuoles or onset of apoptosis.

In this final image, each pixel takes the color associated to the cluster with a higher posterior probability multiplied by the normalized number of counts in the particular channel. Suppose a pixel which in one channel has very low photon counts and one of the clusters results in a relatively high probability, and in another channel, the same pixel shows the same probability of one of the clusters, but with a very high number of counts. In that case, the pixel is colored using the color associated with the cluster in the second channel. An example of this can be seen in Fig. 4 where in the central region overlapping the nucleus, the first channel shows some fluorescence belonging to the lysotracker probe (Fig. 4(D1) in red), and the third channel shows fluorescence belonging to the Golgi Bodipy probe (Fig. 4(D3) in blue). In the final panel Fig. 4(E), those pixels are colored in blue, meaning that the product of the posterior probabilities by the photon counts was higher in the cluster associated to the Golgi than the one associated to the lysosomes.

It is interesting to observe that the method can resolve relatively close phasor populations (see Fig. 4(C1) and 4(C2)) and, as shown in the simulations, it can also deal with heavily skewed cases (Fig. 4(C3)). Noticeable is also what is already reported in the literature [44]; lipids processed in the Golgi after staining start to appear in the plasma membrane as can be seen by the presence of a particular membrane region, where the segmentation algorithm assigned to the so called Golgi cluster (top-right section of the membrane in blue, Fig. 4(D3) and 4(E)).

#### 4. Discussion

In this paper we address the issue of automatic clustering of phasor transform data. The need for such a method has become apparent because of the wide use of the phasor approach in order to describe both spectral and lifetime properties in a sample. Automatic segmentation of images through the phasor plot as opposed to manual region selection is crucial for the reproducibility of experiments and for robustly extracting unbiased and quantitative data from lifetime or spectral experiments. Furthermore, the automation provides high throughput in batch analyzing sets of images. We envision the application as part of an automated pipeline in which for instance large fields of view, *e.g.* tissue samples, can be analyzed to identify regions of interest with the presence of particular targets are detected where one can subsequently perform a much higher resolution image.

The error in the phasor transform of a photon histogram depends on the square root of the number of counts, and the phasor data spreads out following a normal distribution [19]. For this reason, we make use of Gaussian Mixture Models to describe phasor data and successfully cluster the data points in the phasor space. In order to quantify the power of such method we simulated several sets of data and attempted to cluster them using a battery of the most widely used methods in the literature. The simulations were done on a photon-by-photon basis, drawing them out of predefined distributions, and allowing us to generate data from virtually any combination of lifetimes. We quantify the performance of each method under each set by means of the Jaccard index and show that GMM outperforms the rest of methods. These simulations convincingly establish GMM as the most appropriate machine learning clustering method considering the morphology of the data that is obtained in spectral and/or lifetime fluorescence measurements.

The one question that the reader may be asking is what resolution one can achieve on the phasor plot i.e. how close can two populations be for the method to be capable of separating them. The answer is that that depends on the number of photons one is collecting, theoretically if one has enough photons the resolution can get as small as one wishes. In fact, in some of the examples we provide, the distributions are so close that one would doubt there are actually two populations present (Fig. 4(B1 and B3)).

We demonstrate the use of GMM on a particular segmentation problem in which live cells have organelles stained with different probes, some overlapping in spectra, some overlapping in lifetime. The method performs an automatic and unsupervised clustering in the phasor space which translates to a segmentation in the image space where one can then label each cluster. We show the results using colors as labels, but the method allows for a fully quantitative analysis because it assigns a probability of each pixel to belong to each of the labels.

We show a simple example of the emission of three probes collected in a single spectral channel and then we show a more complex example with six probes with the emission collected in three spectral channels. These are good examples of application because each organelle is spatially resolved, meaning that we should not find the many pixels containing more than one fluorescent probes. A high proportion of these pixels will be a problem for the method; the pixels that share the contribution of two probes, have phasor coordinates that fall in a linear combination of the positions of the two pure species. The phasor plot of such a sample would have the two normally distributed clusters where GMM could be successfully fit, but additionally a whole range of points in between the two clusters, each of which corresponding to pixels with different fractions of the components. This situation would skew the shape of the distributions away from normality. The problem would become even worse if the probes were to interact with each other for the same reason, the distributions would move away from normality.

This is the main caveat we must point out; the system will lack validity when combinations of probes are present. In such situations, the method will tend to require additional clusters to describe the combinations/interactions which, on one hand will hinder the automation of the process requiring human input to determine how many additional clusters need to be used to model the data, and on the other hand, these additional clusters are not representative of a real physical fluorescent component.

**Funding.** NIGMS (P41-GM103540).

**Acknowledgements.** The authors would like to thank Raúl Benítez Iglesias, Universitat Politècnica de Catalunya, for inspiration behind this work.

**Disclosures.** The authors declare no conflicts of interest.

**Data Availability.** The set of MATLAB functions to generate the simulations of lifetime phasor data together with the code to draw Fig. 3 in the paper are publicly available in [Dataset 1](#), Ref. [33]. The live cell measurement files used for Figs. 2 and 4 are also included.

**Supplemental document.** See [Supplement 1](#) for supporting content.

## References

1. G. Weber, "Resolution of the fluorescence lifetimes in a heterogeneous system by phase and modulation measurements," *J. Phys. Chem.* **85**(8), 949–953 (1981).
2. D. M. Jameson, E. Gratton, and R. D. Hall, "The measurement and analysis of heterogeneous emissions by multifrequency phase and modulation fluorometry," *Appl. Spectrosc. Rev.* **20**(1), 55–106 (1984).
3. A. H. A. Clayton, Q. S. Hanley, and P. J. Verveer, "Graphical representation and multicomponent analysis of single-frequency fluorescence lifetime imaging microscopy data," *J. Microsc.* **213**(1), 1–5 (2004).
4. G. I. Redford and R. M. Clegg, "Polar plot representation for frequency-domain analysis of fluorescence lifetimes," *J. Fluoresc.* **15**(5), 805–815 (2005).
5. M. A. Digman, V. R. Caiolfa, M. Zamai, and E. Gratton, "The phasor approach to fluorescence lifetime imaging analysis," *Biophys. J.* **94**(2), L14–L16 (2008).
6. F. Fereidouni, A. N. Bader, and H. C. Gerritsen, "Spectral phasor analysis allows rapid and reliable unmixing of fluorescence microscopy spectral images," *Opt. Express* **20**(12), 12729–12741 (2012).



7. S. Ranjit, L. Malacrida, D. M. Jameson, and E. Gratton, "Fit-free analysis of fluorescence lifetime imaging data using the phasor approach," *Nat. Protoc.* **13**(9), 1979–2004 (2018).
8. C. Stringari, A. Cinquin, O. Cinquin, M. A. Digman, P. J. Donovan, and E. Gratton, "Phasor approach to fluorescence lifetime microscopy distinguishes different metabolic states of germ cells in a live tissue," *Proc. Natl. Acad. Sci. U. S. A.* **108**(33), 13582–13587 (2011).
9. E. Hinde, M. A. Digman, C. Welch, and K. M. Hahn, "Biosensor FRET detection by the phasor approach to fluorescence lifetime imaging microscopy (FLIM)," *Microsc. Res. Tech.* **75**(3), 271–281 (2012).
10. Z. Liang, J. Lou, L. Scipioni, E. Gratton, and E. Hinde, "Quantifying nuclear wide chromatin compaction by phasor analysis of histone Förster resonance energy transfer (FRET) in frequency domain fluorescence lifetime imaging microscopy (FLIM) data," *Data Br.* **30**, 105401 (2020).
11. W. Shi, D. E. S. Koo, M. Kitano, H. J. Chiang, L. A. Trinh, G. Turcatel, B. Steventon, C. Arnesano, D. Warburton, S. E. Fraser, and F. Cutrale, "Pre-processing visualization of hyperspectral fluorescent data with Spectrally Encoded Enhanced Representations," *Nat. Commun.* **11**(1), 726 (2020).
12. F. J. Vergeldt, A. Prusova, F. Fereidouni, H. Van Amerongen, H. Van As, T. W. J. Scheenen, and A. N. Bader, "Multi-component quantitative magnetic resonance imaging by phasor representation," *Sci. Rep.* **7**(1), 861 (2017).
13. H. Szmecinski, V. Toshchakov, and J. R. Lakowicz, "Application of phasor plot and autofluorescence correction for study of heterogeneous cell population," *J. Biomed. Opt.* **19**(4), 046017 (2014).
14. S. Ranjit, A. Dvornikov, M. Levi, and E. Gratton, "Multicomponent analysis of phasor plot to decipher changes in metabolic trajectory of biological systems," *Biophys. J.* **114**(3), 167a (2018).
15. N. Ma, M. A. Digman, L. Malacrida, and E. Gratton, "Measurements of absolute concentrations of NADH in cells using the phasor FLIM method," *Biomed. Opt. Express* **7**(7), 2441–2452 (2016).
16. S. C. Warren, A. Margineanu, D. Alibhai, D. J. Kelly, C. Talbot, Y. Alexandrov, I. Munro, M. Katan, C. Dunsby, and P. M. W. French, "Rapid global fitting of large fluorescence lifetime imaging microscopy datasets," *PLoS One* **8**(8), e70687 (2013).
17. M. A. K. Sagar, K. P. Cheng, J. N. Ouellette, J. C. Williams, J. J. Watters, and K. W. Eliceiri, "Machine learning methods for fluorescence lifetime imaging (flim) based label-free detection of microglia," *Front. Neurosci.* **14**, 1–11 (2020).
18. V. Mannam, Y. Zhang, X. Yuan, C. Rivasio, and S. S. Howard, "Machine learning for faster and smarter fluorescence lifetime imaging microscopy," *JPhys Photonics* **2**(4), 042005 (2020).
19. S. Ranjit, A. Dvornikov, M. Levi, S. Furgeson, and E. Gratton, "Characterizing fibrosis in UUO mice model using multiparametric analysis of phasor distribution from FLIM images," *Biomed. Opt. Express* **7**(9), 3519 (2016).
20. S. Ranjit, A. Dvornikov, E. Dobrinskikh, X. Wang, Y. Luo, M. Levi, and E. Gratton, "Measuring the effect of a Western diet on liver tissue architecture by FLIM autofluorescence and harmonic generation microscopy," *Biomed. Opt. Express* **8**(7), 3143 (2017).
21. D. Fu and X. S. Xie, "Reliable cell segmentation based on spectral phasor analysis of hyperspectral stimulated raman scattering imaging data," *Anal. Chem.* **86**(9), 4115–4119 (2014).
22. Y. Zhang, T. Hato, P. C. Dagher, E. L. Nichols, C. J. Smith, K. W. Dunn, and S. S. Howard, "Automatic segmentation of intravital fluorescence microscopy images by K-means clustering of FLIM phasors," *Opt. Lett.* **44**(16), 3928 (2019).
23. R. Brodwolf, P. Volz-Rakebrand, J. Stellmacher, C. Wolff, M. Unbehauen, R. Haag, M. Schäfer-Korting, C. Zoschke, and U. Alexiev, "Faster, sharper, more precise: Automated Cluster-FLIM in preclinical testing directly identifies the intracellular fate of theranostics in live cells and tissue," *Theranostics* **10**(14), 6322–6336 (2020).
24. C. M. Bishop, *Pattern Recognition and Machine Learning* (Springer, 2006).
25. R. A. Colyer, C. Lee, and E. Gratton, "A novel fluorescence lifetime imaging system that optimizes photon efficiency," *Microsc. Res. Tech.* **71**(3), 201–213 (2008).
26. L. Zelnik-manor, L. Zelnik-manor, P. Perona, and P. Perona, "Self-tuning spectral clustering," *Adv. Neural Inf. Process. Syst.* **17**, 1601–1608 (2004).
27. A. P. Dempster, N. M. Laird, and D. B. Rubin, *Maximum Likelihood from Incomplete Data Via the EM Algorithm* (Journal of the Royal Statistical Society, 1977), Vol. 39.
28. H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control* **19**(6), 716–723 (1974).
29. G. Schwarz, "Estimating the dimension of a model," *Ann. Stat.* **6**(2), 461–464 (1978).
30. T. Huang, H. Peng, and K. Zhang, "Model selection for Gaussian mixture models," *Stat. Sin.* **27**, 147–169 (2017).
31. G. Bianchetti, M. De Spirito, and G. Maulucci, "Unsupervised clustering of multiparametric fluorescent images extends the spectrum of detectable cell membrane phases with sub-micrometric resolution," *Biomed. Opt. Express* **11**(10), 5728 (2020).
32. R. D. Spencer and G. Weber, "Measurement of subnanosecond fluorescence lifetimes with a cross-correlation phase fluorometer," *Ann. N. Y. Acad. Sci.* **158**, 361–376 (1969).
33. A. Vallmitjana, "Phasor Clustering Dataset 1," figshare, (2021), <https://doi.org/10.6084/m9.figshare.13232534.v1>.
34. M. Levandowsky and D. Winter, "Distance between sets," *Nature* **234**(5323), 34–35 (1971).
35. J. MacQueen, "Some methods for classification and analysis of multivariate observations," *Proc. Fifth Berkeley Symp. Math. Stat. Prob* **1**, 281–297 (1967).
36. L. Kaufman and P. J. Rousseeuw, *Clustering by Means of Medoids* (North Holland / Elsevier, 1987).

37. J. H. Ward Jr, "Hierarchical grouping to optimize an objective function," *J. Am. Stat. Assoc.* **58**(301), 236–244 (1963).
38. B. J. Frey and D. Dueck, "Clustering by passing messages between data points," *Science* **315**(5814), 972–976 (2007).
39. K. Fukunaga and L. D. Hostetler, "The estimation of the gradient of a density function, with applications in pattern recognition," *IEEE Trans. Inf. Theory* **21**(1), 32–40 (1975).
40. M. Ester, H.-P. Kriege, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," *Proc. 2nd Int. Conf. Knowl. Discov. Data Min.* 226–231 (1996).
41. A. Y. Ng, M. I. Jordan, and Y. Weiss, "On spectral clustering: analysis and an algorithm," in *Proceedings of the 14th International Conference on Neural Information Processing Systems: Natural and Synthetic*, NIPS'01 (MIT Press, 2001), pp. 849–856.
42. J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(8), 888–905 (2000).
43. P. Sarder, D. Maji, and S. Achilefu, "Molecular probes for fluorescence lifetime imaging," *Bioconjugate Chem.* **26**(6), 963–974 (2015).
44. R. E. Pagano, "A fluorescent derivative of ceramide: physical properties and use in studying the Golgi Apparatus of animal cells," *Methods Cell Biol.* **29**, 75–85 (1988).