**University of
Zurich^{UZH}**

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2011

A novel algorithm for the determination of bacterial cell volumes that is unbiased by cell morphology

Zeder, M ; Kohler, E ; Zeder, L ; Pernthaler, J

Abstract: The determination of cell volumes and biomass offers a means of comparing the standing stocks of auto- and heterotrophic microbes of vastly different sizes for applications including the assessment of the flux of organic carbon within aquatic ecosystems. Conclusions about the importance of particular genotypes within microbial communities (e. g., of filamentous bacteria) may strongly depend on whether their contribution to total abundance or to biomass is regarded. Fluorescence microscopy and image analysis are suitable tools for determining bacterial biomass that moreover hold the potential to replace labor-intensive manual measurements by fully automated approaches. However, the current approaches to calculate bacterial cell volumes from digital images are intrinsically biased by the models that are used to approximate the morphology of the cells. Therefore, we developed a generic contour based algorithm to reconstruct the volumes of prokaryotic cells from two-dimensional representations (i.e., microscopic images) irrespective of their shape. Geometric models of commonly encountered bacterial morphotypes were used to verify the algorithm and to compare its performance with previously described approaches. The algorithm is embedded in a freely available computer program that is able to process both raw (8-bit grayscale) and thresholded (binary) images in a fully automated manner.

DOI: <https://doi.org/10.1017/S1431927611012104>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-59993>

Journal Article

Accepted Version

Originally published at:

Zeder, M; Kohler, E; Zeder, L; Pernthaler, J (2011). A novel algorithm for the determination of bacterial cell volumes that is unbiased by cell morphology. *Microscopy and Microanalysis*, 17(5):799-809.

DOI: <https://doi.org/10.1017/S1431927611012104>

A novel algorithm for the determination of bacterial cell volumes that is unbiased by cell morphology

M. Zeder ^{1,2,3*}, E. Kohler ³, L. Zeder ², J. Pernthaler ³

¹ Max Planck Institute for Marine Microbiology, Department of Molecular Ecology,
Celsiusstrasse 1, 28359 Bremen, Germany

² technobiology GmbH, Rütliweidhalde 7a, 6033 Buchrain, Switzerland

³ Department of Limnology, Institute of Plant Biology, University of Zürich, Seestrasse 187,
8802 Kilchberg, Switzerland

Running headline: determination of bacterial cell volumes

* corresponding author

Michael Zeder

Rütliweidhalde 7a

CH-6033 Buchrain

Email: mzeder@technobiology.ch

20 **Abstract:**

21

22 The determination of cell volumes and biomass offers a means of comparing the
23 standing stocks of auto- and heterotrophic microbes of vastly different sizes for
24 applications including the assessment of the flux of organic carbon within aquatic
25 ecosystems. Conclusions about the importance of particular genotypes within microbial
26 communities (e.g. of filamentous bacteria) may strongly depend on whether their
27 contribution to total abundance or to biomass is regarded. Fluorescence microscopy and
28 image analysis are suitable tools for determining bacterial biomass that moreover hold the
29 potential to replace labor-intensive manual measurements by fully automated approaches.
30 However, the current approaches to calculate bacterial cell volumes from digital images
31 are intrinsically biased by the models that are used to approximate the morphology of the
32 cells. Therefore, we developed a generic contour based algorithm to reconstruct the
33 volumes of prokaryotic cells from two-dimensional representations (i.e., microscopic
34 images) irrespective of their shape. Geometric models of commonly encountered bacterial
35 morphotypes were used to verify the algorithm and to compare its performance with
36 previously described approaches. The algorithm is embedded in a freely available
37 computer program that is able to process both raw (8-bit grayscale), and thresholded
38 (binary) images in a fully automated manner.

39

40

41

42 cell volume / high-throughput microscopy / model based object oriented image
43 analysis / microbial ecology / screening

44 **Introduction**

45 Microscopy has been a key instrumentation since the beginning of microbiology
46 (Gest 2004). Recent advances in fluorescent staining, digital imaging and microscope
47 motorization have paved the way for the high-throughput analysis of microbial cells
48 (Pernthaler et al. 2003; Daims & Wagner 2007; Singleton et al. 2001; Zhou et al. 2007).
49 Automation allows for processing of vast numbers of samples, thus increasing spatial and
50 temporal resolution of studies on the distribution of different bacterial populations
51 (Schattenhofer et al. 2009), thereby providing new insight in the functioning of microbial
52 ecosystems. So far, fully automated microscopy and image analysis primarily deliver cell
53 counts in multiple fluorescent channels, yielding total bacterial numbers and relative
54 abundances of specific groups at a precision that can be superior to manual counting (Seo
55 et al. 2010).

56 However, the mere counting of bacteria does not take into account their variable
57 sizes. In phytoplankton ecology, the measurement of cell sizes has been a standard
58 procedure for decades, and complex geometrical models are used to calculate the
59 biovolumes and carbon content of different algal species (Hillebrand et al. 1999). This is
60 deemed essential because phytoplankton communities feature cells with a large variety of
61 morphologies and size classes. Cell numbers, therefore, do not adequately reflect the
62 respective importance of different species, e.g., their contribution to the standing stock
63 (Carlson et al. 1996) or to the flux of organic carbon through food webs (Falkowski et al.
64 1998). Moreover, large and small organisms may exhibit basic metabolic differences that
65 lead to specific annual succession patterns (Sommer et al. 1986). The same may hold true
66 for prokaryotic microbes, which may also substantially differ in cell size (Bertoni et al.
67 2010) and, to a lesser extent, in shape (Young 2006). This information may be ecologically
68 highly relevant, e.g. with respect to their growth potential and predation vulnerability (Zeder
69 et al. 2009; Chrzanowski & Simek 1990). In fact, strikingly different conclusions may be
70 reached about the relative importance of particular populations within microbial
71 assemblages if analyzed in terms of biomass rather than abundances (Posch et al. 2009;
72 Pernthaler et al. 2004).

73 In addition to the determination of biovolumes of entire communities, it has also
74 become more important to accurately measure the volume of individual microbes in order
75 to study their ecophysiology. Nano-scale secondary-ion mass spectroscopy (nanoSIMS) is

76 a novel powerful single cell analysis technology to investigate and assign specific
77 metabolic activities to the identities of individual cells (Musat et al. 2008). For such
78 analysis it is crucial to relate the highly accurate measurement of matter content with an
79 equally precise determination of its volume.

80 The estimation of bacterial cell volumes by means of microscopy and image
81 analysis has been a long standing topic in the literature, and many methods have been
82 proposed for this purpose (Boyde & Williams 1971; van Veen & Paul 1979; Krambeck et
83 al. 1981; Fry & Davies 1985; Bratbak 1985; Bjørnsen 1986; Sieracki et al. 1989; Bloem et
84 al. 1995; Massana et al. 1997; Blackburn et al. 1998; Posch et al. 2009). The standard
85 procedure is to acquire digital images of bacteria and to measure their dimensions by
86 image analysis. The microscopic image of a bacterial cell (Fig. 1 A) is a two-dimensional
87 projection, which, in the optimal case, represents its cross-sectional area (Fig. 1 B). Since
88 it is not possible (or at least not practicable) to analyze the third dimension by wide-field
89 microscopy at the resolution required for the imaging of bacteria, it is generally assumed
90 that cells are rotationally symmetric (Sieracki et al. 1989). Primary features such as area,
91 perimeter and longest chord can be directly extracted from a cell's cross-section by image
92 analysis, and secondary features (that depend on additional model assumptions) such as
93 length and width are often also calculated (Fig. 1 C). From these data, the cell volume is
94 reconstructed, typically by using simple geometrical approximations, the most common
95 being the rod model. It assumes that any cell can be reconstructed as a cylinder with two
96 hemispherical ends (Fig. 1 D).

97 This rather drastic simplification obviously does not correspond with the observed
98 variety of bacterial morphologies (Young 2006). Consequently, not all cell shapes are
99 covered by this model with equal precision. Volume estimates even of simple
100 morphologies such as prolate spheroids or diplococci are, therefore, biased. A more
101 sophisticated approach has been proposed by Sieracki et al. in 1989. It is based on the
102 cell contour, which is segmented into solids of revolution that are oriented along the major
103 axes (Fig. 1 E). This generalization allows precise reproduction of all morphologies with a
104 straight axis. Unfortunately, the assumption of a straight major axis is only valid for some
105 classes of cell shapes. Specifically, the volumes of curved morphologies such as
106 filamentous and vibrioid cells are strongly overestimated, because the segments at
107 locations where the local axis is not parallel to the major axis are reconstructed as
108 cylinders which are too large (Fig. 1 F). A solution would be to not presume a straight

109 major axis, but to segment a cell contour into triangles that are perpendicular to the local
110 axis (Fig. 1 G). These triangles can be accurately approximated as solids that are half
111 cylinders (Fig. 1 H).

112 We developed a general algorithm for microbial cell volume reconstruction from cell
113 contours that is free from *a priori* geometrical assumptions besides rotational symmetry.
114 The algorithm was numerically validated using geometrical models of 22 different bacterial
115 morphotypes. We compared our algorithm with four previously described approaches that
116 are both, theoretically sound and described in sufficient detail to be reproduced (Fry &
117 Davies 1985; Sieracki et al. 1989; Bloem et al. 1995; Blackburn et al. 1998). In addition,
118 images of fluorescent beads and a freshwater bacterial strain with high phenotypic
119 plasticity were analyzed, and critical issues with regard to fully automated imaging and
120 image processing are discussed. The algorithm is embedded in a freely available
121 computer program for batch analysis of digital images, and the source code of the
122 algorithm is provided (<http://www.technobiology.ch>).

123 **Materials & Methods**

124 ***Description of the computer program for image analysis***

125 A computer program, named YABBA (Yet Another Bacterial Biovolume Algorithm),
126 was created in VB.Net (Visual Studio 2005, Microsoft). It includes functionality to process
127 single images, or entire folders of 8-bit grayscale images in three common file formats
128 (.JPG, .TIF, .BMP). In a first step, grayscale images (Fig. 2 A) are converted into binary
129 objects (Fig. 2 B). For object detection (i.e. image segmentation), the user has the option
130 to choose between a dynamic threshold (based on background subtraction by a lowpass
131 filtered image [Gaussian filter] followed by a fixed threshold), or a simple fixed threshold,
132 allowing use of pre-processed (binary) images. In a second step, binary objects, defined
133 as groups of adjacent pixels, are analyzed (Burger 2008). A counting frame is implemented
134 (Gundersen et al. 1988), testing the bounding boxes of objects. The cell contour is of
135 particular importance for the accurate determination of cell volumes by our algorithm. From
136 the chain code contour (Fig. 2 C) (Freeman 1961), a sub-pixel contour (Fig. 2 D) is derived
137 which is then used to calculate elliptical Fourier descriptors (EFD) (Kuhl & Giardina 1982).
138 A normalized contour (EFD contour, Fig. 2 E) consisting of 200 points is reconstructed
139 from these descriptors, using a variable number of Fourier descriptors (2 – 9 harmonics),
140 depending on the perimeter (in pixel) of the object. The number of harmonics is important
141 to adequately represent the cell shape of objects consisting of different number of pixels
142 (e.g. imaged at different magnifications): too few harmonics result in an oversimplified
143 shape representation, whereas too many harmonics produce artificial contour details due
144 to the pixel discretization step (Fig. 2 F). Primary features such as area, perimeter and
145 longest chord are calculated from the EFD contour.

146 Biovolumes of each object are calculated, based on the EFD contour and its derived
147 features. In addition to our new algorithm, four described algorithms (Fry & Davies 1985;
148 Sieracki et al. 1989; Bloem et al. 1995; Blackburn et al. 1998) were also implemented for
149 comparison, subsequently referred to as Fry, Sieracki, Bloem and Blackburn (for
150 implementation details please refer to the source code in the supplementary data). Finally,
151 an output report (plain text) is created and, optionally, annotated images are generated. In
152 the single image processing mode, enlarged depictions of all single cells, including
153 detailed annotations related to the different algorithms are created.

155 ***Strategy of the algorithm***

156 The basic idea of our algorithm is that a microbial cell, independently of its
157 morphology, can be cut into in half-cylinder like segments along the local major axis (Fig. 1
158 G, H). In the two dimensional projection, as given by wide-field microscopic imaging, the
159 cell contour (subsequently also referred to as 'polygon') has to be cut into triangles. Those
160 triangles can be used for the reconstruction of solids of revolution that are half-cylinders.
161 The summation of the volumes of those volume elements finally results in the volume of
162 the microbial cell.

163

164 ***Detailed description of the algorithm***

165 The input for the algorithm to calculate the biovolume of a cell consists of its EFD
166 based contour, in the form of a list of n points (x,y -coordinates, $n = 200$ by default),
167 defining a polygon. The polygon points are listed in adjacent order. In a first step, the list of
168 simple contour points is converted into a list of more complex point objects (Fig. 2 G, for
169 better visibility, only 41 instead of 200 points are depicted), containing additional
170 information for each point such as its index I , the index of (link to) the next point (nI), the
171 previous point (pI), its coordinates (X,Y), and its normal vector (pointing inwards, X_n, Y_n)
172 which is calculated based on the coordinates of the actual point I and its two neighboring
173 points (nI, pI).

174 The second step consists of the recursive processing of the polygon (n points) in
175 order to subdivide it into $n-2$ triangles. The recursion step aims to find the optimal triangle
176 within a polygon (Fig. 2 H). The optimal triangle has to be perpendicular to the cell's local
177 axis (Fig. 2 G–I, dashed line). This is achieved by first considering all possible triangles
178 that can be constructed with two adjacent points $P1$ and $P2$ and a third point $P3$ on the
179 polygon, and selecting the most appropriate of these triangles by applying an optimization
180 function (described below). In each recursion, after the optimal triangle has been found,
181 the polygon is segmented into two new polygons by excision of this triangle (Fig. 2 I). The
182 newly defined polygons are again subjected to the recursion until subdivision is no longer
183 possible. The recursion distinguishes between two cases: a new polygon is an end region
184 of the entire contour or it is not (Fig. 2 J, K: blue and purple triangles). In the first case, the

185 polygon is consecutively divided into triangles, starting with the largest possible triangle. In
186 the second case, the optimization function is again applied and the recursion continues.

187 By using the point objects described above, it is not necessary to create new
188 polygons (i.e. new lists of points). Instead, it is sufficient to alter the links (nl, pl) between
189 the existing points at the location where the polygon is split. Since, after every subdivision,
190 one single point (Fig. 2 I, Point 36) will be a member of both of the new polygons; a new
191 point has to be added to the existing list of points. In our example, the properties of points
192 4, 5 and 36 were changed as follows: P4.nl: 5 $\rightarrow$ 36; P36.pl: 35 $\rightarrow$ 4; P4.pl: 4 $\rightarrow$ 42; and
193 the new point 42 was created with pl = 35 and nl = 5). At the end of the recursion, the
194 contour is subdivided in n-2 triangles that all are perpendicular to the local major axis (Fig.
195 2 J, K). The volume of the cell is finally reconstructed as the sum of solids of revolutions of
196 the single triangles. Depending on whether a triangle is part of an end region or not, a
197 different formula is used for the calculation (Fig. 2 J).

198 **Optimization function:** As a prerequisite for finding the optimal triangle, the longest
199 distance $D_{\max}$ between any two points within the polygon is calculated. Then, for every
200 possible triangle, consisting of a point P1 a second, adjacent point P2 and any third point
201 P3 in a polygon, an optimization value is calculated (Fig. 2 H). It is defined as the sum of
202 the two vectors $\mathbf{r}$ and $\mathbf{s}$ ($\mathbf{r}$ is the difference of the normal vector on P1 ($\mathbf{P1n}$) and the
203 normalized vector from P1 to P3 ($\mathbf{P13n}$) and $\mathbf{s}$ is the difference of the normal vector on P3
204 ($\mathbf{P3n}$) and the normalized vector from P3 to P1 ($\mathbf{P31n}$)), multiplied by the distance from P1
205 to P3 divided by $D_{\max}$. The triangle featuring the lowest value is then chosen as the optimal
206 triangle.

207

208 ***Verification of the algorithm***

209 For the verification of the algorithm, 22 geometrical models of realistic bacterial
210 morphotypes (Bergey 1994) were designed (Fig. 3 A). The models were constructed using
211 a CAD program (Pro/E Wildfire 4, PTC, Needham, USA) according to the specifications
212 given in Fig. 3 B. In addition, geometric formulae were derived for each model to express
213 its cross-section area and volume as a function of the radius r , and to calculate its volume
214 as a function of the cross-section area (Table 1). These formulae were validated
215 numerically using the CAD software. Finally, cross-section images (Fig. 3 C, $r = 10$ pixel)
216 were created from these models, which were then subjected to different cell volume

217 reconstruction algorithms from the literature (Fry, Sieracki, Bloem, Blackburn) and our own
218 algorithm. By this approach it was possible to exactly compare the volume estimates of the
219 different algorithms with the “true volume” of the models, as determined by the
220 corresponding formulae and the cross-section area measured by image analysis.

221

222 ***Analysis of fluorescent beads and bacteria***

223 Fluorescent beads of two size classes (diameter, $1.13 \pm 0.04 \mu\text{m}$ and 0.474 ± 0.012
224 μm , FlowCheck High Intensity Green Alignment, Polysciences Inc., Warrington, PA) were
225 analyzed to test the dynamic thresholding method for object detection in the YABBA
226 program and to assess the effect of exposure time and offset to the focal plane on the
227 apparent size of objects. In addition, cells from a bacterial strain with high phenotypic
228 plasticity (*Flectobacillus major* (DSMZ 103)) were imaged and their cell volumes as given
229 by the different algorithms were subsequently determined. Imaging was carried out on an
230 AxioImager.Z2 (Zeiss, Germany), using an AxioCam MRm and a 63 x Objective (Plan-
231 Apochromat, NA: 1.4). Image acquisition was partially automated using the software
232 AxioVision 4.8 (Zeiss) and the Visual Basic for Applications module therein. Bacteria were
233 filtered onto 0.22 μm pore size membrane filters (Osmonics, diameter: 25 mm) and stained
234 with DAPI (Hobbie et al. 1977; Porter & Feig 1980). Fluorescent beads were dropped onto
235 microscope slides and embedded in immersion oil after evaporation of the solvent.

236 **Exposure time series:** A single field of view (FOV) containing fluorescent beads (n
237 = 44, bead diameter: $0.474 \mu\text{m}$) was imaged at different exposure times (0 – 300 ms, Δ =
238 10 ms). The procedure was repeated on a preparation with $1.13 \mu\text{m}$ beads (n = 15). The
239 diameter of beads (calculated from the area) as well as the mean gray value and the
240 fraction of saturated pixels per bead were determined by the YABBA program.

241 **Focal plane offset series:** A single FOV containing 15 beads (diameter: $1.13 \mu\text{m}$)
242 was imaged at different offsets from the focal plane ($-2 - 2 \mu\text{m}$, Δ = $0.2 \mu\text{m}$). At every
243 location, a single image and a stacked image were acquired. Stacked images consisted of
244 seven images $0.4 \mu\text{m}$ apart from each other and were converted into a single image using
245 the wavelet-based extended depth of field algorithm (EDF) provided in AixoVision. Images
246 were processed by the YABBA program to count the beads and to measure their size
247 (calculated from the area).

248 Results

249 **Validation of the new algorithm and comparison:** The error of the new algorithm
250 was assessed by the analysis of digital images derived from well-defined geometrical
251 models of commonly observed bacterial morphologies (Figure 3). The formulae in Table 1
252 allowed for the determination of the true volume from the measured area from each model.
253 Images as depicted in Table 1 were subjected to the YABBA program that calculated cell
254 volumes of each shape by 5 different algorithms (Fry, Sieracki, Bloem, Blackburn, Zeder).
255 The error of each method is shown as the ratio of the true volume to the algorithms volume
256 in Table 1. The ratio of volumes determined by our new algorithm to the actual volumes
257 ranged from 0.98 to 1.03 (Mean = 1.00; SD = 0.01). The other algorithms performed as
258 follows: Fry: Min = 0.70; Max = 1.01; Mean = 0.96; SD = 0.08; Sieracki: Min = 0.98; Max =
259 6.30; Mean = 1.51; SD = 1.29; Bloem: Min = 0.70; Max = 1.41; Mean = 1.02; SD = 0.17;
260 Blackburn: Min = 0.71; Max = 2.06; Mean = 1.16; SD = 0.35.

261
262 **Analysis of bacterial cells:** Images of bacterial cells from a culture of
263 *Flectobacillus major* were analyzed to demonstrate the effect of cell morphology on the
264 error of different cell volume calculation methods. The estimates of the four previously
265 described methods were compared to our new method on a single cell level (Fig. 4 A-D).
266 The analyzed bacteria comprised rods of different lengths as well as filamentous and
267 curved morphologies. These morphotypes correspond to the models 3, 6-8, 12-16 and 18-
268 20 in Table 1. Eight bacterial cells with different morphologies (Fig. 4 E) were specifically
269 highlighted by individual symbols in the plots of Figure 4. As predicted by the validation
270 procedure, all analyzed bacterial morphologies gave similar results upon application of
271 Fry's algorithm and our new algorithm (Fig. 4 A). The results by the algorithms of
272 Blackburn and Sieracki however differed from ours (Fig. 4 B, C): the volumes of larger
273 filamentous cells, especially when they were curved were overestimated, as predicted in
274 Table 1. The algorithm of Bloem resulted in similar results for almost all cells with the
275 exception of two short, ellipsoid rods (Fig 4 D).

276
277 **Analysis of fluorescent beads:** Fluorescent beads of two different size classes
278 were imaged under different conditions to quantify the effect of exposure time and

279 deviation from the focal plane on size measurement (Fig. 5 A–C). The exposure time
280 series revealed that the measured sizes of beads increases with increasing exposure time
281 (Fig. 5 A, B; triangles). In addition to size, the mean gray value and the proportion of
282 saturated pixels per bead were also determined (Fig. 5 A, B; circles and squares). The
283 correct size (i.e. nominal diameter, indicated by horizontal lines in Fig. 5 A, B) was
284 obtained at the exposure time at the transition from no to few saturated pixels (indicated by
285 vertical lines in Fig. 5 A, B). In the examined range from 0 to 300 ms exposure time, the
286 determined sizes ranged between approx. 0.6 and 1.5 times the correct diameter for 0.5
287 μm beads and 0.3 to 1.3 for the 1 μm beads, respectively, corresponding to 22–338 % and
288 3–220 % of the nominal volumes. The time interval to obtain the correct diameter with ± 5
289 % error was roughly ± 10 ms.

290 The effect of deviation to the focal plane on size measurement of beads was
291 performed by imaging 1 μm beads at different offsets to the focal plane. Single images
292 (Fig. 5 C, white symbols) as well as images obtained by z-stacking and application of an
293 Extended Depth of Field algorithm (EDF, Fig. 5 C, black symbols) were analyzed. Image
294 quality decreases by increasing distance from the focal plane in both positive and negative
295 directions. Our object detection by dynamic thresholding allowed for a correct
296 determination of bead numbers within a distance of approximately - 1 to + 1 μm to the
297 focal plane for single images, and of - 1.4 to + 2 μm for the EDF image. On single images,
298 the offset range that allowed for obtaining the correct diameter with ± 5 % error was
299 between - 0.2 μm to + 0.4 μm . In the case of EDF images, the range was larger, i.e. from -
300 0.6 μm to + 1.6 μm .

301

302 Discussion

303 *Morphology independent determination of bacterial cell volumes*

304 Up to now, the rod-model represents the quasi standard for the estimation of
305 bacterial cell volumes on microscopic images (Fig. 1 A–E). Presumably this is due to its
306 conceptual simplicity, ease of implementation (that vary in details), and lack of alternatives.
307 The natural diversity of bacterial morphologies is not considered in this model, and, so far,
308 no attempts have been made to systematically investigate the effect of this variability on
309 the precision of cell volume determination. The goal of this work was to quantify this bias,

310 and to present a new algorithm that is largely unaffected by cell morphology. A crucial
311 issue therein was to create a reliable system to test the quality of any cell volume
312 calculation algorithm on specific bacterial morphologies. This was achieved by
313 constructing geometrical models of 22 commonly observed bacterial morphotypes (Fig. 3).
314 Subsequently, formulae were deduced (Table 1) that allowed calculating the true volume of
315 these morphotypes on the basis of their cross-section area. The cross-section area (in the
316 ideal case) is probably the most easy and accurate feature that can be observed when
317 bacteria are imaged by wide field fluorescence microscopy and measured by image
318 analysis. With this analytical approach we could show that four common algorithms for cell
319 volume determination are biased by cell morphology (Table 1). This was obviously due to
320 the fact that the outline of the cell is not considered, except by Sieracki's algorithm (Fig. 1
321 E). This algorithm is, however, limited by the assumption of a straight major cell axis. In
322 contrast, our algorithm is a generic solution that is based solely on the contour of a cell,
323 with no more assumptions than local rotational symmetry.

324 Microscopic images of bacterial cells usually consist of only a few pixels (Viles &
325 Sieracki 1992). Due to the discretization, their outline primarily is an artificial construct of
326 horizontal and vertical lines of 1 pixel length (or of 1.414 pixel length at 45° angles,
327 respectively). A more 'realistic' contour representation (in sub-pixel resolution) could be
328 obtained by using elliptical Fourier descriptors (Kuhl & Giardina 1982) (Fig. 2 A–G). Our
329 algorithm then recursively divides this contour into triangles that are all perpendicular to
330 the local axis of the cell. Finally the cell volume is reconstructed by summing up the
331 volumes of half-cylinders on the basis of the triangles. Our validation approach shows that
332 this algorithm is largely unbiased by different cell morphologies and that its errors are
333 minimal (< 4%) for all tested shapes (Table 1). Thus our approach will yield better overall
334 precision when samples are analyzed that feature cells with a variety of morphologies
335 (Figure 4), as is the case in many natural environments. However, our algorithm is also
336 more complicated to implement than others, and as well computationally more demanding.
337 We thus provide a stand-alone computer program as well as the documented source code
338 of the algorithm. On an average personal computer, the program is able to analyze
339 approximately 2 cells per second (image segmentation and determination of cell volumes
340 with the five algorithms), and there is potential for optimization.

341 ***Unresolved problems of bacterial cell sizing by microscopy***

342 It should be noted, that there are other issues related to the estimation of bacterial
343 biomass by microscopy that have not yet been resolved at a satisfactory level. A major
344 problem seems to be the choice of the staining procedure used to visualize a cell by
345 fluorescence microscopy. Nucleic acid stains, e.g. DAPI (4',6-diamidino-2-phenylindole) or
346 AO (acridine orange) are commonly used, but they yield different apparent cell sizes
347 (Suzuki et al. 1993; Posch et al. 2001). The same holds true when nucleic acid stains are
348 compared to protein stains (Straza et al. 2009). Moreover, fluorescent signals may also
349 emerge from other small objects, e.g. viruses, which can make up most of the particles in
350 water (Wommack & Colwell 2000; Isao et al. 1990). Irrespective of these issues, the actual
351 size measurement of fluorescently stained cells itself is far from trivial and requires proper
352 calibration. Many authors refer to fluorescent beads for this purpose, although correct
353 sizing of beads by microscopy is complicated (Gretz & Duling 1995). Moreover, beads are
354 not necessarily good models for bacterial cells with respect to fluorescence properties, as
355 the latter usually show lower and greater variance in brightness (Viles & Sieracki 1992).
356 Nonetheless, beads are the only widely available objects of known and constant size for
357 size-measurement calibration, and we, therefore, also used beads to verify our
358 thresholding method.

359 In order to demonstrate pitfalls of microscopic cell sizing that cannot be
360 compensated for by a more precise algorithm for cell volume determination, we
361 exemplarily assessed the effects of variation in exposure time and of deviation from the
362 focal plane. Both aspects are obviously of extreme importance for correct size
363 measurements (Fig. 5). Even small deviations in exposure time from the optimal value (i.e.
364 the time where the brightest pixel of a bead is just close to saturation) will cause significant
365 changes in measured bead size (Fig. 5 A, B). Moreover, any deviation from the focal plane
366 will quickly result in biased estimation of diameter that in turn will result in gross under- or
367 overestimations of cell volumes. While precise focusing in manual microscopy is less
368 problematic, it represents a critical issue in automated imaging. We have previously
369 reported that the acquisition of z-stacks and subsequent application of an extended depth
370 of field algorithm can substantially improve image quality and the precision of cell counting
371 (Zeder & Pernthaler 2009). Here we show that the same holds true for size measurement
372 (Fig. 5 C).

373 In summary, we present a novel algorithm to calculate bacterial cell volumes from
374 microscopic images, that – in contrast to other approaches – is largely unaffected by cell
375 morphology. We also present an analytical strategy for the verification of cell volume
376 algorithms that should be considered whenever new algorithms for this purpose are
377 developed or compared. It has to be stressed, that this work is strictly focused on the
378 correct volume reconstruction from the microscopic image of a cell, which is just one step
379 on the way from environmental sampling to final biomass estimates. Nevertheless, our
380 solution resolves the bias emerging from cell morphology which already may be apparent
381 in pure culture (Fig. 4), but certainly plays a role in natural microbial communities that often
382 harbor a variety of cell shapes (Young 2006). However, depending on the morphological
383 composition in a sample, e.g. if exclusively cocci and rod-shaped bacteria are present, it
384 may also be legitimate to use a simpler algorithm. In this case we recommend to apply the
385 algorithm described by Fry (Fry & Davies 1985; Fry 1990) that performed well for all but a
386 few morphologies (Table 1). By offering a free and easy-to-use stand-alone computer
387 program we hope that this method may find application in other studies.

388

389 **Acknowledgements**

390 Judith Blom is acknowledged for providing the *F. major* strain. Rudolf Amann,
391 Bernhard Fuchs and Marcel Kuypers are acknowledged for valuable discussion.

- Bergey D. (1994). *Bergey's manual of determinative bacteriology*. 9th ed. Williams & Wilkins: Baltimore.
- Bertoni R, Callieri C, Corno G, Rasconi S, Caravati E, Contesini M. (2010). Long-term trends of epilimnetic and hypolimnetic bacteria and organic carbon in a deep holo-oligomictic lake. *Hydrobiologia* 644:279-287.
- Bjørnsen PK. (1986). Automatic determination of bacterioplankton biomass by image analysis. *Appl Environ Microbiol* 51:1199-1204.
- Blackburn N, Hagstrom A, Wikner J, Cuadros-Hansson R, Bjørnsen PK. (1998). Rapid determination of bacterial abundance, biovolume, morphology, and growth by neural network-based image analysis. *Appl Environ Microbiol* 64:3246-3255.
- Bloem J, Veninga M, Shepherd J. (1995). Fully automatic determination of soil bacterium numbers, cell volumes, and frequencies of dividing cells by confocal laser scanning microscopy and image analysis. *Appl Environ Microbiol* 61:926-936.
- Boyde A, Williams RA. (1971). Estimation of the volumes of bacterial cells by scanning electron microscopy. *Arch Oral Biol* 16:259-267.
- Bratbak G. (1985). Bacterial biovolume and biomass estimations. *Appl Environ Microbiol* 49:1488-1493.
- Burger W. (2008). *Digital image processing an algorithmic introduction using Java*. Springer: New York.
- Carlson C, Ducklow H, Sleeter T. (1996). Stocks and dynamics of bacterioplankton in the northwestern Sargasso Sea. *Deep Sea Research Part II: Topical Studies in Oceanography* 43:491-515.
- Chrzanowski TH, Simek K. (1990). Prey-Size Selection by Freshwater Flagellated Protozoa. *Limnol Oceanogr* 35:1429-1436.
- Daims H, Wagner M. (2007). Quantification of uncultured microorganisms by fluorescence microscopy and digital image analysis. *Appl Microbiol Biotechnol* 75:237-248.
- Falkowski PG, Barber RT, Smetacek V. (1998). Biogeochemical Controls and Feedbacks on Ocean Primary Production. *Science* 281:200-207.
- Freeman H. (1961). On the Encoding of Arbitrary Geometric Configurations. *IEEE Trans Electron Comput* EC-10:260-268.
- Fry JC. (1990). Direct Methods and Biomass Estimation. *Methods Microbiol.* 22:41-85.
- Fry JC, Davies AR. (1985). An assessment of methods for measuring volumes of planktonic bacteria, with particular reference to television image-analysis. *J Appl Bacteriol* 58:105-112.

- Gest H. (2004). The discovery of microorganisms by Robert Hooke and Antoni Van Leeuwenhoek, fellows of the Royal Society. *Notes Rec R Soc Lond* 58:187-201.
- Gretz JE, Duling BR. (1995). Measurement uncertainties associated with the use of bright-field and fluorescence microscopy in the microcirculation. *Microvasc Res* 49:134-140.
- Gundersen HJ, Bagger P, Bendtsen TF, Evans SM, Korbo L, Marcussen N, Møller A, Mnielsen K, Nvengaard JR, Pakkenberg B, Sørensen FB, Vesterby A, West JM. (1988). The new stereological tools: disector, fractionator, nucleator and point sampled intercepts and their use in pathological research and diagnosis. *APMIS* 96:857-881.
- Hillebrand H, Durselen CD, Kirschtel D, Pollinger U, Zohary T. (1999). Biovolume calculation for pelagic and benthic microalgae. *J Phycol* 35:403-424.
- Hobbie JE, Daley RJ, Jasper S. (1977). Use of nucleopore filters for counting bacteria by fluorescence microscopy. *Appl Environ Microbiol* 33:1225-1228.
- Isao K, Hara S, Terauchi K, Kogure K. (1990). Role of sub-micrometre particles in the ocean. *Nature* 345:242-244.
- Krambeck C, Krambeck HJ, Overbeck J. (1981). Microcomputer-assisted biomass determination of plankton bacteria on scanning electron micrographs. *Appl Environ Microbiol* 42:142-149.
- Kuhl F, Giardina C. (1982). Elliptic Fourier features of a closed contour. *Computer Graphics and Image Processing* 18:236-258.
- Massana R, Gasol JM, Bjørnsen PK, Blackburn N, Hagstrom Å, Hietanen S, Hygum BH, Kuparinen J, Pedrós-Alió C. (1997). Measurement of bacterial size via image analysis of epifluorescence preparations: description of an inexpensive system and solutions to some of the most common problems. *Sci Mar* 61:397-407.
- Musat N, Halm H, Winterholler B, Hoppe P, Peduzzi S, Hillion F, Horreard F, Amann R, Jørgensen BB, Kuypers MM. (2008). A single-cell view on the ecophysiology of anaerobic phototrophic bacteria. *Proc Nat. Acad Sci U.S.A* 105:17861-17866.
- Pernthaler J, Pernthaler A, Amann R. (2003). Automated enumeration of groups of marine picoplankton after fluorescence in situ hybridization. *Appl Environ Microbiol* 69:2631-2637.
- Pernthaler J, Zöllner E, Warnecke F, Jürgens K. (2004). Bloom of filamentous bacteria in a mesotrophic lake: identity and potential controlling mechanism. *Appl Environ Microbiol* 70:6272-6281.
- Porter KG, Feig YS. (1980). The use of DAPI for identifying and counting aquatic microflora. *Limnol Oceanogr* 25:943-948.
- Posch T, Franzoi J, Prader M, Salcher MM. (2009). New image analysis tool to study biomass and morphotypes of three major bacterioplankton groups in an alpine lake. *Aquat Microb Ecol* 54:113-126.
- Posch T, Loferer-Krossbacher M, Gao G, Alfreider A, Pernthaler J., Psenner R. (2001). Precision of bacterioplankton biomass determination: a comparison of two fluorescent

dyes, and of allometric and linear volume-to-carbon conversion factors. *Aquatic Microbial Ecology* 25:55-63.

Schattenhofer M, Fuchs BM, Amann Rudolf, Zubkov MV, Tarran GA, Pernthaler J. (2009). Latitudinal distribution of prokaryotic picoplankton populations in the Atlantic Ocean. *Environ Microbiol* 11:2078-2093.

Seo EY, Ahn TS, Zo YG. (2010). Agreement, Precision, and Accuracy of Epifluorescence Microscopy Methods for Enumeration of Total Bacterial Numbers. *Appl. Environ. Microbiol.* 76:1981-1991.

Sieracki ME, Viles CL, Webb KL. (1989). Algorithm to estimate cell biovolume using image analyzed microscopy. *Cytometry* 10:551-557.

Singleton S, Cahill JG, Watson GK, Allison C, Cummins D, Thurnheer T, Guggenheim B, Gmür R (2001). A fully automated microscope bacterial enumeration system for studies of oral microbial ecology. *J Immunoassay Immunochem* 22:253-274.

Sommer U, Gliwicz ZM, Lampert W, Duncan A. (1986). The Peg-Model of seasonal succession of planktonic events in fresh waters. *Archiv Für Hydrobiologie* 106:433-471.

Straza TRA, Cottrell MT, Ducklow HW, Kirchman DL. (2009). Geographic and Phylogenetic Variation in Bacterial Biovolume as Revealed by Protein and Nucleic Acid Staining. *Appl Environ Microbiol.* 75:4028-4034.

Suzuki MT, Sherr E, Sherr BF. (1993). DAPI direct counting underestimates bacterial abundances and average cell size compared to AO direct counting. *Limnol Oceanogr* 38:1566-1570.

van Veen JA, Paul EA. (1979). Conversion of biovolume measurements of soil organisms, grown under various moisture tensions, to biomass and their nutrient content. *Appl Environ Microbiol* 37:686-692.

Viles CL, Sieracki ME. (1992). Measurement of marine picoplankton cell size by using a cooled, charge-coupled device camera with image-analyzed fluorescence microscopy. *Appl Environ Microbiol* 58:584-592.

Wommack KE, Colwell RR. (2000). Virioplankton: viruses in aquatic ecosystems. *Microbiol Mol Biol Rev* 64:69-114.

Young KD. (2006). The selective value of bacterial shape. *Microbiol Mol Biol Rev* 70:660-703.

Zeder M, Pernthaler J. (2009). Multispot live-image autofocus for high-throughput microscopy of fluorescently stained bacteria. *Cytometry A* 75A:781-788.

Zeder M, Peter S, Shabarova T, Pernthaler J. (2009). A small population of planktonic Flavobacteria with disproportionately high growth during the spring phytoplankton bloom in a prealpine lake. *Environ Microbiol* 11:2676-2686.

Zhou Z, Pons MN, Raskin L, Zilles JL. (2007). Automated image analysis for quantitative fluorescence in situ hybridization with environmental samples. *Appl Environ Microbiol* 73:2956-2962.

393 Legend to Figures:

394

395 Figure 1

396 **Different approaches for cell volume determination:** The microscopic image of a
397 three-dimensional bacterial cell (A) appears as the cross-section area (B). Computer
398 assisted image analysis delivers features such as area, contour and perimeter, wherefrom
399 features such as length and width can be derived (C). There are several approaches to
400 calculate the cell volume, e.g. the rod-model (D) or contour subdivision along the major
401 axis (E) and reconstruction as cylinders (F). The method described here performs a
402 contour subdivision into triangles perpendicular to the local axis (G) and volume
403 reconstruction as half-cylinders (H).

404

405 Figure 2

406 **Description of the algorithm:** A digital image of a bacterial cell (A) is represented
407 as a group of adjacent pixels (B). Its primary contour (C) is transformed into a sub-pixel
408 contour (D) that is used to calculate an elliptical Fourier descriptor based contour (E) by an
409 adequate number of harmonics (F). This contour is converted into a list of point objects
410 which contain important properties such as the index (I), the coordinates (X,Y), normal
411 vectors (Xn,Yn) and links to the previous (pl) and next (nl) point (G). The algorithm
412 recursively subdivides the contour into triangles using an optimization function (H). After
413 excision of the optimal triangle, the contour is divided into two contours. This is realized by
414 altering point properties (pl, nl) and adding a new point to the list (I). The resulting
415 subdivision of the cells (A) EFD contour (E) yields triangles that are perpendicular to the
416 local axis of the cell (J). The generic algorithm works for different bacterial shapes (K).

417

418 Figure 3

419 **Geometrical models of 22 bacterial morphotypes:** 22 bacterial cell morphologies
420 were modeled using a CAD software. The models are parametric and solely depend on
421 the variable r. A three-dimensional representation (A), a dimensioned drawing (B) and the

422 cross-sectional area that can be subjected to image analysis (C) are shown for each
423 model.

424

425 Figure 4

426 **Analysis of bacterial cells.** Images of bacteria from a culture were analyzed by the
427 YABBA program. The results (bacterial cell volumes (BCV)) of four published methods (A:
428 Fry, B: Sieracki, C: Blackburn, and D: Bloem) were compared to the new method
429 described here. Each data point represents a single cell. For illustration purposes, images
430 of some of the cells in the graphs (A – D) are shown (E).

431

432 Figure 5

433 **Effect of exposure time and focal plane offset on apparent cell size:**
434 Fluorescent beads (nominal size: 0.47 μm (A) and 1.13 μm (B)) were imaged at different
435 exposure times. Diameter, mean gray value and the fraction of saturated (i.e.
436 overexposed) pixels per bead were measured. Single images and image stacks were
437 acquired at different distances off the focal plane (C), and the numbers of detected beads
438 and bead diameters were assessed.

439

440 Table 1

441 **Errors of different methods:** Formulae were developed for 22 models of bacterial
442 cell morphologies, that express volume and area as a function of the radius, and, that
443 express volume as a function of the area. Images of the projected areas of these models
444 (left column) were subjected to five different methods for cell volume calculation. The
445 relative deviations from the true volume (determined by area measurement and application
446 of the corresponding formulae) are given as the ratios of the output by the respective
447 method divided by the true volume. Deviations of more than 3 % are printed in bold letters.

448

A

B

C

D

E

G

H

F

A

B

C

Table 1

		Geometrical Models of Bacteria			Errors of Different Algorithms				
#		Volume $V = f(r)$	Area $A = f(r)$	Volume $V = f(A)$	1985Fry	1989Sieracki	1995Bloem	1998Blackburn	2011Zeder
1		$\frac{4}{3} p r^3$	πr^2	$0.752 A^{\frac{3}{2}}$	0.97	1.01	1.00	0.98	0.99
2		$\frac{4}{3} + \frac{3\sqrt{3}}{4} p r^3$	$2 p - \frac{p}{3} - \frac{\sqrt{3}}{2} r^2$	$0.549 A^{\frac{3}{2}}$	0.85	0.98	0.86	0.97	0.99
3		$\frac{8}{3} p r^3$	$2 p r^2$	$0.532 A^{\frac{3}{2}}$	1.01	0.99	1.41	0.95	1.00
4		$\frac{16}{3} p r^3$	$4 p r^2$	$0.376 A^{\frac{3}{2}}$	0.98	0.99	0.99	0.94	0.98
5		$\frac{16}{3} + \frac{3\sqrt{3}}{4} p r^3$	$\frac{5 p}{3} + \frac{\sqrt{3}}{2} + 8 r^2$	$0.393 A^{\frac{3}{2}}$	0.94	1.00	0.95	0.99	1.01
6		$\frac{10}{3} p r^3$	$(\pi + 4) r^2$	$0.549 A^{\frac{3}{2}}$	1.01	1.01	1.37	0.99	1.00
7		$\frac{28}{3} p r^3$	$(\pi + 16) r^2$	$0.350 A^{\frac{3}{2}}$	1.00	1.00	1.01	1.00	1.00
8		$\frac{58}{3} p r^3$	$(\pi + 36) r^2$	$0.248 A^{\frac{3}{2}}$	1.00	1.00	1.00	1.00	1.00
9		$\frac{11}{6} + \frac{3\sqrt{3}}{4} p r^3$	$2 + \frac{\sqrt{3}}{2} + \frac{5 p}{3} r^2$	$0.427 A^{\frac{3}{2}}$	0.83	0.98	0.83	0.93	1.02
10		$\frac{11}{3} p r^3$	$(\pi + 6) r^2$	$0.417 A^{\frac{3}{2}}$	0.89	0.98	0.90	0.96	1.02
11		$\frac{46}{3} p r^3$	$\frac{5 p}{2} + 12 r^2$	$0.545 A^{\frac{3}{2}}$	0.95	1.00	1.38	0.95	1.03
12		$3 p^2 + \frac{4 p}{3} r^3$	$7 p r^2$	$0.328 A^{\frac{3}{2}}$	0.99	1.09	0.99	1.08	0.99
13		$6 p^2 + \frac{4 p}{3} r^3$	$13 p r^2$	$0.243 A^{\frac{3}{2}}$	0.99	1.62	0.99	1.49	0.99
14		$3 p^2 + \frac{40 p}{3} r^3$	$(7 p + 24) r^2$	$0.229 A^{\frac{3}{2}}$	1.00	6.30	1.00	2.05	1.00
15		$3 p^2 + \frac{40 p}{3} r^3$	$(7 p + 24) r^2$	$0.229 A^{\frac{3}{2}}$	1.00	1.14	1.00	1.12	1.00
16		$\pi^2 + \frac{40 p}{3} r^3$	$(4 p + 24) r^2$	$0.256 A^{\frac{3}{2}}$	1.00	1.30	1.00	1.28	1.00
17		$(14 - \sqrt{3}) p - \frac{4}{3} + p \sqrt{3} r^3$	$\frac{3 p}{2} + 24 - \sqrt{3} r^2$	$0.297 A^{\frac{3}{2}}$	1.00	2.06	1.00	1.57	1.01
18		$\frac{3 p^2}{16} + \frac{p}{6} r^3$	πr^2	$0.426 A^{\frac{3}{2}}$	0.99	1.00	1.00	1.00	0.99
19		$\frac{3 p^2}{2} + \frac{4 p}{3} r^3$	$4 p r^2$	$0.426 A^{\frac{3}{2}}$	0.99	1.09	1.00	1.07	0.99
20		$3 p^2 + \frac{4 p}{3} r^3$	$7 p r^2$	$0.328 A^{\frac{3}{2}}$	0.99	1.52	0.99	1.44	0.98
21		$\frac{9 p^2}{2} + \frac{4 p}{3} r^3$	$10 p r^2$	$0.276 A^{\frac{3}{2}}$	0.99	4.26	0.99	2.06	0.98
22		$(100\sqrt{6} + 1312 r - 245) \frac{p}{375} r^2$	$\frac{\sqrt{3}}{100} + \frac{377 p}{150} + \frac{8}{5} r^2$	$0.374 A^{\frac{3}{2}}$	0.70	0.98	0.70	0.71	1.00