TOWARDS AUTOMATED SEGMENTATION OF DENSE RANGE SCANS

Jan Böhm
Institute for Photogrammetry, University of Stuttgart, Geschwister-Scholl-Str. 24, 70174 Stuttgart, Germany - jan.boehm@ifp.uni-stuttgart.de

KEY WORDS: Segmentation, Classification, Extraction, Surface Feature, CAD

ABSTRACT

This paper addresses the problem of segmenting dense range data containing curved surfaces. Segmentation is a crucial step in the processing of range data for applications in object recognition, measurement, reengineering and modeling. We propose a two stage process using model-based curvature classification as an initial grouping. Features based on differential geometry, mainly curvature features, are ideally suited for processing objects of arbitrary shape including of course curved surfaces. The second stage uses a modified region growing algorithm to perform the final segmentation. The approach is demonstrated on a test scene acquired with a stripe projection sensor.

1 INTRODUCTION

Today a full variety of range sensing devices is available for all kinds of object sizes, resolution demands and purposes. Applications range from industrial inspection, multimedia, as-built-documentation to cultural heritage and others. Be it active triangulation or time-of-flight systems, scanners are able to densely sample surfaces with great accuracy. Oftentimes hundreds of thousands or even millions of points are obtained from only a single scan. Usually the result of a single scan is represented as an ordered point cloud. Most of the times the topology is based on a grid either given by a frame sensor or by the scan movement.

While this type of data is easily triangulated and directly suited for visualization purposes, most applications require data of another quality. Be it for measurement purposes, reengineering, construction or design purposes it is necessary to give an interpretation to the point cloud, that is to group individual points into meaningful entities. This grouping is called a segmentation. Just what 'meaningful' really is, is determined by the application. In the context of range data, most often a segmentation into individual surfaces is desired.

Several possibilities exist for the segmentation of dense range data. Among the first and most prominent are region growing approaches. The crucial point in region growing is the selection of seed regions. When seed regions are selected to close to the rim of a surface, the resulting region is usually meaningless. Today we see the first implementations of region growing approaches in commercial applications. The problem of seed region selection is left to the operator, who selects regions manually. This manual operation can be quite time consuming considering the size of the point cloud and is generally a tedious task. Thus it is evidently desirable to automate the process.

The work by J.P. Besl (Besl, 1988) has been among the first to present automated region growing for range image segmentation containing curved surfaces. By initially grouping the pixels of a range image according to the sign of the Gaussian and Mean curvature an over segmentation is achieved. By successively applying morphological operators the region are shrunk to minimal size. These minimal regions are then used as seed regions for region growing. While the algorithms were shown to produce good results, the complexity of the approach and the large number of user supplied parameters have prevented the wide-spread adoption of the approach in practice.

In the course of our work on industrial inspection using range imaging, we have developed methods for range image processing and feature extraction. We have earlier reported on our developments of a range image classification framework using differential geometry (Böhm et al., 2000). Features based on differential geometry, mainly curvature features, are ideally suited for processing objects of arbitrary shape including of course curved surfaces. Furthermore these features are suited for object recognition purposes as they are invariant to rotation and translation.

While the curvature-based classification approach gives an interpretation to each individual pixel and by performing a cluster analysis in curvature space effectively gives a grouping of the pixels, this grouping does not take into account the topology of the point cloud. This is a major shortcoming when comparing the classification result to a segmentation. Using the classification result as an initial grouping we now present our work on extending the classification to a segmentation of the point cloud. This two stage process uses a modified region growing algorithm to perform the segmentation based on the topology. While our approach shares some concepts with the one mentioned above, it extends the idea in several ways. First the absolute value of curvature is used for initial classification, not only the sign of curvature. Secondly model knowledge of the object is incorporated into the process and thus greatly enhances reliability. When the type of surface can already be derived from model knowledge, this information can be used to check for the homogeneity of a group of points during the growing process.

While the proposed framework in principle is general to range data, we show its application in the context of industrial measurement purposes. All range data for this project has been acquired with a stripe projection system developed at the Institute for Photogrammetry. The scanner system itself is described in detail in the next chapter including

an examples of the sensor's performance. Chapter 3 describes the fundamental quantities characterizing the local behavior of a surface and gives the mathematical formulas to compute these quantities. In chapter 4 the actual classification and segmentation process is presented detailing our approach to curvature approximation from range data. The proposed approach is demonstrated on a test scenario acquired with our sensor.

2 DENSE RANGE ACQUISITION

For dense surface measurement several alternative measurement techniques are available. If the object's surface shows sufficient radiometric detail, image matching techniques can be employed to recover surface geometry. However in most industrial cases surfaces are not very cooperative with respect to texture detail. One method thus is to use an artificial static texture pattern projected onto the object by a slide projector. Instead of using a static 2D texture pattern, one can just use a single spot, most often generated by a laser beam, which is moved across the surface. Alternatively one-dimensional structures, most often a line, can be projected. To speed up the process several lines can be projected in parallel, leading to the method of coded light projection.

All the methods mentioned above are triangulation based methods. Completely different approaches use the time-of-flight or interferometer principle to determine distance. These approaches have just recently become popular in close range applications now that several commercial laser scanners are available on the market. While each of the methods has its unique advantages and disadvantages, triangulation has the best potential for accurate measurement at very close distances. Because of the speed of measurement a stripe projection system is the most frequent choice for industrial applications in the measurement of small parts and was thus chosen for our project. Since every sensor system has its unique features and also unique problems, we will detail below the sensor system we use.

2.1 Sensor Hardware

We use a LCD type projector for our experiments. The line pattern is generated by switching lines on a two dimensional LCD backlit from behind. This type of projector has the advantage that there are no moving parts. On the other hand, due to the LCD with polarizing filters, brightness is inferior to projectors using metal coated glass plates. While normal LCD stripe projectors use two glass plates with conducting stripes aligned precisely, a cross-pattern projector has one of the glass plates turned by 90 degrees. Since all stripes can be switched individually, arbitrary vertical and horizontal stripe patterns can be generated (albeit no arbitrary 2D patterns can be generated, since the 2D pattern always results from a XOR of the two line patterns). In the context of a photogrammetric evaluation, this means that the projector can be modeled as an inverse camera delivering 2D 'image' coordinates. On the down side, twice as many stripe patterns have to be projected per sequence in order to obtain x and y coordinates.

Figure 1: The sensor hardware used for the experiments consists of a LCD stripe projector and a digital camera.

The projector we use features a LCD with 640×640 lines, line spacing of 0.09 mm (LCD size 57.6^2 mm, and a halogen light source of 400W). Patterns can be switched in 14 milliseconds making it feasible to acquire images in video realtime, although we do not use this option since it requires hardware support. Commands and pattern sequences can be sent to the projector via a RS-232 interface. The camera we use is a digital CCD camera with a resolution of 1300×1030 pixels and approximately 0.0067 mm pixel size. Projector and camera are mounted on a stable aluminum profile with a fixed relative orientation.

2.2 Sensor Calibration

Sensor calibration is a fundamental prerequisite for any vision system that relies on quantitative measurements of the observed scene. Although it is very common to calibrate optical 3-D systems, like stripe projectors, by means of direct calibration techniques (e.g. polynomial models) we found it favorable to use model based calibration (Brenner et al., 1999), where parameters of a geometric model of the sensor, so called intrinsic and extrinsic parameters, are determined. The fact, that model parameters hold true for all the measurement volume of the sensor increases flexibility and omits problems with measurements lying outside the volume originally covered by calibration. In addition, residuals and the obtained covariance matrix give a clear diagnosis for the determination of the calibration parameters and object point coordinates. The model parameters describe how points in 3-D space are projected onto the image plane, considering imperfect cameras and lenses. For a camera this means to find appropriate values for the focal length, principal point position and lens distortion. If a sensor consists of multiple components their relative position and orientation must also be determined. The stripe projection systems are either modeled as inverse cameras or used as an aid to establish point correspondences between at least two cameras e.g. the cameras of the stereo head.

Despite the existence of techniques in photogrammetry to simultaneously estimate these parameters during measurement tasks, we are using a specially designed test object to effectively compute the desired quantities from a few cali-

bration measurements. Since any short-term geometrically stable object can be used for calibration, there is no need for an accurate calibration normal. Nonetheless, since absolute measurements are required, one accurate distance is given from a scale bar to fix the scale.

2.3 Sensor Performance

To be able to assess the accuracy of our sensor system and to determine the amount and nature of noise we have performed a series of tests using a certified test object. The object is a precise sphere with a diameter of 100 mm which is certified to 0.001 mm in shape. We acquired several shots of the object with the sensor. The range data was then fit to a sphere. We have used an implicit polynomial for the fitting. Least squares adjustment was performed using an eigenvalue approach. The minimization criterion we used is the algebraic distance. On a measurement area of about 300×300 mm we have found the standard deviation of the error of fit to be 0.02 mm. The sampling distance on the object's surface is approximately 2 tenth of a millimeter. Of the 25000 points tested 99.8% were below 0.07 mm in deviation. The exact distribution of deviations is given in figure 2.

More important than the magnitude of noise is the nature of the noise on the object's surface. Since the system uses two components incorporating a grid structure we experience Moiré effects on the surface data. They result in concentric ripples across the surface (see figure 2). The amplitude of the ripples is approximately one hundredth of a millimeter. This type of noise is actually worse than purely statistical noise since it is more difficult to filter out and it locally changes surface characteristics. However the sensors accuracy is still more than adequate and allows to capture shape in great detail.

3 THREE-DIMENSIONAL SHAPE

In order to find an initial grouping of the pixels of a range image we have to establish quantities characterizing the local behavior of a surface. The following describes the most fundamental quantities and gives the mathematical formulas to compute these quantities (do Carmo, 1976).

3.1 Fundamentals

Any parameterized surface X in three-dimensional space is given by the projection of an open set U over $\Re^2$ into the space $\Re^3$:

$$X : U \subset \Re^2 \rightarrow \Re^3$$

There are three fundamental ways to describe a surface:

- using a vector-function

$$\overline{X}(u,v) = \begin{bmatrix} x(u,v) \\ y(u,v) \\ z(u,v) \end{bmatrix}, \ (u,v) \in U$$

- using an explicit function

$$X : w = F(u,v)$$

- using an implicit expression

$$X : F(u,v,w) = 0$$

The partial derivatives $\frac{\partial X}{\partial u}$ and $\frac{\partial X}{\partial v}$ are noted as X_u and X_v. In the case of a vector function these partial derivatives are easily computed as:

$$\begin{aligned} \overline{X}_u = & \left(\tfrac{\partial x(u,v)}{\partial u}, \tfrac{\partial y(u,v)}{\partial u}, \tfrac{\partial z(u,v)}{\partial u} \right) \\ \overline{X}_v = & \left(\tfrac{\partial x(u,v)}{\partial v}, \tfrac{\partial y(u,v)}{\partial v}, \tfrac{\partial z(u,v)}{\partial v} \right) \end{aligned}$$

Every explicit function can be converted to vector notation:

$$\overline{X} = (u, v, F(u,v))$$

and its partial derivatives are given by:

$$\overline{X}_u = \left(1, 0, \frac{\partial F(u,v)}{\partial u}\right), \ \overline{X}_v = \left(0, 1, \frac{\partial F(u,v)}{\partial v}\right)$$

It is important to note, that every surface can be locally described by an explicit function.

3.2 Fundamental Forms

A differentiable surface X is given with the condition $\overline{X}_u \times \overline{X}_v \neq 0$. The *unit normal vector* $\overline{N}$ is then given as

$$\overline{N} = \frac{\overline{X}_u \times \overline{X}_v}{\left\| \overline{X}_u \times \overline{X}_v \right\|}$$

The expression

$$\begin{aligned} I &= d\overline{X} \cdot d\overline{X} \\ &= (\overline{X}_u du + \overline{X}_v dv) \cdot (\overline{X}_u du + \overline{X}_v dv) \\ &= E du^2 + 2F du dv + G dv^2 \end{aligned}$$

with

$$E = \overline{X}_u \cdot \overline{X}_{u,} \qquad F = \overline{X}_u \cdot \overline{X}_{v,} \qquad G = \overline{X}_v \cdot \overline{X}_v$$

is called *first fundamental form.*

The expression

$$\begin{aligned} II &= d\overline{X} \cdot d\overline{N} \\ &= -(\overline{X}_u du + \overline{X}_v dv) \cdot (\overline{N}_u du + \overline{N}_v dv) \\ &= L du^2 + 2M du dv + N dv^2 \end{aligned}$$

with

$$L = -\overline{X}_u \cdot \overline{N}_u, M = -\frac{1}{2}(\overline{X}_u \cdot \overline{N}_v + \overline{X}_v \cdot \overline{N}_u), N = \overline{X}_v \cdot \overline{N}_v$$

is called *second fundamental form*. The above condition can be rewritten as

$$L = \overline{X}_{uu} \cdot \overline{N}, \qquad M = \overline{X}_{uv} \cdot \overline{N}, \qquad N = \overline{X}_{vv} \cdot \overline{N}$$

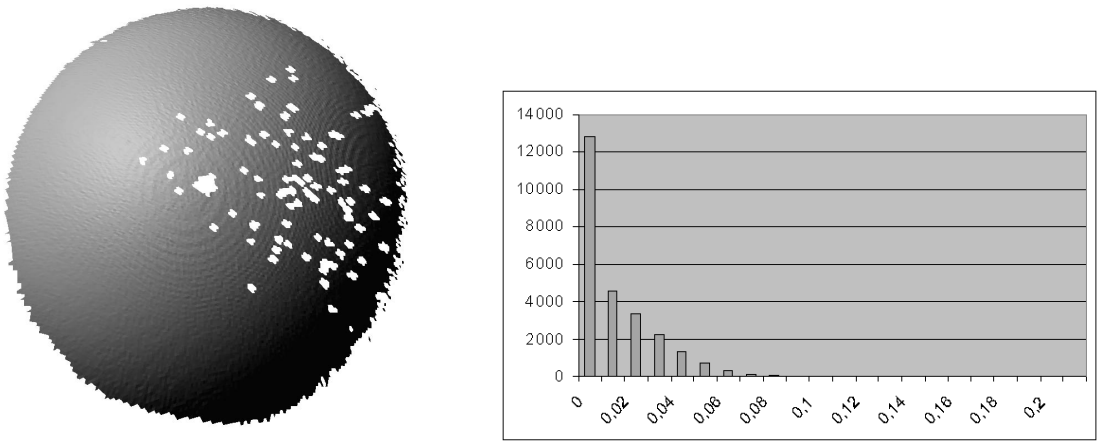

Figure 2: Rendering of a data set acquired with our stripe projection sensor. The object is a precise sphere. Deviations of the sensed data from the ideal sphere are shown on the right. The number of points is plotted over the deviations in millimeters.

3.3 Curvature

Surface curvature is derived from the fundamental forms given above. The principal curvatures are the *maximum curvature* $k_1$and *minimal curvature* k_2. Alternatively the mean curvature H and Gaussian curvature K can be used to describe the surface locally:

$$k_{1,2} = H \pm \sqrt{H^2 - K}$$

and

$$\begin{aligned} K &= \frac{LN - M^2}{EG - F^2} \\ H &= \frac{EN + GL - 2FM}{2(EG - F^2)} \end{aligned}$$

They are translation- and rotation-invariant. While k_1, k_2 and H, K are both valid pairs for local surface characterization, as is noted in (Besl, 1988), there are further considerations which may favor the one over the other. For one to compute the principal curvatures is computationally slightly more expensive. Since the expression of which the square root is taken can become negative due to numerical instabilities, additional precautions have to be taken. The mean curvature is the average of the two principal curvatures and is therefore less sensitive to noise. On the other hand since Gaussian curvature is the product of the two it is much more sensitive to noise. Using only the signs of the curvatures six basic surface types can be determined using principal curvature while eight can be determined using mean and Gaussian curvature.

Based on principal curvature further local properties of a surface can be derived. (Koenderink and van Doorn, 1992) have proposed a shape classification scheme based on two quantities called S and C:

$$S = \frac{2}{\pi} \cdot \arctan\left(\frac{k_1 + k_2}{k_1 - k_2}\right) \quad k_1 \geq k_2$$

$$C = \sqrt{\frac{k_1^2 + k_2^2}{2}}$$

Where S describes the shape, and C the strength of curvature. (C is the square root of the deviation from flatness, another derived quantity in differential geometry). Points of same value for S but differing C, can be seen as points of same shape with stronger curvature. The main difference to the description using mean and Gaussian curvature is the possibility to describe surface flatness with a single quantity $C < \theta$. We will detail the analogy to our approach below. A study comparing both description schemes (Cantzler and Fisher, 2001)found no significant difference of the two. Other authors have extended the SC scheme and have given different formulas for the shape parameter. For our studies we have decided to use the HK scheme.

4 CLASSIFICATION AND SEGMENTATION

In the previous chapter we have given the mathematical quantities used to describe a surface locally. In order to apply these quantities in a classification we have to compute them from range data. Due to the nature of the data described earlier reliable curvature estimation becomes a difficult task crucial to the success of the segmentation process. We have tested the algorithms described below on a dataset of a test scene consisting of a planar, a cylindrical and a spherical region, which was acquired with our sensor (see figure 3).

4.1 Curvature Estimation

Several methods for curvature estimation have been presented in the past. An overview of the most prominent methods has been given in (Flynn and Jain, 1989). For simple approximation the curvature can be computed from the change of orientation from the point of interest to its neighbor. Some methods based on this idea have been presented especially for triangulated surfaces, where surface normals are computed per mesh. These simple techniques are often used for edge pixel detection for example in mesh

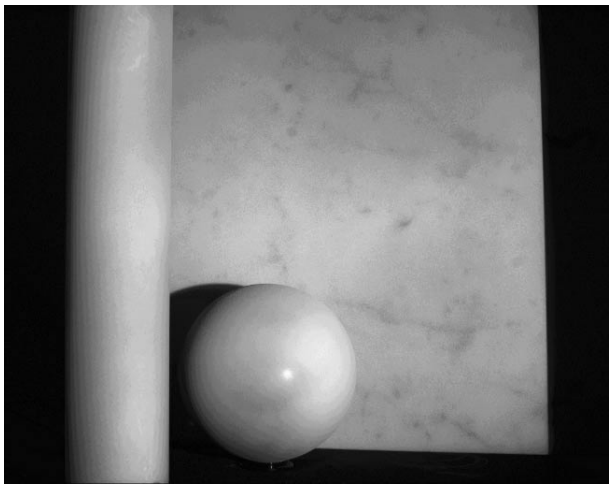

Figure 3: Test scene containing basic shapes as acquired with our stripe projection sensor.

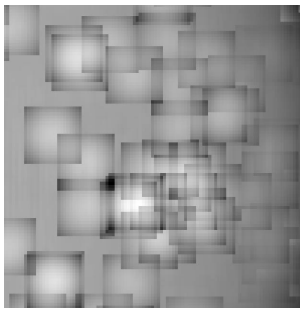

Figure 4: Box shaped artifacts in curvature estimation using convolution approach.

simplification. Since these methods use information only in the direct neighborhood of a point they are extremely sensitive to noise. Since we require exact quantities for our classification these methods were not considered.

Precise estimates of curvature can be obtained from analytic methods. The general strategy of analytical methods is to fit a surface in the local neighborhood of the point of interest and then compute the partial derivatives needed to determine curvature. The main difference of the analytic methods is in the method for local surface fitting. (Besl and Jain, 1986) have proposed a method that is implemented as a series of separable convolution operations one for each partial derivative. The approach is known as orthogonal polynomial approximation. The advantage of their approach is the potential speed of the process using optimized convolution operations. On the down side using a convolution mask does not allow for individual elimination of single points. Thus if the data set contains small holes or outliers within the area of the convolution mask the curvature will deviate by a large amount an cause box shaped artifacts (see figure 4).

We have chosen to use an estimation process based on classical least squares. Moving a square filter mask over the data set individual points are added as observation to the estimation process. Invalid pixels can be easily discarded. This approach even allows us to give individual weights to

SURFACE LABEL	H	K
PLANE	-0.0000	0.0000
CYLINDER	-0.0200	0.0000
SPHERE	-0.0200	0.0004

Figure 5: Model information used for the test scene.

every pixel in the range image. We fit the data to a second degree explicit polynomial:

$$z = ax^2 + bxy + cy^2 + dx + ey + f$$

where the minimization criterion is $\sum(f(x_i, y_i) - z_i)^2$. The partial derivatives are derived directly from the surface parameters.

In order to reliably estimate curvature from range data the data has to be filtered to eliminate noise artifacts. As mentioned above we experience considerable aliasing effects. Median filtering otherwise popular is not suited for this type of noise (Karbacher et al., 2001). A simple weighted averaging is suitable for the purpose. The size of the filter mask has to be established during a test run on a known geometry beforehand, since the noise is specific to every sensor.

4.2 Classification

After mean and Gaussian curvature have been computed for each valid pixel in the range image, each pixel is classified according to its curvature. The model data is entered in the form of HK clusters (see figure 5). A simple minimum distance classification is then used to map every pixel to its corresponding curvature cluster. An additional threshold is used to prevent classification of points with curvature information which deviates strongly from its nearest cluster. Since mean and Gaussian curvature are two different quantities with different magnitudes two separate thresholds can be used to eliminate outliers. (Cantzler and Fisher, 2001) have proposed a formula to derive the threshold for the Gaussian curvature from the threshold for mean curvature. However this formula uses the maximum of mean curvature, a quantity which is difficult to obtain reliably. We have chosen to use only a single threshold for both curvatures. For a planar patch this effectively is similar to the approach of Koendrink mentioned above, since it uses a single threshold to determine planarity of a local region. The result of initial classification is shown in figure 6.

4.3 Region Growing

After initial classification each pixel is either labeled according to its corresponding surface or remains unlabelled. The results show a considerable amount of points which were misclassified, i.e. were assigned a wrong label or were not assigned a label but should have. These misclassifications are caused by false curvature estimation. Especially on the cylinder it becomes evident that not all of the ripples were removed during smoothing. These classification errors have to be removed in a second processing step.

Figure 6: Segmentation results on the test scene. On the left the results after initial classification are shown. On the right the result after region growing and removal of sliver regions is shown.

The most dominant regions, i.e. regions above a certain size threshold, are selected as seed regions for a region growing process. Region growing is implemented as a morphological operation. A 3×3 mask is moved over the dataset. When a neighbor to the point of interest (the center of the mask) has a label assigned, the point of interest is checked for compatibility to that region. In case it is found to be compatible it is assigned the label of the corresponding region. If there are conflicting regions, i.e. there is different regions adjacent to the point of interest, the largest region is preferred. This is also the case if the center pixel is already labeled.

The compatibility check is performed by a least squares fit to a second degree explicit polynomial as described above. If the error of fit is below a certain threshold the point is accepted as compatible. The threshold has to be established beforehand, when evaluating the sensor system. The result of the region growing is shown in figure 6.

5 CONCLUSION

We have presented an efficient technique for the model-based segmentation of dense range scans. The joint use of model-based classification and region growing results in a reliable segmentation and overcomes most of the problems caused by misclassification. Curvature estimation still is a crucial part of the process and remains a topic of intense research. The proposed method is aimed at inspection and measurement task of industrial objects, but also has potential for the application in the automated segmentation of laser scans. In the future we plan to extend the process by improving the compatibility check during region growing. Since the surface type is assumed to be known from initial classification, the surface fit can be constrained to a specific surface type.

REFERENCES

Besl, P., 1988. Surfaces in Range Image Understanding. Springer-Verlag.

Besl, P. and Jain, R., 1986. Invariant surface characteristics for 3d object recognition in range images. Computer Vision, Graphics, Image Processing 33, pp. 33–80.

Böhm, J., Brenner, C., Gühring, J. and Fritsch, D., 2000. Automated extraction of features from cad models for 3d object recognition. In: ISPRS Congress 2000, Vol. 33number 5, Amsterdam, Netherlands.

Brenner, C., Böhm, J. and Gühring, J., 1999. Photogrammetric calibration and accuracy evaluation of a cross-pattern stripe projector. Proc. SPIE Videometrics VI 3641, pp. 164–172.

Cantzler, C. and Fisher, R., 2001. Comparison of hk and sc curvature description methods. In: Third international conference on 3-D Digital Imaging and Modeling, IEEE, Quebec City, Canada, pp. 285–291.

do Carmo, M., 1976. Differential Geometry of Curves and Surfaces. Prentice-Hall.

Flynn, P. and Jain, A., 1989. On reliable curvature estimation. In: IEEE Conference on Pattern Recognition, pp. 110–116.

Karbacher, S., Labourex, X., Schoen, N. and Haeusler, G., 2001. Processing range data for reverse engineering and virtual reality. In: Third international conference on 3-D Digital Imaging and Modeling, IEEE, Quebec City, Canada.

Koenderink, J. and van Doorn, A., 1992. Surface shape and curvature scales. IVC 10, pp. 557–565.