Spatio-temporal interaction of bacteria mixture within biofilms

Y. Li, K.S. Kim, Julien Deschamps, Romain Briandet, Alain Trubuil

Available online at www.sciencedirect.com

ScienceDirect

Procedia Environmental Sciences 26 (2015) 11 – 18

Procedia Environmental Sciences

Spatial Statistics 2015: Emerging Patterns

Spatio-temporal interaction of bacteria mixture within biofilms

Y. Li[a], K. S. Kim[bc], J. Deschamps[bc], R. Briandet[bc], A. Trubuil[a]*

[a]INRA, UR341 Mathématiques et informatique appliquées, F-78350 Jouy-en-Josas
[b]INRA, UMR1319, F-78350 Jouy-en-Josas [c]AgroParisTech, Micalis Institute, F-78350 Jouy-en-Josas

Abstract

The biofilm ubiquitously exists on most wet surfaces. It is a protective shield of the bacteria and causes the difficulty in the disinfection. The irrigation of the biofilm by the specific swimmer bacteria can exacerbate killing of biofilm bacteria. Therefore, we precisely investigate the tunneling of swimmers bacteria within biofilms. These bacterial stealth swimmers create transient opened spaces in the biofilm. We found that these opened spaced in the biofilm is the obvious indication of the motion of the swimmers. We exploit both detected swimmer in one swimmer channel and the opened space in the other biofilm channel of confocal microscope video, in order to interactively improve the tracking of the swimmer's motion in the biofilm, which is implemented by improving a successful algorithm of single particle tracking. Moreover, we quantitatively study the swimmer trajectory, the opened space and their relation in the large-scale microscope video when the biofilm is separately irrigated by many different species of bacteria. We try to discover the influence of the irrigated bacteria on the biofilm.

Peer-review under responsibility of Spatial Statistics 2015: Emerging Patterns committee

Keywords: Kalman filter; Linear assignment problem; Particle tracking; Biofilm; Multi-sets

1. Introduction

Along the food chain, most wet surfaces are covered by biological pellicles called biofilms. Biofilms are composed of spatially organized micro-organisms embedded in an extracellular matrix [1][8]. This organic cement can act as a protective shield against the action of antimicrobials, thus raising serious problems of pathogens persistence and extra-use of chemical biocides in industrial settings. Houry et al. [3] discovered that planktonic bacilli propelled by flagella are able to tunnel deep within the biofilms structures. These bacterial stealth swimmers

* Corresponding author. Tel.: (+33) (0)1 34 65 22 22; fax: (+33) (0)1 34 65 22 17.
E-mail address: Alain.Trubuil@jouy.inra.fr

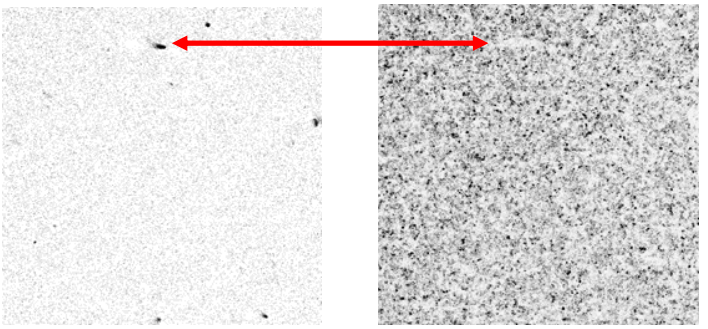
Peer-review under responsibility of Spatial Statistics 2015: Emerging Patterns Committee
doi:10.1016/j.proenv.2015.05.009

create transient opened spaces that increase macromolecular transfer within the biofilm. It has been proved [2][3] that the irrigation of the biofilm by swimmer bacteria can exacerbate killing of biofilm bacteria by facilitating penetration and action of disinfectants from the environment. So the researchers become interested in studying more precisely the tunneling of swimmers bacteria within biofilms. The time-lapse confocal microscopy 2D+Time multichannel videos show the dynamic interaction of swimmers and biofilm may exist. In the present study, dynamics are studied in the vicinity of the adherence surface where the biofilm sticks. Nevertheless bacteria can enter or leave this region so occlusions of trajectories occur. Depending on bacteria strains, swimmers bacteria may adopt separated dynamics or aggregate in chains of several bacteria. Herein we will consider strains which do not aggregate or divide.

In this paper we will explain how we process such video data in order to extract and analyze quantitative information with respect to the interaction of the swimmers within the biofilm. First, swimmers are localized in each frame and then linked across the frames. Transient opened spaces or pores created by swimmers within the biofilm are correlated with swimmers tracks. So we decide to exploit the image information between the channel of swimmer bacteria and the channel of biofilm background from the microscope video to achieve the better quality of the swimmer and pore tracking. Once swimmers and pores are localized and tracked several features (e.g. speed, persistence, pore area …) are computed. So the comparison and the classification of swimmers dynamics within the biofilm itself can be addressed with classical statistical tools. Moreover questions related to the interactions between swimmers and between swimmers and biofilm can be considered. Hereafter on one hand we question if the swimmers reuse preferentially space within the biofilm already visited and propose a simple statistical test. On the other hand we also show preliminary results related to mixtures of swimmers within the biofilm.

2. Swimmer tracking by the features of two channels

Since the biofilm video from the microscope is composed of two channels – the swimmer channel (S-channel) and biofilm channel (B-channel), as shown in Fig. 1, we will explore features in both channels to improve the tracking of the bacteria's motion behavior in the biofilm.



(a) Swimmer channel. (b) Biofilm channel.

Fig. 1. Both channels in the biofilm video.

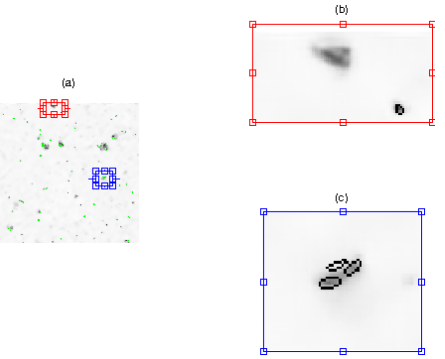


Fig. 2. The detection of the swimmer in S-channel.

2.1. The detection of swimmers in the swimmer channel

Since the images from the microscopy as Fig. 1 are extremely noisy, we will first denoise both kinds of images by ND-SAFIR toolkit [4]. Then, depending on the difficulty in the segmentation of S-channel either a simple threshold method or a specific segmentation approach based on the detection of elliptic points was used. The elliptic point is defined as the point on a regular surface, whose Gaussian curvature or equivalently, the principal curvatures have the same sign. And the elliptic point is detected by the Gaussian filtered Hessian matrix of the image. Once detected we filter the elliptic points by several criteria related to intensity, including the intensity inside ellipses, the distance between elliptic points, and the intercepted areas between ellipses. This step is followed by a classical watershed segmentation method. An example of swimmer detection is shown in Fig. 2.

2.2. The detection of the pore in the biofilm channel

We detect the pore in biofilm frames by the superpixel [5] because there is not pixel-level segmentation in biofilm frames. We partition the frame image into superpixels shown in Fig. 3. In the collaboration of two channels of the biofilm, the superpixels around the swimmer are of high intensity in swimmer channel, while low intensity in biofilm channel. We use this property as the criterion to decide the pore together with the Euclidean distance to the swimmer. An example result is in Fig. 4. Each detected pore is assigned to its closest swimmer by Euclidean distance.

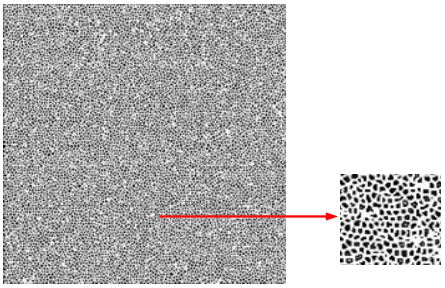


Fig. 3 The partition by the superpixel.

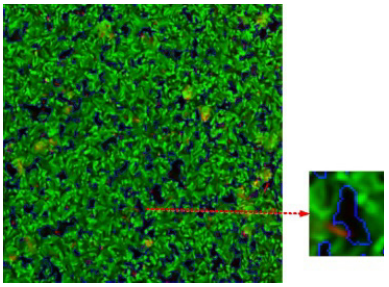
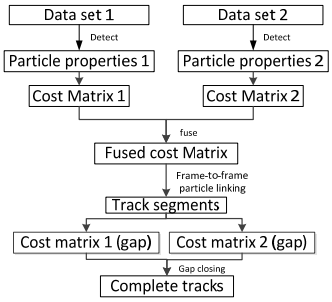


Fig. 4 The detection of the pore around the swimmer.

2.3. The swimmer tracking by both channels

In the past many different approaches of particles tracking have been proposed by researchers, but most approaches are proposed to track the particles in a data set. Here we are meeting with two or more collections of particles (one or two swimmers species, bacteria from the biofilm) moving in the same space and observed in more or less specific channels images. There are correlations between the images from these two channels. For instance a portion of space occupied by a swimmer is not occupied by the bacteria from the biofilm. Also as soon as a swimmer moves within the biofilm it opens a portion of space which stays free of bacteria from the biofilm for some time. Therefore, more accurate tracking by both sets of particles than by one set of particles should be possible. In this paper we propose an algorithm to consider the tracking by using two sets of correlated particles. The approach in [6], as well its toolkit U-Track, has been a successful approach in tracking the particle objects. U-Track includes two steps. In the first step, the Kalman filter is exploited to estimate the propagation of particles, compute the linking cost between particles in the consecutive frames and construct the cost matrix of consecutive frames. Then the linear assignment problem (LAP) [7] globally minimizes the linking cost defined by the cost matrix of particles between consecutive frames to find the linking of the particles. Similarly the track segments from the first step are linked to form the complete trajectories.

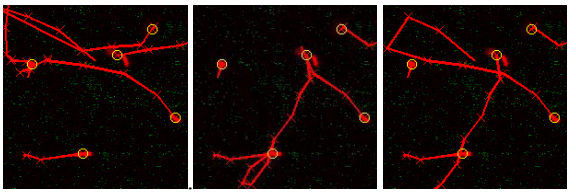
Inspired by U-Track we propose DU-Track suitable for two sets of related particles. The framework of DU-Track is illustrated in the following figure:



We regard each set of data as the individual data and implement the Kalman filter on each data to get its cost matrix. Then two cost matrices are fused according to the relation of two sets of data. The fusion of two cost matrix is formulated as Eq. 1:

$$c_{ij} = \lambda c_{ij}(x_i, x_j) + (1 - \lambda) c'_{ij}(y_i, y_j) \quad (1)$$

where ℓ_{ij} is the element of cost matrix for the data i and j in set S, while ℓ'_{ij} is the element of cost matrix for the data i and j in the other set P. c_{ij} is the new element of cost matrix for the data s_i in the set S, together with its related data at p_i in the other set P at frame t, and the data s_j, together with its related data p_j in the other set P at frame $t+1$. The weight λ is decided according to the correlation between two sets S and P. In the proposed DU-Track, we take into account the histogram and shape of the swimmer, and sum up shape and intensity into the normalized n-bins histogram [9]. In addition, we implement the same procedure for gap closing after linking. We have used real biofilm videos to prove that DU-Track is better than U-Track, especially when swimmers meets with each other. We can see that DU-Track tracks better for swimmers in crowd and create the simpler trajectory. However, it is hard to get the ground truth even with the help from the expert. So we also create the simulation video with the swimmer number of 40, 80, 120, 160 and 200, each with 5 repetitions. It is easy in Fig. 6 to find that DU-Track achieves the higher recall rates together with the lower wrong tracking in the simulation videos. And IU-Track, which is the cost matrix of the pore by using the motion propagation from Kalman filter of the swimmer, is normally better than U-Track.



(a) U-Track (b) IU-Track (c) DU-Track

Fig. 5. The trajectory of real biofilm videos.

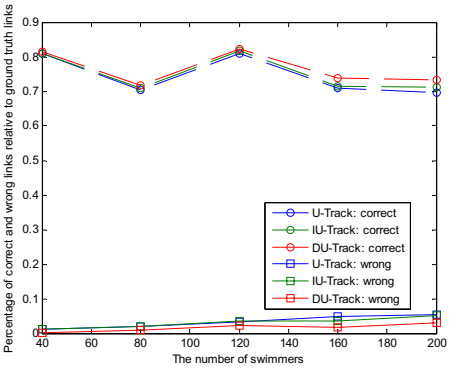


Fig. 6. The correct and wrong rates in the simulation.

We have applied DU-TRACK on the large-scale biofilm videos from the confocal microscope. The tested videos contain the following bacteria species: *Bacillus cereus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus mycoides*, *Bacillus polymyxa*, *Bacillus pumilus*, *Bacillus sphaericus*, and *Bacillus subtilis and the mixture of 2 species*. Because of the limited page, we only show several groups of data about swimmer speed, the pore and the number of the swimmer in Table 1. And we display the maximal projection of S-channel in each video of *B. licheniformis* in Fig. 7. We will analyze Table 1 in detail in this section.

Table 1. 12 kinds of data about speed and pore in several species.

Species / Video/strain Name	Licheniformis	Megaterium	Mixture of 2 species	Polymyxa	Pumilus	Sphaericus	Subtilis
	7B1.lif-Series007-C_1	11A5.lif-Series011-Z0C_1	cere11B15+sph11G3.lif-Series015-C_1	11B1.lif-Series003-Z0C_1	11B16.lif-Series002-C_1	10C3.lif-Series008-C_1	20B9.lif-Series004-C_1
Mean speed	3.065358	8.823769	13.629228	10.136891	6.483279	9.223361	9.960491
Maximum speed	7.52994	19.105562	31.555106	19.464718	14.795022	20.409907	29.634086
Mean of max speeds of frames	13.94485	28.812816	64.768949	17.822918	20.845176	25.009459	32.070953
Mean persistence	0.258692	0.514866	0.498213	0.619235	0.454129	0.600873	0.373457
Maximum persistence	0.827319	0.990321	0.982525	0.983908	0.932129	0.994437	0.836825
Pore area	8220.399	3644.096675	11520.92651	3084.904678	5072.765021	4109.576313	2079.521767
Pore area excluding eDNA	7172.127	3606.011099	9586.329589	52.925336	4229.369963	NaN	318.815566
Pore area per swimmer	170.2346	66.841019	50.811605	219.656099	148.932873	48.903859	84.796033
Pore area per swimmer excluding eDNA	148.5261	66.142443	42.279307	3.768471	124.171378	NaN	13.000246
Mean number of swimmers	48.28866	54.518868	226.738095	14.044248	34.060748	84.033784	24.52381

Swimmer number in and around eDNA	0.088204	0.034219	0.190118	0.306985	0.253886	0.495888	0.48097
Swimmer number in and around eDNA per eDNA size	0.000084	0.000898	0.000098	0.000101	0.000301	0.000117	0.000273

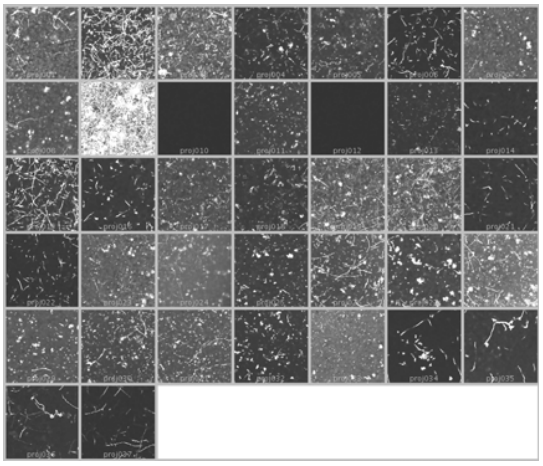


Fig. 7. The maximal projection of each video of *B. Licheniformis*.
Light regions are associated to visited space by swimmers.

2.4. The discussion about the speed and pore features

2.4.1. speeds

In Table 1, we have shown 3 kinds of data about the motion speed of the swimmer: mean speed, maximum speed, and mean of max speeds of frames. For better understanding, we separately describe the definition of the mean speed of swimmer tracks in a video, maximum speed of swimmer tracks in a video, and the mean of max speeds in each frame as following equations:

$$V_{mean\ speed}(e,s,k) = mean_{t\in T(e,s,k)}(V_t); V_{max\ speed}(e,s,k) = quantile(V_{t\in T(e,s,k)}, 0.9); \quad (2)$$
$$V_{mean\ of\ max\ speed}(e,s,k) = mean_t\left(\max_{t\in T(e,s,k)} V(i,t)\right)$$

where the set (*species; strain; sample*) means a video belonging to a sample *k* of a strain *s* of a species *e*. *T* is the set of tracks in the trajectories. V_i denotes the speed of the track *i*. *t* is the time of the frame in a video (*species; strain; sample*).

In Fig. 8 and 9, we separately show the notched box plots and p-value from one-way Anova analysis of each species of the mean speed of swimmer tracks in a video and the mean of max speeds in each frame for different species. From this global analysis speeds are different for this collection of species. However considering multiple test *B. licheniformis* is different from *B. sphaericus* at a 5% significance level, *B. Polymyxa* is significantly different from *B. sphaericus*, and *B. pumilus* species and *B. subtilis* is not significantly distinctive from others.

2.4.2. persistence

For a particle denoted *i*, the persistence at time *t* is defined as the ratio of the distance between the positions at time t_0 and *t* of the particle to the covered distance by the particle during this time: $P_{t_0}(t) = \frac{|x_i(t)-x_i(t_0)|}{s_i(t_0,t)}$ (3)

with $s_i(t_0,t) = \int_{t_0}^{t}\int \|x_i'(u)\|\, du$. In Fig. 10 we display the mean persistence in the whole video of each video in all the species. The persistence of *B. cereus* and *B. polymyxa* are higher than others.

2.4.3. pores size

We want to check if the swimmer reuses the pore opened by other swimmers in the biofilm. So we have tested the pore size with only 1 species or 2 kinds of species. But we did not find the obvious difference. The distribution is tested as similar in two cases and different in two other cases. However, we have shown the box plot of the pore size

of all the species in Fig. 11. We can found that their properties are distinct.

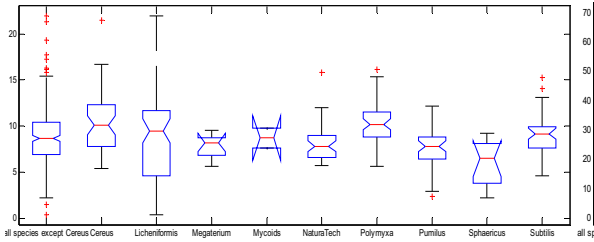


Fig. 8. the mean speed of swimmer tracks in the whole video.

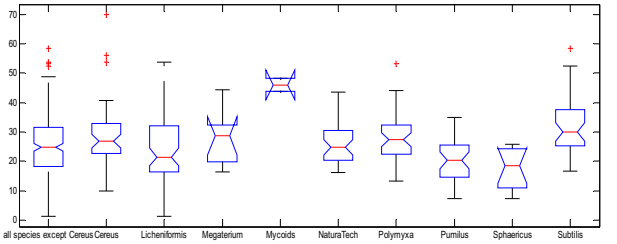


Fig. 9. The mean of max speeds in each frame

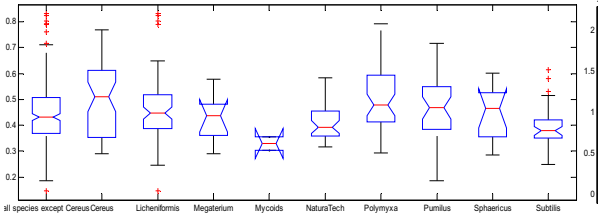


Fig. 10. The mean persistence in the whole video.

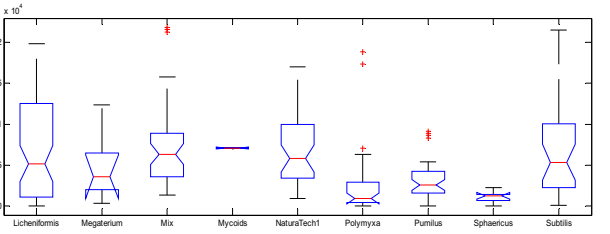


Fig. 11. The pore size at different frame time in a video with mixture of species

We have the bag of swimmer features of 12 values, however 12 dimensions are too high to get the inner property of swimming characteristics of different species. Partial least squares Discriminant Analysis (PLS-DA) [10] is a classical PLS regression [11] but its response variables are classes. PLS-DA is popularly used in the classification and discrimination problems. In PLS-DA, we use the variables of motion properties as the predictor variables and swimmer species as the response variable. The PLS-DA of our data is shown in Fig. 12. It is obvious that PLS-DA can discriminate the strain species.

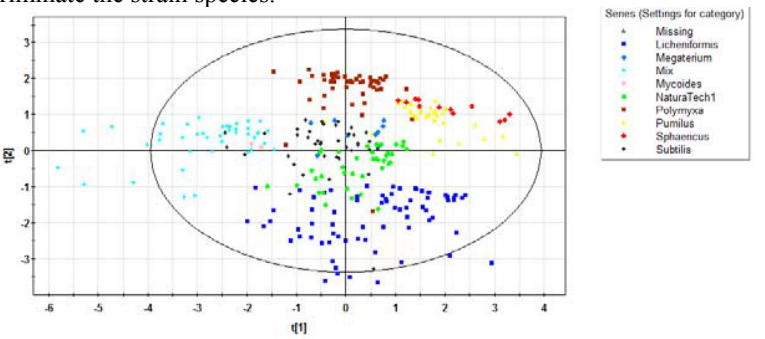


Fig. 12. The PLS-DA analysis of the bag of swimmer motion features in different strains. Several species are well separated on the two first axes.

2.5. The further discussion of the biofilm

We can conclude that the mean speeds of each swimmer species are not so different in Fig. 8, however, the variance of the speed in different species are obviously different in Fig. 9. Through the large-scale experiments, we also found that most swimmers stay in one focal plane only for 3-10 frames (large swimmers like *B. sphaericus* could stay longer), and then randomly move up or down to the other focal planes of the 3d structure of the biofilm.

Swimmers and eDNA

Some of the variability in the speed could be explained by the existence of e-DNA pockets inside the extra-cellular matrix from the biofilm. The eDNA is one of natural extracellular environment of biofilm, which can be stained by the cell permeant SYTO 61 red fluorescent nucleic acid. SYTO 61 red fluorescent nucleic acid cannot immerse and penetrate into the biofilm in a short time, but it can color eDNA. So if the eDNA exists in the biofilm, the eDNA can be always visible in S-channel and sometime in B-channel. The swimmers may stick to eDNA.

Do the swimmers reuse paths already visited?

First of all, we can see the proportion of space visited by swimmers grow regularly along time in Fig. 13. However to address in more details this question we proceed in several steps. First we have built the progressive projections of space visited by swimmers along time. In Fig. 14 we show (in gray) the space visited from time step *t=0* (frame 1) to a couple of time steps. The second step consists in considering swimmers at time *t+Delay* and comparing their locations to the space visited up to time *t*. The parameter *Delay* is fixed in such a way that the bacteria observed at time *t+Delay* has no chance to be present at time *t* except if they were more or less static or they have been leaving the focal plane in between time *t=0* and time *t* and come back. Bacteria that are more or less static are filtered out (red points). They are present in more or less fixed locations from time *t* to time *t+Delay*. For the number *N(t+Delay)* moving swimmers at time *t+Delay* (blue points) the distances from their centroid position to the visited space at time *t* are computed. At the same time we generate *N(t+Delay)* uniformly random locations in the image plane and compute their distances to the visited space. In Fig. 14 we also show the distributions of distances to the visited space for the swimmers on the one hand and random locations on the other hand. As the visited space is more and more filled the distances to the visited space are more and more short. The high proportion of swimmers close to the visited space may suggest swimmers reuse the paths already visited. In Fig.14, we consider two samples, one from *B. licheniformis* (left panel), and one from *B. sphaericus* (right panel). Applying the Kolmogorov-Smirnov test of distributions comparison we accept null hypothesis most of the times for *B. licheniformis* and reject most of the times for *B. sphaericus*. In the first case it seems that swimmers bacteria preferentially do not reuse visited space at contrary to the second case. In the first case the bacteria are much smaller than in the second case.

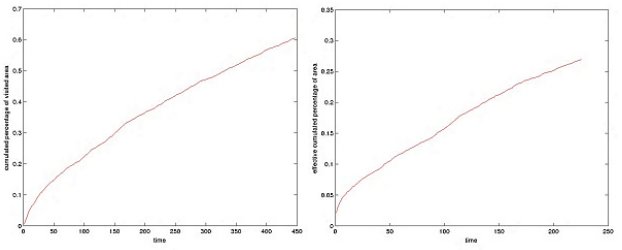


Fig.13. The proportion of visited space along time for a sample from *B. Licheniformis* (resp. *B. sphaericus*) on the left (resp. right).

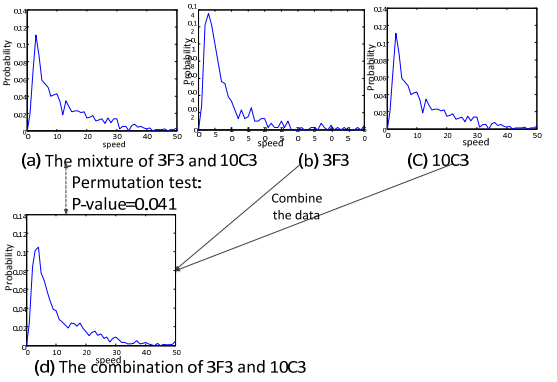


Fig. 15. The distributions of mean speed for the mixture of swimmers from B. megaterium and B. Pumilus.

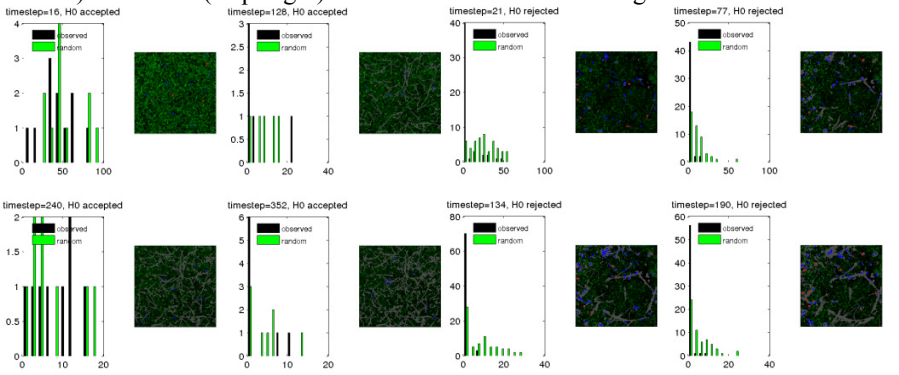


Fig. 14. Distribution of distances of swimmers (centers in blue on images) detected at different time steps to previously visited space (in gray on images) of biofilm (in green). Also we show distributions of distances for randomly distributed points. The panel on the left (resp. right) side is from a sample of *B. licheniformis* (resp. *B. sphaericus*) swimmers.

Does mixture of swimmers from different species result in interaction effects?

We compare the mixture of *B. pumilus* (strain 3F3) and *B. megaterium* (strain 10C3) to the pure strains. In Fig. 15 we can found the mixture of 3F3 and 10C3 is similar to 10C3. From left to right they are the mixture, the individual strains and the distribution created from the random sampling of individual strains in equal proportions. At least from this mixture, we can found that one species in the mixture dominates the swimming behavior of both in the mixture, while the impact from the other species is little. In Fig.16 it seems the two strains taken individually may reuse the previously visited space. Taken together it is not so clear.

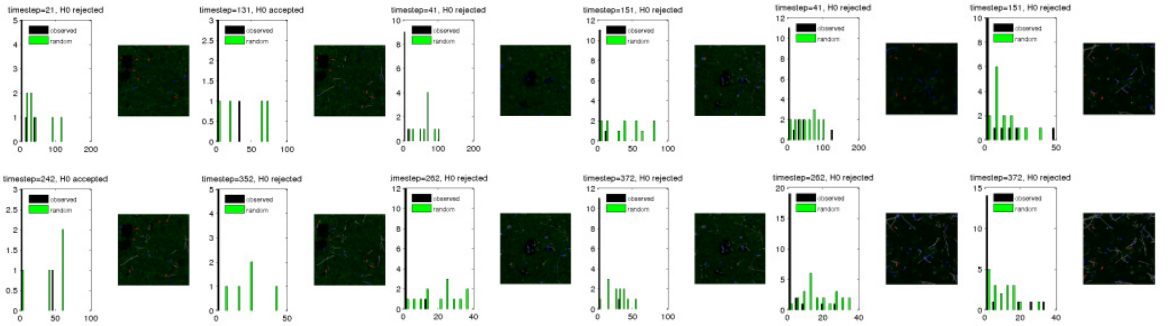


Fig.16. Distribution of distances of swimmers to visited space. From left to right a sample from a mixture of *B. pumilus* and *B. megaterium*, *B. pumilus*, *B. megaterium*.

3. Conclusions

In this paper we have firstly proposed an algorithm to achieve better performance of tracking two correlated sets of particles by exploiting their correlation and the Kalman filter of each set, especially for the crossing or close particles, based on the successful single particle tracking algorithm, U-Track. We have proved that the proposed DU-Track can better track two correlated sets of data in the biofilm, which are the detected swimmers and the pores in the biofilm. Then we have analyzed in detail the features of biofilm from detecting and tracking the swimmer and the pore in order to reveal the behavior of the swimmer and its influence on the biofilm. The swimmer speeds of various species are different but the sizes of the created pores are not significantly different. However according the large-scale data of swimmer species it is possible to classify the swimmer species if we have the features of the swimmer, the pore and the corresponding trajectory. It seems that swimmers do not reuse visited space but this may depend on their size, and further investigation is needed. Detailed study of mixture of swimmers from different species is undergoing and can benefit from the tools already developed for single species tracking studies. Above results can promote the knowledge of the best swimmer species in destroying the biofilm.

References

1. Costerton, J. William. "Introduction to biofilm." International journal of antimicrobial agents 11.3 (1999): 217-221.APA
2. Bridier, Arnaud, et al. "Resistance of bacterial biofilms to disinfectants: a review." Biofouling 27.9 (2011): 1017-1032.
3. Houry, A. et al., (2012) Bacterial swimmers that infiltrate and take over the biofilm matrix. Proceedings of the National Academy of Sciences 109.32: 13088-13093.
4. J. Boulanger. Non-parametric estimation and contributions to image sequence analysis: Modeling, simulation and estimation of the intracellular traffic in video-microscopy image sequences. Université de Rennes 1, Mention Traitement du Signal et des Télécommunications, Jan 2007.
5. A. Lucchi, K. Smith, R. Achanta, V. Lepetit and P. Fua, A Fully Automated Approach to Segmentation of Irregularly Shaped Cellular Structures in EM Images, International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), Beijing, 2010.
6. Jaqaman, Khuloud, et al. "Robust single-particle tracking in live-cell time-lapse sequences." *Nature methods* 5.8 (2008): 695-702.
7. Jonker, Roy, and Anton Volgenant. "A shortest augmenting path algorithm for dense and sparse linear assignment problems." *Computing* 38.4 (1987): 325-340.
8. Costerton, J.W. (2007). "The Biofilm Primer", Springer.
9. Li, Y, Briandet, R, and Trubuil A., "Tracking Swimmers Bacteria and Pores within a Biofilm ", IEEE International Symposium on Biomedical Imaging (2014).
10. M. Pérez-Enciso and M. Tenenhaus, "Prediction of clinical outcome with microarray data : a partial least squares discriminant analysis (pls-da) approach," Human genetics, vol. 112, no. 5-6, pp. 581–592, 2003.
11. S. Wold, M. Sjöström, and L. Eriksson, "Pls-regression : a basic tool of chemometrics," Chemometrics and intelligent laboratory systems, vol. 58, no. 2, pp. 109–130, 2001.