1 **A geometrical model generator for quasi-axisymmetric biological products**

2

3 Seppe Rogge ^a , Shiferaw D. Beyene ^a , Els Herremans ^a , Maarten L Hertog ^a , Thijs Defraeye ^a , Pieter
4 Verboven ^a , Bart M. Nicolai ^{a,b*}

5

6 ^a *MeBioS, Department of Biosystems, University of Leuven, Willem de Croylaan 42, 3001 Heverlee, Belgium*

7 ^b *VCBT, Flanders Centre of Postharvest Technology, Willem de Croylaan 42, 3001 Heverlee, Belgium*

* Corresponding author. Address: *MeBioS, Department of Biosystems, University of Leuven, Willem de Croylaan 42, 3001 Heverlee, Belgium*. Tel.: +32 16 322 375
E-mail address: bart.nicolai@biw.kuleuven.be

8 **Abstract**

9 A geometrical model generator for biological products is presented, which uses X-ray computed tomography
10 images of quasi-axisymmetric biological products as input. It was tested with a dataset of 73 scanned
11 Braeburn apples. For each sample, the generator constructed different cross sections. From these sections,
12 contours were extracted and selected. The contours were expressed as a series of shape descriptors. For this
13 purpose, elliptical Fourier descriptors were used. The obtained frequency distributions were transformed to
14 standard normal distributions. On these transformed distributions, the covariance decomposition algorithm
15 was applied. This algorithm generated new sets of descriptors, which opened up a large range of possibilities
16 for generation of representative shape contours. After reverse transformation of the (generated) descriptor
17 distributions, new contours were obtained from the new descriptors. These new contours were converted to
18 3D geometrical models of biological products by interpolation and revolving. By comparing the volumes of
19 the generated models with those of the scanned fruit, it was shown that the resulting geometrical models
20 have the same variability as the biological variability in the original dataset. This generator is a fast method,
21 which requires minimal user intervention, and creates 3D models including the biological variability as
22 observed in the scanned fruit. Because these 3D geometrical models are directly available as CAD models,
23 they are useful for numerical modelling of transport phenomena in and around biological products.

25 **Keywords**

26 elliptic Fourier descriptors; shape description; shape generation; biological variability; virtual fruit

28 **1. Introduction**

29 Structure, shape and size are features of plant organs which have an impact on gas and water transport within
30 the organs and their exchange with the environment, in relation to physiological processes such as growth
31 (Abera et al., 2012), respiration (Ho et al., 2011) and photosynthesis (Ho et al., 2012), as well as to
32 postharvest quality of plant products such as fruit (Nguyen et al., 2006; Veraverbeke et al., 2003a, 2003b).
33 Numerical modelling is often used to analyse such processes (Ho et al., 2013). Examples include convective
34 cooling of produce (Dehghannya et al., 2010; Nicolai et al., 1999; Verboven et al., 1997, 2006), the non-
35 destructive analysis of firmness of fruit (Jancsó et al., 2001), convective drying of food (Kaya et al., 2006),
36 long-term storage of horticultural products (Ambaw et al., 2012; Delele et al., 2008, 2009), and individual
37 quick freezing (Peralta et al., 2010). These numerical methods need a geometrical model of the studied
38 object but no standard geometrical models exist for biological objects. The main reasons are that the shape of
39 biological products is often complex and rather difficult to acquire and describe, and that a large variability is
40 found between species and even cultivars. Though modelling techniques and software already exist to
41 generate models of some kinds of plants and plant organs (Pradal et al., 2009; Prusinkiewicz & Runions,

2012), they are not yet available to generate models of common types of fruit and vegetables. Hence, their complex shapes are often simplified to basic geometries, like ellipsoids (Rashidi & Gholami, 2008) or spheres (Ambaw et al., 2012; Delele et al., 2008, 2009).

There are two steps in generating realistic models of biological products: image acquisition and model generation from the obtained image. Image acquisition can be done with both destructive and non-destructive techniques. Destructive techniques require product manipulation (e.g., Mebatsion et al., 2011), which can induce deformation of the produce. This is particularly the case for soft products, such as tomato. Non-destructive techniques can be split in two types: fruit can simply be observed with multiple cameras from the sides (wire frame models, e.g. Imou et al., 2006; Jancsó, 2001), but this technique fails to capture concave regions in fruit. The other kind of non-destructive techniques are 3D methods that also reveal the inner fruit structure, such as magnetic resonance imaging (MRI) or X-ray computed tomography (CT) (Herremans et al., 2013; Lammertyn et al., 2003).

Regarding model generation from the images, different strategies exist. For 3D imaging, complex and realistic 3D geometric models can be directly constructed from the data using reverse engineering software (Goñi et al., 2007, 2008; Moustakides et al., 2000). However, the procedure is quite time-consuming, even with specialized software. If hundreds of different products have to be processed, e.g., to fill a virtual container (Ambaw et al., 2012), this option is not realistic. In most cases, 3D models are generated based on 2D contours, where the most important step is the description of the shape of the contour. Many different methods for 2D shape description exist. Overviews can be found in Costa et al. (2011), Moreda et al. (2012), and Zhang & Lu (2004). Applications include analysis during dehydration, automatic grading, and quantitative evaluation based on shape (Fernández et al., 2005; Liming & Yanchao, 2010; Saad et al., 2011; Tanabata et al., 2012).

Fourier descriptors (FD) are very popular shape descriptors and are mainly used in a 2D shape description context. The general idea is to create a 1D function from a 2D boundary contour of for example the exterior shape of a fruit: the shape signature. An example is the centroid distance signature which is a (periodic) function that represents the distance from the boundary to the centroid of the image. This shape signature can be approximated by a Fourier series, where the obtained coefficients are called Fourier descriptors. The larger the set of derived descriptors, the better the accuracy for shape retrieval will be. The advantages of FD are (i) their low complexity, (ii) each descriptor has a physical meaning, (iii) they can easily be normalised, and (iv) they describe shape features at all scales (Zhang & Lu, 2004, 2005).

Some variations on FD exist. An example are Elliptic Fourier descriptors (EFD), introduced by Kuhl & Giardina (1982). EFD describe a closed contour with a series of rotating phasors with elliptical loci; the contour is hence represented as a set of harmonically related ellipses. Elliptical Fourier analysis removes three of the limitations encountered in conventional Fourier analysis: (i) the sampled interval has to be equally divided, (ii) the descriptors depend on the chosen coordinate system, and (iii) the difficulty of dealing with outlines that curve back on themselves (Lestrel, 1989). The drawback of EFD is that many

descriptors have to be used, as each harmonic (ellipse) consists of four descriptors. However, this is not really of concern, because the harmonics are computed fairly easy and fast. EFD have often been used for describing shape variation of biological products, ranging from rice to petals or a stallions sperm heads (e.g. Iwata et al., 2010; Kawabata et al., 2009; Severa et al., 2010).

Based on aforementioned image acquisition and shape description techniques, different methods have been explored concerning 3D model generation of biological materials. 3D models of complex food shapes have been constructed based on images of slices of the objects (Goñi et al., 2007), and based on MRI scans (Goñi et al., 2008; 2010). Mebatsion et al. (2011) described the contour of sections of plant organs with Fourier descriptors and interpolated different contours to construct 3D models of the plant organs. The more complex the shape, the higher the number of contours used for one image. Lu et al. (2007, 2009) presented a method to construct fruit models by searching a mathematical function to approximate the shape (starting from ellipsoids) and by applying random perturbations. This method introduces random variations in the dataset, though these variations were not connected with measured variations in fruit shapes. Image acquisition combined with shape description analysis on a set of biological products, such as fruit, opens perspectives to automatic generation of random shapes, based on statistical analysis of the dataset, but was not extensively explored. The resulting database of biological products would allow introducing biological shape variability, as present in reality, into numerical simulations. This has been explored by Jancsó et al. (1997), but this work did not use the benefits of EFD, was not able to capture concave parts of fruit (and therefore not able to generate models of fruit with concave parts such as apples), and required normal distributions for all shape descriptors.

In this study, we present a new automatic generator of fruit shapes. The inputs of the generator were X-ray CT scans. From the images of each fruit, a single contour was extracted. These contours were described with elliptic Fourier descriptors. From the database with described contours, new random contours were generated, based on statistical analysis of the obtained descriptors. From each new contour a 3D model was constructed by revolving and interpolating the two halves of the contour. This generator is a fast method, requiring little manual intervention. It is able to create a large amount of 3D models of quasi-symmetric fruit that have the same biological variability as the original set of scanned fruit. The generator was validated by testing it with a set of 73 scanned Braeburn apples, and a set of 37 Conference pears.

2. Materials and methods

2.1 Fruit

A dataset consisting of full-object CT scans of 73 apples (cv. 'Braeburn') was acquired to test the geometrical model generator. All apples were grown on a field of the Centre of Fruit Growing (pcfruit, Sint-Truiden, Belgium). 18 apples were harvested on 26th of October 2009, the others on 21st of October 2011, which was in the optimal commercial picking window for 'Braeburn' apples, determined by the Flanders

113 Centre of Postharvest Technology (VCBT, Belgium). The first 18 apples were scanned on the 17th and 26th
114 of October 2009, the others in February and March 2012.

115 A second dataset consisted of 37 pears (cv. ‘Conference’), also grown at pcfruit (Sint-Truiden, Belgium). 10
116 pears were harvested on 24th of September 2009 and scanned on 26th of November 2009; the other 27 pears
117 were harvested on 2nd of September 2011 and scanned in October and November 2011. These harvest dates
118 are also in the optimal commercial picking window.

119 **2.2 Image acquisition**

120 X-ray CT was used to acquire the images required for the geometrical model generator. This technique has
121 the advantage that a suitable cross section can be selected after the object has been scanned in 3D.

122 X-ray CT is an imaging method with which cross-sectional slices can be obtained, representing the X-ray
123 attenuation properties of the fruit. The X-ray CT scans were made on a microfocus X-ray CT (AEA
124 Tomohawk, Philips, The Netherlands) using a Philips HOMX 161 X-ray source, operating at 85 kV and 0.41
125 mA for the first 18 apples and first 10 pears, and at 75 kV and 0.463 mA for the others. The size of the
126 voxels in the images ranged from 81 to 106 μm for the apples and from 110 to 138 μm for the pears. 3D
127 information was obtained by rotating the X-ray source in discrete steps over an angle of 180°. The first 18
128 apples and first 10 pears were scanned with an angular increment of 0.5 degrees, the others with an
129 increment of 0.3 degrees.

130 **2.3 Geometrical model generator**

131 The geometrical model generation described below, was entirely coded in Matlab (The MathWorks Inc.,
132 Natick, MA). The algorithm was optimised to reduce the amount of necessary manual intervention to a
133 minimum to make the procedure fast and user friendly. All computations were performed on a computer
134 with an Intel Core2 Quad Q9650 @ 3.00 GHz processor with 8 GB RAM.

135 **2.3.1 Shape extraction and description**

136 For each scanned object, a 2D cross section was extracted. In order to avoid artefacts in the resulting 3D
137 models, the 2D cross section needed to be quasi-symmetric. Large stems frequently occurred in the cross
138 sections, disrupting the symmetry. Therefore, eight different cross sections were automatically constructed
139 from the acquired 3D stack. The contours of these sections were automatically extracted by Matlab’s built-in
140 edge detection and boundary tracing routines. Afterwards, one of the eight contours was selected manually to
141 avoid large stems in the dataset. Subsequently, the chosen contours were expressed in an appropriate
142 coordinate system: the origin was positioned at the centroid of the contour, and the z-axis was defined to
143 coincide with the stem-calyx axis (all mentioned coordinates are cylindrical coordinates). An example of a
144 contour in this coordinate system is shown in Figure 1.

145 EFD were used to describe the shape of the contours. The iterative process used to compute these descriptors
146 can be found in Kuhl & Giardina (1982). From a sensitivity analysis, 25 harmonics were found to be
147 sufficient to describe the contours. This is visualised in Figure 2: the yellow line in this figure (25 harmonics)
148 almost completely coincides with the red (100 harmonics) and black (original contour) lines.

149 **2.3.2 Random generation of new contours**

150 The EFD statistics were used for random generation of new 2D contours. These new contours should have
151 the same variability as that of the original dataset. Therefore, not only the average value and variance of each
152 descriptor should match, but also the covariance between the different descriptors. If the descriptors follow
153 Gaussian distributions, new sets of descriptors can be generated using the covariance decomposition
154 algorithm (Rubinstein, 1981). Using this method, first the covariance matrix $\mathbf{V}$ was Cholesky decomposed as
155 $\mathbf{V} = \mathbf{L}\mathbf{L}^T$ where $\mathbf{L}$ is a lower triangular matrix and $\mathbf{L}^T$ is its transpose. It can be shown that a new set of
156 descriptors $\mathbf{Y}$, with the same average and covariance as the original data, can be generated as $\mathbf{Y} = \mathbf{L}\mathbf{Z} + \boldsymbol{\mu}$,
157 where $\mathbf{Z}$ is a vector with random numbers (Gaussian distribution) and $\boldsymbol{\mu}$ is a vector with the average of each
158 descriptor.

159 The normality of the descriptor distributions was tested with Lilliefors tests (Lilliefors, 1967). When these
160 tests detected significant deviations from normal distributions, the distributions of the descriptors were
161 transformed to the standard normal distribution using a proposed SKN-distribution (a distribution function
162 that can exhibit both skewness (S) and kurtosis (K), based on the standard normal distribution (N); Hertog et
163 al., 2009). In that case, the covariance decomposition method was applied on the transformed distributions,
164 and by performing the inverse transformation on the generated descriptors, new descriptors representing the
165 original distributions were also obtained. Two-sample Kolmogorov-Smirnov tests (Massey, 1951) were used
166 to check if the distributions of the generated descriptors were the same as the original distributions. For this
167 purpose, 10000 new sets of EFD were generated.

168 Next, new contours were obtained by reversely transforming each new set of harmonics from the Fourier
169 domain. The same coordinate system as the one for the original data was applied for the new contours.
170 Contours having more than two intersections with the z -axis were rejected and replaced with newly
171 computed contours. Otherwise the resulting geometrical models of these contours would have had
172 discontinuities after revolving. This did not occur often (i.e. $< 0.5\%$ of the contours); mostly it was caused by
173 an odd-shaped calyx.

174 **2.3.3 3D shape construction**

175 3D geometric models were generated from the new, randomly-generated 2D contours. Because the contours
176 were not symmetric, interpolation between the right and the left half of the contour was necessary when
177 revolving the contour. This way, the constructed 3D models were not entirely axisymmetric. Instead of
178 interpolating between the coordinates of the points of the left and the right half of the contour, interpolation
179 of the descriptors representing the two sides of the contour was preferred.

180 In practice, first the EFD of the mirror image of the contour along the axis of (quasi)symmetry was
 181 calculated. Subsequently, a number of new sets of descriptors (50) were generated by interpolating the
 182 values of these new descriptors and those of the original contour. The interpolation could be done linearly.
 183 However, with this kind of interpolation, a sharp edge appeared at the position of the largest side of the
 184 original contour after constructing a 3D model. This edge was caused by the (too) rapid change in size of the
 185 half contour when interpolating. Therefore, it was chosen to add more sets of descriptors with values close to
 186 the descriptors of the original contour or its mirror image. This way, the 3D geometrical model became
 187 smoother around the original contour. For this purpose, a sine and a cosine factor were added in the
 188 interpolation formula. In order to preserve the size of the contour, the added factors were squared:

$$189 \quad EFD_i = EFD \times \cos^2(i) + EFD_{mir} \times \sin^2(i) \quad (1)$$

190 In this formula, the increment i ranges linearly from 0 to $\pi/2$, so that the 50 new sets of descriptors (EFD_i)
 191 range from the original set (EFD) to the set of descriptors of the mirrored contour (EFD_{mir}). After the
 192 interpolation, each set of descriptors was converted to a contour.

193 The computed contours were rotated around the z -axis (θ -direction) at an angle between 0 and π , so that 3D
 194 wire frames of the fruit were obtained. Next, a NURBS (non-uniform rational B-spline) surface was fitted to
 195 the points of each wire frame, creating 3D geometrical models. These surfaces were exported as IGES files,
 196 which can be directly imported as CAD models in commercial software for finite element modelling or
 197 computational fluid dynamics.

198

199 **3. Results**

200 **3.1 Geometrical model generator**

201 For each apple in the database, a suitable contour was extracted and transformed to EFD. The result was a
 202 database with 100 descriptors for each of the 73 contours. The (frequency) distributions of the descriptors of
 203 the first 6 harmonics are shown in Figure 3. Lilliefors tests confirmed that not all the descriptors were
 204 normally distributed: ~20% of the p-values were below 0.05. A histogram of the obtained p-values is shown
 205 in Figure 4a. Therefore, all descriptor distributions were transformed with SKN-distributions before applying
 206 the covariance decomposition algorithm to generate new sets of EFD. The resulting p-values of the two-
 207 sample Kolmogorov-Smirnov tests are shown in Figure 4b. These tests revealed that only one descriptor had
 208 a new distribution significantly different from the original distribution (descriptor b_4 in Figure 3; $p=0.033$).

209 For each randomly generated contour, 50 additional contours were created with the described interpolation
 210 method. An example (with only 15 contours for clarity) can be seen in Figure 5. It should be noted that this
 211 interpolation method did not put any constraints on the position where the new contours intersected with the
 212 z -axis. In Figure 5 can be seen that these intersections are slightly different for the different contours. In order

213 to avoid discontinuities in the resulting 3D models, corrections were applied, putting all contours at the top
214 or bottom together in the same point (the position of the original contour). From this point, the contours were
215 adjusted so that they returned rapidly, but smoothly, back to their original outlines. Sometimes, the necessary
216 corrections were quite large, in some cases leading to unrealistic artefacts in the 3D model. For this reason, if
217 a contour needed corrections larger than a chosen threshold (1/10 of the maximum value of the original
218 contour in the z -direction), the whole set of contours was automatically omitted, and a new contour was
219 randomly generated. This happened quite often (~25% of the models).

220 3D wire frame models were constructed by placing the different contours in a set at different angles. An
221 example is shown in Figure 6, where again only 15 contours are shown for clarity. NURBS surfaces were
222 fitted to these wireframes, and they were subsequently exported as IGES files. As such, they could directly
223 be used as CAD input in numerical software for modelling transport and physiological processes. In this
224 software, the user first has to create a mesh of the geometrical model. Some examples of surface meshes,
225 created with ANSYS ICEM CFD 13.0 (Ansys, Canonsburg, PA), can be found in Figure 7.

226 **3.2 Volume comparison**

227 To evaluate the quality of the constructed geometric models, the volumes of the 73 apples in the dataset and
228 of a set of 10000 new, randomly generated models were calculated and compared. The volumes of the
229 original apples were determined directly from the CT scan images: the number of pixels with a value higher
230 than a preset threshold, which delineated the apple, was counted and multiplied with the 3D voxel size. This
231 threshold was stable: when raising or lowering the threshold with 10%, the calculated volume changed with
232 less than 0.3%. To obtain the volume of the generated models, each model was considered to be a
233 compilation of 100 slices (the revolved half-contours). The volumes of these slices were calculated and
234 summed. From these calculations, the volumes of all apples were obtained. They are shown in a histogram in
235 Figure 8a. In this histogram can be seen that the volumes of the scanned apples are on average slightly lower
236 than those of the generated models, while the spread of the distributions looks similar. This was confirmed
237 by the statistics of the two distributions: the scanned apples had an average volume of 208 cm³ with a
238 standard deviation of 31 cm³, while the generated models had an average volume of 217 cm³ with a standard
239 deviation of 30 cm³.

240 The flexibility of the generator to other quasi-axisymmetric species was tested by applying the generator on a
241 dataset of scanned Conference pears. Only one adjustment to the generator was necessary to produce pear
242 shapes: the rotating algorithm, aligning the stem-calyx axis with the z -axis was omitted, because this
243 function can only handle apple shapes. However, this routine is not needed for pears, because they generally
244 are more asymmetric and have much smaller concave regions than apples. Surface meshes of resulting
245 geometrical models can be seen in Figure 9. The same volume comparison as for the apples was also
246 performed for the pear models. The results are shown in a histogram in Figure 8b. The 37 pears had an
247 average volume of 212 cm³ with a standard deviation of 38 cm³, while 10000 new models had an average
248 volume of 214 cm³ with a standard deviation of 38 cm³.

249 3.3 Computation time

250 The created generator uses a two-step procedure for generating geometrical models with biological
251 variability. The first part takes CT scan images as input and returns statistical data on Fourier descriptors (the
252 covariance matrix and parameters necessary for the inverse transformation of the distribution of each Fourier
253 descriptor). The total CPU-time necessary for executing this part on the dataset of 73 apples was 1.3×10^4 s.
254 This does not include the time necessary for manually selecting suitable contours. Next to this manual
255 intervention, the most time consuming step was the generation of a 2D cross section (loading the images).
256 Once the output of the first part is stored, an arbitrary amount of new models can be generated using the
257 second part of the generator. Because this part is fully automatic, it was rather fast: the CPU-time necessary
258 for computing 1000 models was 2.4×10^3 s. Here, the most time consuming step was the conversion of the
259 NURBS surface into an IGES file. In conclusion, the total time to process the dataset of 73 apples into 1000
260 geometrical models, including manual intervention, was about 5 hours, which is very fast.

261

262 4. Discussion

263 Because of the large size of the CT images, the 2D cross section generation stage can be rather time
264 consuming, certainly because eight cross sections were constructed for each scanned fruit. However, this
265 stage is fully automatic thus does not ask a lot of time of the user. The contour extraction phase is the only
266 phase that requires manual intervention: the user has to select one of the eight extracted contours. Next to the
267 ability to avoid asymmetric features such as stems, the manual intervention also alerts the user of the
268 occasional situations in which the edge detection or boundary tracing routines failed. This happens either
269 when the quality of the dataset is not high enough (e.g., tomography constructed with a large angular
270 increment; this was not of concern for the current datasets), or when the observed fruit is not completely in
271 the image. The EFD of the selected contours are calculated easily, and therefore, shape description is a fast
272 part of the process. Though other shape descriptors would also have been appropriate, Elliptic Fourier
273 descriptors were chosen, mainly because each descriptor has a physical meaning, and because the sampled
274 points do not have to be at equal distances of one another. Random contour generation also was a fast part of
275 the process. Transforming the distributions of the descriptors to normal distributions significantly improved
276 the results.

277 In the final stage, a 3D model was constructed. For this purpose, it was chosen to interpolate between the
278 EFD of the contour and those of the mirror image of the contour. In most cases the differences between this
279 interpolation, and the more obvious way of interpolating between the points of the right and the left half of
280 the contour were subtle. However, sometimes the latter method gave less realistic results. When a small stem
281 appeared in the contour, for example, interpolating between points spread the stem out, while through
282 interpolating between descriptors, this feature could be confined more locally. Converting the contour back
283 to Fourier descriptors for the interpolation and then reconstructing contours may appear inefficient.

284 However, because this routine is fully automatic and because converting a contour to descriptors or vice
285 versa is by far not the most time consuming part of the process, the additional complexity is not really of
286 concern.

287 The generated models were created from an entire contour. Consequently, they were not axisymmetric, as
288 variations in the θ -direction occurred. These variations were however limited: the variations were the same
289 in both directions from one half of the contour to the other (from $\theta=0$ to π , or from $\theta=0$ to $-\pi$; the $y=0$ -plane
290 in Figure 6 is a mirror plane). The variability in the θ -direction was also limited in size: the variations were
291 global (from one side to the other); local variations did not occur. Despite these limitations in variation, the
292 resulting models were still realistic. Apples are known to be quasi-axisymmetric, so not much variation in
293 the θ -direction was expected. In addition, the variation in the z - and r -direction was detailed enough to create
294 realistic fruit models.

295 Therefore, the generator mostly produced models that appear realistic. However, it should be noticed that
296 from time to time (~2% of the models) the random generator failed to create a contour with typical apple-like
297 features. This can be explained by the relatively large spread in shape features in this dataset: the generator
298 combined different apple features in an unrealistic way.

299 Because an important part of the generated models was omitted due to the necessity of large continuity
300 corrections (~25% of the models, as mentioned in section 2.3.3), one could argue that the spread in
301 biological variability was also altered. However, Figure 8 shows that both for apples and pears, the
302 distributions of the volumes of the scanned fruit are similar to those of the created models. This agreement
303 proves that the size variations of the new models still resemble those of real fruit. An agreement in size
304 variations in turn is a strong indication that the biological variability in shape is also preserved.

305 The presented geometrical model generator is a stable algorithm. It is fast, requires little manual intervention,
306 and introduces biological variability in the created models. In practise, for most purposes, one should be able
307 to create enough models in about a day's work, when a decent dataset (~between 30 and 100 scanned
308 objects) of CT scanned objects is available. It was designed to create geometrical models of apples. The
309 good resemblance between the volumes of the 37 Conference pears and the pear geometrical models shows
310 that the generator also works fine for smaller datasets and for other quasi-axisymmetric fruit. In fact, the
311 generator can easily be optimized for other quasi-axisymmetric biological products, such as nuts, berries,
312 oranges or carrots. Furthermore, when revolution around a curved instead of a straight axis would be
313 considered, model generation of biological products such as bananas and squash could be considered.

314 Instead of describing one cross section of an object the generator will in the future also be adapted to use
315 multiple edges of the 3D full object scans. The complete outer surface of the object can then be described
316 with a 2D Fourier series. This would enable the generator to use all available shape information of the
317 scanned objects and to produce geometrical models of asymmetric objects. Furthermore, instead of the outer
318 surface, this method could also be applied to internal features such as the apple core.

319

320 **5. Conclusions**

321 The generator presented here is a fast algorithm to create geometric models of biological materials. From a
322 representative set of CT scan images of a certain fruit cultivar, different cross sections were constructed.
323 Next, the contours of these cross sections were extracted, and the user could then choose one contour for
324 each scanned object. In order to apply statistical analysis on the features of these extracted contours, the
325 contours were transformed into a series of numbers (elliptic Fourier descriptors). From the statistical
326 analysis, new contours were constructed and from these new contours, 3D geometrical models were created.

327 The first stages of the generator (from the image acquisition till the statistical analysis) can be quite time
328 consuming, among others since it includes some manual intervention. However, the generator can be
329 regarded as fast, because this task only needs to be done once for a particular species or cultivar. Afterwards,
330 the statistical data are stored and new models can be generated quickly. As the generator's Matlab routines
331 are user friendly and straightforward, a new user should be able to process a dataset of less than 100 objects
332 in a day's work. Processing the studied dataset of 73 Braeburn apples, for example, was done in about 4
333 hours. Generating 1000 models from the obtained results, took less than one hour.

334 The constructed geometrical models show the same biological variability as the original dataset, as proved
335 by a good comparison of the volumes of the generated models and the original fruit. Though most of the
336 created models are of good quality, some have artefacts that do not appear realistic. Most of these models are
337 already omitted in the generator. However, because the generator is not able to detect all of them, the user
338 will have to select the models he/she regards being fit for use. These models can directly be used as a CAD
339 input in software for numerical modelling of transport and physiological processes in or around apples or
340 other biological products. They can replace the commonly-used simplified geometries such as spheres and
341 ellipsoids in numerical calculations based on finite element modelling or computational fluid dynamics,
342 which could improve the accuracy and realism of the simulation results considerably.

343

344 **Acknowledgements**

345 Financial support of FWO Vlaanderen (project G.0645.13) and the University of Leuven (project OT
346 12/055) is gratefully acknowledged. Thijs Defraeye is postdoctoral fellow of FWO and acknowledges its
347 support.

References

- Abera MK, Fanta SW, Verboven P, Ho QT, Carmeliet J & Nicolai BM (2012) Virtual Fruit Tissue Generation Based on Cell Growth Modelling. *Food and Bioprocess Technology*, in press DOI: 10.1007/s11947-011-0775-4.
- Ambaw A, Verboven P, Delele MA, Defraeye T, Tijskens E, Schenk A & Nicolai BM (2012) CFD Modelling of the 3D Spatial and Temporal Distribution of 1-methylcyclopropene in a Fruit Storage Container. *Food and Bioprocess Technology*, in press DOI: 10.1007/s11947-012-0913-7.
- Costa C, Antonucci F, Pallottino F, Aguzzi J, Sun DW & Menesatti P (2011) Shape analysis of agricultural products: A review of recent research advances and potential application to computer vision. *Food and Bioprocess Technology*, 4(5), 673-692.
- Dehghannya J, Ngadi M & Vigneault C (2010) Mathematical modeling procedures for airflow, heat and mass transfer during forced convection cooling of produce: A review. *Food Engineering Reviews*, 2(4), 227-243.
- Delele MA, Tijskens E, Atalay YT, Ho QT, Ramon H, Nicolai BM & Verboven P (2008) Combined discrete element and CFD modelling of airflow through random stacking of horticultural products in vented boxes. *Journal of Food Engineering*, 89(1), 33-41.
- Delele MA, Schenk A, Tijskens E, Ramon H, Nicolai BM & Verboven P (2009) Optimization of the humidification of cold stores by pressurized water atomizers based on a multiscale CFD model. *Journal of Food Engineering*, 91(2), 228-239.
- Fernández L, Castellero C & Aguilera JM (2005) An application of image analysis to dehydration of apple discs. *Journal of Food Engineering*, 67(1), 185-193.
- Goñi SM, Purlis E & Salvadori VO (2007) Three-dimensional reconstruction of irregular foodstuffs. *Journal of Food Engineering*, 82(4), 536-547.
- Goñi SM, Purlis E & Salvadori VO (2008) Geometry modelling of food materials from magnetic resonance imaging. *Journal of Food Engineering*, 88(4), 561-567.
- Goñi SM & Purlis E (2010) Geometric modelling of heterogeneous and complex foods. *Journal of Food Engineering*, 97(4), 547-554.
- Herremans E, Verboven P, Bongaers E, Estrade P, Verlinden BE, Wevers M, Hertog ML & Nicolai BM (2013) Characterisation of 'Braeburn' browning disorder by means of X-ray micro-CT. *Postharvest Biology and Technology*, 75, 114-124.
- Hertog ML, Scheerlinck N & Nicolai BM (2009) Monte Carlo evaluation of biological variation: Random generation of correlated non-Gaussian model parameters. *Journal of Computational and Applied Mathematics*, 223(1), 1-14.
- Ho QT, Verboven P, Verlinden BE, Herremans E, Wevers M, Carmeliet J & Nicolai BM (2011) A three-dimensional multiscale model for gas exchange in fruit. *Plant Physiology*, 155(3), 1158-1168.
- Ho QT, Verboven P, Yin X, Struik PC & Nicolai BM (2012) A microscale model for combined CO₂ diffusion and photosynthesis in leaves. *PloS one*, 7(11), e48376.
- Ho QT, Carmeliet J, Datta AK, Defraeye T, Delele MA, Herremans E, Opara L, Ramon H, Tijskens E, van der Sman R, Van Liedekerke P, Verboven P & Nicolai, BM (2013) Multiscale modeling in food engineering. *Journal of Food Engineering*, 114(3), 279-291.
- Imou K., Kaizu Y, Morita M & Yokoyama S. (2006) Three-dimensional shape measurement of strawberries by volume intersection method. *Transactions of the ASAE*, 49(2), 449-456.
- Iwata H, Ebana K, Uga Y, Hayashi T & Jannink JL (2010) Genome-wide association study of grain shape variation among *Oryza sativa* L. germplasms based on elliptic Fourier analysis. *Molecular Breeding*, 25(2), 203-215.

390 Jancsó PT, Nicolai BM, Coucke P & Baerdemaeker JD (1997) 3D finite element model generation of fruits based on
 391 image processing. *Mathematical and control applications in agriculture and horticulture*, 131-135.

392 Jancsó PT, Clijmans L, Nicolai BM & De Baerdemaeker JD (2001) Investigation of the effect of shape on the acoustic
 393 response of ‘conference’ pears by finite element modelling. *Postharvest Biology and Technology*, 23(1), 1-12.

394 Kawabata S, Yokoo M & Nii K (2009) Quantitative analysis of corolla shapes and petal contours in single-flower
 395 cultivars of *lisianthus*. *Scientia Horticulturae*, 121(2), 206-212.

396 Kaya A, Aydın O & Dincer I (2006). Numerical modeling of heat and mass transfer during forced convection drying of
 397 rectangular moist objects. *International Journal of Heat and Mass Transfer*, 49(17), 3094-3103.

398 Kuhl FP & Giardina CR (1982) Elliptic Fourier features of a closed contour. *Computer Graphics and Image Processing*,
 399 18(3), 236-258.

400 Lammertyn J, Dresselaers T, Van Hecke P, Jancsó P, Wevers M & Nicolai BM (2003) Analysis of the time course of
 401 core breakdown in ‘Conference’ pears by means of MRI and X-ray CT. *Postharvest Biology and Technology*, 29(1),
 402 19-28.

403 Lestrel PE (1989) Method for Analyzing Complex Two-Dimensional Forms: Elliptical Fourier Functions. *American*
 404 *Journal of Human Biology*, 1, 149-164.

405 Lilliefors HW (1967) On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the*
 406 *American Statistical Association*, 62(318), 399-402.

407 Liming X & Yanchao Z (2010) Automated strawberry grading system based on image processing. *Computers and*
 408 *Electronics in Agriculture*, 71, S32-S39.

409 Lu L, Xu H, Song W & Liu G (2007) Research on visualisation of fruits based on deformation. *New Zealand Journal of*
 410 *Agricultural Research*, 50(5), 593-600.

411 Lu L, Song WL & Wang L (2009) A simulation method for the fruitage body. In: *Society of Photo-Optical*
 412 *Instrumentation Engineers Conference Series*, PIAGENG 2009, China.

413 Massey FJ (1951) The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American Statistical Association*,
 414 46(253), 68-78.

415 Mebatsion HK, Boudon F, Godin C, Pradal C, Génard M, Goz-Bac C & Bertin N (2011) A novel profile based model
 416 for virtual representation of quasi-symmetric plant organs. *Computers and Electronics in Agriculture*, 75(1), 113-124.

417 Moreda GP, Muñoz MA, Ruiz-Altisent M & Perdignes A (2012) Shape determination of horticultural produce using
 418 two-dimensional computer vision—A review. *Journal of Food Engineering*, 108(2), 245-261.

419 Moustakides G, Briassoulis D, Psarakis E & Dimas E (2000) 3D image acquisition and NURBS based geometry
 420 modelling of natural objects. *Advances in Engineering Software*, 31(12), 955-969.

421 Nguyen TA, Dresselaers T, Verboven P, D'hallewin G, Culeddu N, Van Hecke P & Nicolai BM (2006). Finite element
 422 modelling and MRI validation of 3D transient water profiles in pears during postharvest storage. *Journal of the Science*
 423 *of Food and Agriculture*, 86(5), 745-756.

424 Nicolai BM, Verlinden B, Beuselinck A, Jancsó P, Quenon V, Scheerlinck N, Verboven P & De Baerdemaeker J
 425 (1999) Propagation of stochastic temperature fluctuations in refrigerated fruits: Propagation des fluctuations de
 426 température stochastiques dans des fruits réfrigérés. *International Journal of Refrigeration*, 22(2), 81-90.

427 Peralta JM, Rubiolo AC & Zorrilla SE (2010) Mathematical modeling of the heat transfer and flow field of liquid
 428 refrigerants in a hydrofluidization system with a stationary sphere. *Journal of Food Engineering*, 99(3), 303-313.

429 Pradal C, Boudon F, Nouguié C, Chopard J & Godin C (2009) PlantGL: a Python-based geometric library for 3D plant
 430 modelling at different scales. *Graphical models*, 71(1), 1-21.

431 Prusinkiewicz P & Runions A (2012) Computational models of plant development and form. *New Phytologist*, 193(3),
432 549-569.

433 Rashidi M & Gholami M (2008) Determination of kiwifruit volume using ellipsoid approximation and image-
434 processing methods. *International Journal of Agriculture and Biology*, 10, 375-380.

435 Rubinstein RY (1981) *Simulation and the Monte Carlo method*. John Wiley and Sons, New York, USA.

436 Saad M, Sadoudi A, Rondet E & Cuq B (2011) Morphological characterization of wheat powders, how to characterize
437 the shape of particles? *Journal of Food Engineering*, 102(4), 293-301.

438 Severa L, Máchal L, Švábová L & Mamica O (2010) Evaluation of shape variability of stallion sperm heads by means
439 of image analysis and Fourier descriptors. *Animal Reproduction Science*, 119(1), 50-55.

440 Tanabata T, Shibaya T, Hori K, Ebana K & Yano M (2012) SmartGrain: High-throughput phenotyping software for
441 measuring seed shape through image analysis. *Plant Physiology*, 160(4), 1871-1880.

442 Veraverbeke EA, Verboven P, Van Oostveldt P & Nicolai BM (2003a) Prediction of moisture loss across the cuticle of
443 apple (*Malus sylvestris* subsp. *mitis* (Wallr.)) during storage: Part 1. Model development and determination of diffusion
444 coefficients. *Postharvest Biology and Technology*, 30(1), 75-88.

445 Veraverbeke EA, Verboven P, Van Oostveldt P & Nicolai BM (2003b) Prediction of moisture loss across the cuticle of
446 apple (*Malus sylvestris* subsp. *mitis* (Wallr.)) during storage: part 2. Model simulations and practical applications.
447 *Postharvest Biology and Technology*, 30(1), 89-97.

448 Verboven P, Nicolai BM, Scheerlinck N & De Baerdemaeker J (1997) The local surface heat transfer coefficient in
449 thermal food process calculations: A CFD approach. *Journal of Food Engineering*, 33(1), 15-35.

450 Verboven P, Flick D, Nicolai BM & Alvarez G (2006) Modelling transport phenomena in refrigerated food bulks,
451 packages and stacks: basics and advances. *International Journal of Refrigeration*, 29(6), 985-997.

452 Zhang D & Lu G (2004) Review of shape representation and description techniques. *Pattern recognition*, 37(1), 1-19.

453 Zhang D & Lu G (2005) Study and evaluation of different Fourier methods for image retrieval. *Image and Vision*
454 *Computing*, 23(1), 33-49.

457
458 Figure 1: Extracted contour of a cross section.

459
460 Figure 2: Reconstruction of a contour from EFD with different numbers of harmonics: the thick black line is the
461 original extracted contour, the blue, cyan, green, yellow, and red lines were respectively reconstructed from 1, 3, 6, 25,
462 and 100 harmonics.

463

464

465

Figure 3: Frequency distributions of the elliptic Fourier descriptors (first six harmonics). Each row represents one harmonic, as each harmonic consists of four descriptors.

466

467

468

469

470

Figure 4: Histograms of the p-values from statistical tests. a) p-values obtained with Lilliefors tests, verifying the normality of the original descriptor distributions. (Note that p-values>0.5 are denoted as being exactly 0.5.) b) p-values obtained with two-sample Kolmogorov-Smirnov tests, verifying the agreement of the original and generated descriptor distributions.

471

472 Figure 5: Generation of contours: a new contour (blue), its mirror image (red), and 15 interpolated contours (with
473 colours ranging from blue to red).

474

475 Figure 6: The contours from Figure 5 brought to 3D: the thick blue-red line is the original contour, the other colours
476 denote the interpolated contours.

477

478 Figure 7: Surface meshes created with ANSYS ICEM 13.0, fitted to the geometry of different generated geometrical
479 models.

480

481 Figure 8: Distribution of the volumes of the scanned fruit and the generated models: a) for Braeburn apples, and b) for
482 Conference pears.

483

484 Figure 9: Geometrical models generated from a dataset of 37 Conference pears.