Anatomically Plausible Surface Alignment and Reconstruction

Rasmus R. Paulsen and Rasmus Larsen

Informatics and Mathematical Modelling, Technical University of Denmark
Richard Petersens Plads, Building 321
DK-2800 Kgs. Lyngby, Denmark

Abstract

With the increasing clinical use of 3D surface scanners, there is a need for accurate and reliable algorithms that can produce anatomically plausible surfaces. In this paper, a combined method for surface alignment and reconstruction is proposed. It is based on an implicit surface representation combined with a Markov Random Field regularisation method. Conceptually, the method maintains an implicit ideal description of the sought surface. This implicit surface is iteratively updated by realigning the input point sets and Markov Random Field regularisation. The regularisation is based on a prior energy that has earlier proved to be particularly well suited for human surface scans. The method has been tested on full cranial scans of ten test subjects and on several scans of the outer human ear.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Line and Curve Generation I.4.8 [Image Processing and Computer Vision]: Surface fitting

1. Introduction

With the introduction of fast 3D surface scanners based on sophisticated stereo reconstruction algorithms the clinical relevance of surface based medical image analysis is rapidly increasing. Examples are the analysis of craniofacial dysmorphologies caused by chromosomal changes [HHA*04] and the prediction of cosmetic surgery of for example cleft lip and palate children [KHD06]. While computed tomography and magnetic resonance imaging would provide a complete image of the subject, these modalities are not used for these clinical problems due to the exposure risk and the limited resolution. When using a surface scanner in the clinic, it is crucial to be able to capture the entire face or the entire cranial region with minimal retakes. Especially, for very young children or with patients with syndromes it is not practically possible to continue the acquisition until a perfect scan is available. It is therefore very important to be able to anatomically plausible reconstruct missing parts of the anatomy. Hole-filling algorithms are typically supplied with the modality. However, they are very simple and do not take the differential properties of the surrounding tissue into account. While scanner setups that can capture the entire head are available, there are still many cases, where it is necessary to capture several scans of the same subject from several angles and merge these scans together. A typical example is the scanning of the outer human ear, where the complex anatomy makes it impossible to scan it once. It is our goal to provide a framework that can compute an anatomically plausible surface based on multiple scans of complex anatomies, where parts can be missing and noise present. The data produced by the framework could for example be used for the simulation of the sound field around the head [KN07]. This is currently used in the design of state-of-the-art hearing aids and is done by the boundary element method that requires high-quality meshes. Therefore, the final surface should consist of near equilateral triangles and be free of non-manifold edges.

Surface scanning and reconstruction have mainly been focused on man made objects like engine parts or statues [LPC*00] and the large number of existing algorithms are therefore predominantly focused on this type of data. In our work, we are less focused on sharp corners and straight planes and more on the behaviour of surfaces based on human body scans. The presented method is a combined point

set alignment and surface reconstruction algorithm. Traditionally, this has been treated as separate processes, where surfaces are first computed based on input point sets, then aligned, and finally zippered together [TL94]. The advantage of a combined method is that the surface is represented implicitly during the whole procedure. The triangulated surface is first computed when all scans are aligned and the surface is regularised, enabling sophisticated mesh regularisation based on the implicit surface. While Poisson surface reconstruction [KBH06] is probably the current gold standard algorithm, the membrane-like behaviour of the reconstructed surface in regions with holes is not ideal for human body scans [PBL10]. Another recent method is based on local quadratic functions that are fitted locally and weighted globally using a partition of unity method (MPU) [OBA*03]. Although, the MPU method is both fast and accurate, it does not perform well with point clouds with holes and noise [PBL10]. Recently, a surface reconstruction algorithm based on Markov Random Field regularisation was published [PBL10]. Due to its good hole-filling and noise handling abilities we base our framework on this approach. The major novelty in this work compared to [PBL10] is the groupwise alignment that resembles the approaches taken in recent methods dealing with groupwise registration. In particular methods, that simultaneously estimates an atlas and individual deformation fields [DPTA09]. While these methods typically assume complete samples over a population, we are focused on incomplete samples over one subject. A statistical approach to combined alignment and reconstruction of point clouds is described in [HAW07]. Here, a set of point clouds is roughly aligned. Secondly, a *prototype* surface is computed using a local quadratics approach that is very similar to the MPU method [OBA*03]. In the next step, the poses of the input point clouds are changed to match the prototype. While this method is similar to our method, it cannot handle areas with missing data, due to the previously mentioned shortcomings of the MPU method.

2. Method

The method proposed consists of several steps that are iterated until convergence. The input consists of a set of roughly aligned point sets. Initially, a single oriented point set with consistent normal directions is created from the input data. Secondly, a signed distance field is computed based on the oriented point set. In the next step a Markov Random Field based regularisation is applied to the distance field. In the last step, the input point sets are aligned to the implicit zero level using an implicit variant of the iterative closest point algorithm (ICP) [BM92]. These steps are then repeated in a multi-scale framework. Finally, the resulting surface is extracted from the regularised distance field using an iso-surface extraction algorithm, and the resulting mesh is optimised. Conceptually, the method keeps track of an ideal implicit surface representation of the combined input data. This implicit surface representation is updated and refined until it optimally represents the aligned input point sets.

Step 1: Point set merging. In the first step, the input point sets are merged into one point set. At this point, they are only roughly aligned and the merged set will therefore contain points that are seen as both outliers and noise. An example can be seen in Fig. 1 left, where the polygons from the scanner are kept for visual purposes (they are discarded before point processing). Consistent normals are computed using a local principal component analysis followed by a graph based voting scheme [PBL10]. In the first iteration, a large neighbourhood (25 nearest neighbours) is used in the PCA, which smoothes out the artefacts caused by the initial misalignment.

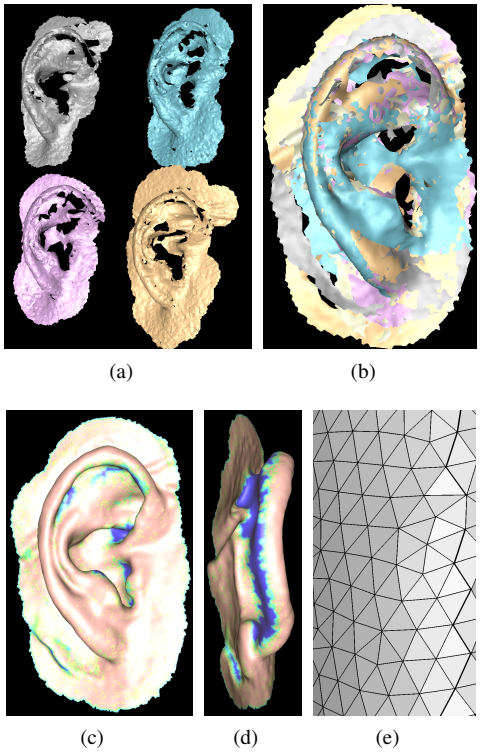

Figure 1: *Aligned and reconstructed scans of the outer human ear.* ***a)*** *Input data with original polygons kept,* ***b)*** *roughly aligned input data,* ***c+d)*** *resulting surface from two angles, and* ***e)*** *close-up of final polygonisation.*

Step 2: Computing the Signed Distance Field The distance field is represented as a uniform voxel volume, where the bounds of the volume are computed to extend five voxels beyond the bounds of the merged point set. The signed distance is computed in each voxel as the distance from the voxel centre to the line spanned by the five closest points and the average of their normals. In practice, a fast method is used to compute the entire signed distance field [JBS06]. Optimally, the zero set of this distance field defines the wanted surface. However, this initial distance field suffers from ambiguities

in regions with holes in the merged point set and in regions with noise. In the following, The initial distance field is defined as $\mathbf{d}^o$ and the initial distance at voxel i as d_i^o. We use a simple one dimensional indexing of the voxels instead of the standard (x,y,z) indexing.

Step 3: Markov Random Field Regularisation To remove the influence of noise and introduce hole-filling capabilities, the signed distance field is regularised using a Bayesian approach. The distance field, $\mathbf{d}$, is formulated as a Markov Random Field and the goal is to compute the distance field $\hat{\mathbf{d}}$ that maximises the posterior probability

$$\hat{\mathbf{d}} = \arg\max_{\mathbf{d}} \; p(\mathbf{d}|\mathbf{d}^o). \tag{1}$$

The Markovianity assumption implies that the involved probabilities can be formulated locally using near voxel neighbourhoods. In the following, we are using a 6-neighbourhood for each voxel. The $n = 6$ neighbours of voxel i is written as $i \sim j$. The local probabilities are based on the Gibbs distribution and are therefore formulated as energies. The local prior probability is based on differences between neighbouring Laplacians:

$$U_{\mathcal{L}}(d_i) = \sum_{i \sim j} (\mathcal{L}(d_i) - \mathcal{L}(d_j))^2, \tag{2}$$

where d_i is the voxel value in the current distance field and $\mathcal{L}(d_i)$ is the Laplacian estimated at voxel i. A discrete approximation of the Laplacian is used:

$$\mathcal{L}(d_i) = \frac{1}{n} \sum_{i \sim j} (d_i - d_j). \tag{3}$$

This approximation is known from image processing, where it is normally formulated as a 3x3 kernel. The normalisation is chosen to be the number of neighbour voxels and this is constant, except at the borders of the volume. Obviously, this prior favours distance fields with smooth Laplacians. Since the Laplacian of a distance field is proportional to its mean curvature [JBS06], the prior in Eq. (2) favours distance fields with small variations in mean curvature. Furthermore, it has been shown that this prior is particularly well suited for surface scans of the human body [PBL10]. The prior model can conceptually be understood as how we would like the surface to behave in regions with no samples. However, a model that forces the surface to follow input data is needed and therefore the following observation model is used:

$$U_{\text{obs}}(d_i^o) = (d_i - d_i^o)^2, \tag{4}$$

As mentioned earlier, d_i^o is the original distance at voxel i and d_i is the current estimate. Using the Gibbs measure, this energy function leads to $p(d_i^o|d_i) = \exp(-U_{\text{obs}}(d_i^o))$, that is, a Gaussian distribution with mean d_i. The observation model describes the distribution of the observed values given a true underlying distance field. In our case, the initial distance field is the observed data and we are seeking an estimate of the true distance field $\hat{\mathbf{d}}$.

In order to balance the prior and observation model a per voxel confidence measure $\alpha_i \in [0:1]$ is introduced. It is based on the Euclidean distance from the voxel centre to the nearest input point d_i^E. Here $\alpha_i = 1 - \min(d_i^E / d_{\max}^E, 1)$, where $d_{\max}^E$ is a user-defined maximum Euclidean distance. A discussion of suitable values of $d_{\max}^E$ can be found in [PBL10] and a discussion of other confidence measures can be found in [CL96]. Confidence values in voxels near input points are therefore one and in regions with no input points zero. In practice, a confidence volume is precomputed. In Fig. 2 the pre-computed α-map has been projected into the resulting surfaces and thereby visualising which part of the resulting surface that is influenced by the prior, the observation, or a combined model. Using Bayes' theorem the prior and the observation terms are combined into a posteriori probability per voxel. Using the Gibbs mea-

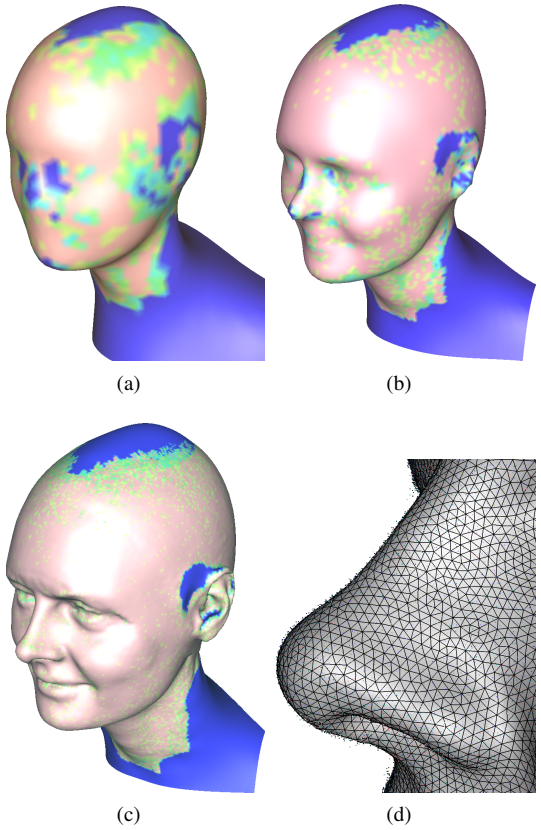

Figure 2: ***a+b+c)*** *The extracted zero-level iso-surface from iteration 0, 1, and 3. The colour coding corresponds to the weights in the α-map. Blue is $\alpha = 0$, where a pure prior energy is used. Skin colour is where $\alpha = 1$ and shows where a pure observation models is used.* ***d)*** *the final polygonised nose is seen with the aligned points.*

sure, it becomes:

$$p(d_i|d_j, i \sim j) = \exp\left(-\alpha_i\beta\, U_{\text{obs}}(d_i) - (1-\alpha_i\beta)\, U_{\mathcal{L}}(d_i)\right),$$

where the global weight β and the local α_i's are used to balance the prior and observation models. Using this, the maximum likelihood estimate of the voxel value d_i given its six neighbours can be computed as a linear combination of neighbouring voxel values and the original distance estimate [PBL10].

We wish to compute the distance field that maximises the a posteriori conditional probability $\hat{\mathbf{d}} = \arg\max_{\mathbf{d}}\ p(\mathbf{d}|\mathbf{d}^o)$. The maximisation of the posterior probability is transformed into the minimisation of the weighted sum of the energy functions for the prior distribution and the observation model over the entire field. The global optimum can be found using the Iterative Conditional Modes (ICM) algorithm [Bes86], where each site is iteratively assigned its local maximum likelihood estimate. Due to the uniform space division, a multiscale ICM solver is fairly easy to implement. Initially, the solution is found in a small voxel volume and the results are propagated to the next level using linear interpolation. In each step, the side lengths of the voxel volume are doubled. Furthermore, a new initial distance estimate is recomputed for the voxels in a narrow band around the input points at each level. In the current implementation, the maximum side length of the initial volume is 16 and the typical number of levels used is four. Other strategies can be used to find the optimal solution, including conjugate gradients and sparse Cholesky factorisation [PBL10]. However, the proposed multi-scale solver is both faster and can handle larger volumes than other tested strategies. The popular graph cut algorithm for solving discrete Markov Random Fields is not applicable to our problem due to its non-discrete nature.

The result of the regularisation is a distance field where the behaviour of the zero-level iso-surface is well defined, even in areas with no input points. In areas with dense sampling, the iso-surface adheres implicitly to the observation energy and in areas with no input points it follows the prior energy. In the first two or three iterations, a smooth distance field is computed by selecting a low β value. In later iterations, a high β value is used putting more weight on the observation model. Note, that the word *distance field* is used in a loose sense since we do not enforce unit length gradients.

Step 4: Realignment In this step the individual point sets are realigned to the zero-level iso-surface. Since this iso-surface is both smoothed and regularised, the influence of the individual point sets have been blended together. A regularised iso-surface from the first iteration can be seen in Fig. 2 left. A specialised version of ICP is used. Each input set is individually realigned to the zero-level iso-surface. For each point in the input point set the closest point on the iso-surface is found. Since we have an approximate distance field this can be done using a Runge-Kutta like gradient descent. The sought after rigid-body realignment transform can now be computed using the closed form solution found in [Hor87]. Finally, a new merged point set is created after the transformation of the input point sets.

Step 1-4 are repeated in a multi-scale approach. The scale is determined by the voxel size used in the distance field representation. Due to the multi-scale ICM solver, the voxel size is specified as a minimum voxel size that determines the last scale in the multiscale ICM. In our experiments, we have used the following parameters: iteration 0: $\beta = 0.1$, max voxels = 20.000, iteration 1: $\beta = 0.2$, max voxels = 100.000, iteration 2: $\beta = 0.4$, max voxels = 500.000, iteration 3: $\beta = 0.8$, max voxels = 1.000.000, iteration 4: $\beta = 0.9$, max voxels = 1.000.000. The resulting iso-surfaces can be seen in Fig. 2 for iteration number 0, 1, and 3. While the head is mostly represented by a smooth spherical shape in the first iteration, all details are seen in iteration 3. In practise, the iso-surfaces are not triangulated during the iterations.

Step 5: Surface extraction In step 1-4 the sought after surface is represented as the zero-level iso-surface of the regularised distance. A polygonised version can be extracted using for example marching cubes [LC87]. Marching cubes have a tendency to create triangles with bad aspect ratios and therefore a mesh optimisation scheme is applied. This is greatly aided by the implicit description of the surface that enables easy calculation of tangents and point reprojections. We use a modified version of the technique described in [BK04]. Details can be found in [PBL10].

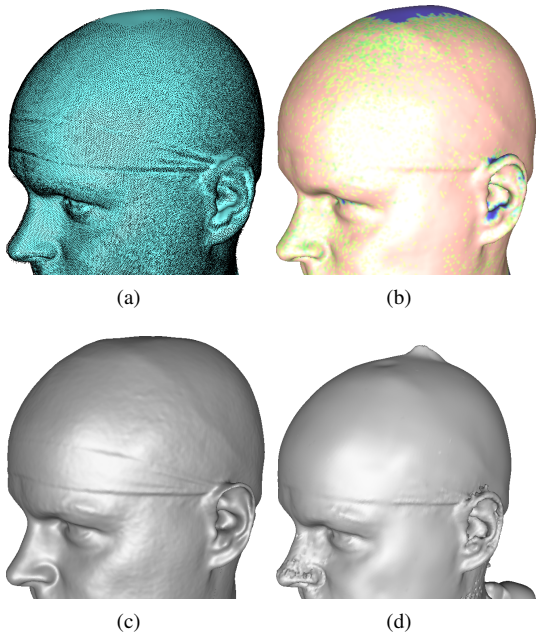

Figure 3: *Comparison.* ***a)*** *Aligned points on MRF surface,* ***b)*** *MRF surface with α-map,* ***c)*** *Poisson surface, and* ***d)*** *MPU Surface.*

3. Results

The algorithm has been tested on scans of the head of ten volunteers. The scans were performed with a 3dMD cranial system in a 2 pod configuration. The subject was placed in a rotating chair and scans were taken from several directions while keeping the head pose fixed. The results in Fig. 2 and 3 are representative for the results obtained from all subjects. The results are highly satisfactory since the algorithm 1) produces a good surface in areas with good quality input data and 2) interpolates in a plausible way in areas with no input samples (the blue regions). Furthermore, the surfaces consist of near equilateral triangles as seen in Fig. 1 right and Fig. 2 right and no non-manifold edges have been detected. The algorithm has also been tested on the complicated geometry of the ear as seen in Fig. 1. It can be seen that even if the input data contains noise and does not cover the full ear, the resulting surface successfully captures the outer ear shape. Even behind the pinna, where the input data is very sparse, the algorithm is able to faithfully reconstruct the anatomy of the ear. On a standard PC (WindowsXP 32 bit, 4 GB RAM, 2 GHz) the running time for the ear example is less than five minutes and for a human head around 10 minutes. The Poisson surface reconstruction algorithm [KBH06] has been tested on a set of aligned point sets as seen in Fig. 3. Compared to the MRF results, the Poisson surface is flattened where there are no input data. This is due to the membrane-like nature of the interpolation used in the Poisson methods. While the Poisson methods does not interpolate as well as the MRF method, the MPU surface reconstruction algorithm [OBA*03] simply fails in regions with noise and outliers as seen in Fig. 3. Neither the Poisson nor the MPU based methods does alignment and the alignment in this comparison was therefore done using the proposed MRF based method. The accuracy of the result obviously depends on the voxel size of the final volume used in the last iteration. For practical applications dealing with for example the human head and torso an accuracy of 0.1 mm can be achieved with a moderately sized volume (1 million voxels). This is currently better than the surface scanner hardware can provide.

4. Summary and Conclusions

A framework for the simultaneous alignment of point clouds and reconstruction of the underlying true surface has been presented. Compared to previous approaches the method performs particularly well in regions with no data and in presence of noise. This is due to the higher-order prior energy term used in regularisation of the implicit surface representation. Currently, the realignment is done using rigid body transformations. To accommodate slight changes in head posture during capture, an obvious extension is the use of non-rigid surface registration in the realignment step. The framework is well suited for complex geometries like the outer human ear that cannot be captured in a single acquisition. It is our belief, that a groupwise alignment is ideal for the type of data used in this paper. A pair-wise alignment algorithm would fail, since the scan of the front of the head does not overlap with the scan of the back of the skull. If an incremental method, that aligns scan 0 to scan 1 and so forth, was chosen errors propagate and a pre-sorting of the scans is necessary. Furthermore, the multi-scale approach, where the scans are re-aligned to a smooth combined surface, plays a major role in the successful alignment. If the scans were re-aligned to a surface that was reconstructed at the original level of detail, they would only be aligned to their own (noise and incomplete) part of the reconstructed surface.

Acknowledgements The authors would like to thank the Oticon Foundation for partly funding this work and the 3D Craniofacial Image Research Laboratory at Copenhagen University for letting us use their scanner. The software and test data can be downloaded from www.imm.dtu.dk/MRFSurface.

References

[Bes86] BESAG J.: On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society, Series B 48*, 3 (1986), 259–302. 4

[BK04] BOTSCH M., KOBBELT L.: A remeshing approach to multiresolution modeling. *Proc. Symposium on Geometry processing* (2004), 185–192. 4

[BM92] BESL P. J., MCKAY N.: A method of registration of 3D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence 14*, 2 (1992), 239–256. 2

[CL96] CURLESS B., LEVOY M.: A volumetric method for building complex models from range images. *Proc. SIGGRAPH* (1996), 303–312. 3

[DPTA09] DURRLEMAN S., PENNEC X., TROUVÉ A., AYACHE N.: Statistical models of sets of curves and surfaces based on currents. *Medical Image Analysis 13*, 5 (2009), 793–808. 2

[HAW07] HUANG Q., ADAMS B., WAND M.: Bayesian surface reconstruction via iterative scan alignment to an optimized prototype. *Proc. Eurographics symposium on Geometry processing* (2007), 213–223. 2

[HHA*04] HAMMOND P., HUTTON T., ALLANSON J., CAMPBELL L., HENNEKAM R., HOLDEN S., PATTON M., SHAW A., TEMPLE I., TROTTER M.: 3D analysis of facial morphology. *American Journal of Medical Genetics 126*, 4 (2004), 339–348. 1

[Hor87] HORN B. K. P.: Closed form solution of absolute orientation using unit quaternions. *Journal of the Optical Society A 4*, 4 (1987), 629–642. 4

[JBS06] JONES M., BÆRENTZEN J., SRAMEK M.: 3D Distance Fields: A Survey of Techniques and Applications. *IEEE Transactions On Visualization and Computer Graphics 12*, 4 (2006), 518–599. 2, 3

[KBH06] KAZHDAN M., BOLITHO M., HOPPE H.: Poisson Surface Reconstruction. *Proc. Symposium on Geometry Processing* (2006), 61–70. 2, 5

[KHD06] KREIBORG S., HERMANN N., DARVANN T.: Characteristics of Facial Morphology and Growth in Infants with Clefts. *Cleft lip and palate. Diagnosis, and management. 2nd ed.* (2006), 225–35. 1

[KN07] KAHANA Y., NELSON P.: Boundary element simulations of the transfer function of human heads and baffled pinnae using accurate geometric models. *Journal of Sound and Vibration 300*, 3-5 (2007), 552–579. 1

[LC87] LORENSEN W. E., CLINE H. E.: Marching cubes: A high resolution 3D surface construction algorithm. *Proc. SIGGRAPH* (1987), 163–169. 4

[LPC*00] LEVOY M., PULLI K., CURLESS B., RUSINKIEWICZ S., KOLLER D., PEREIRA L., GINZTON M., ANDERSON S., DAVIS J., GINSBERG J., SHADE J., FULK D.: The digital Michelangelo project: 3D scanning of large statues. *Proc. SIGGRAPH* (2000), 131–144. 1

[OBA*03] OHTAKE Y., BELYAEV A., ALEXA M., TURK G., SEIDEL H.: Multi-level partition of unity implicits. *Proc. SIGGRAPH* (2003), 463–470. 2, 5

[PBL10] PAULSEN R. R., BÆRENTZEN J. A., LARSEN R.: Markov random field surface reconstruction. *IEEE Transactions on Visualization and Computer Graphics 16* (2010), 636–646. 2, 3, 4

[TL94] TURK G., LEVOY M.: Zippered polygon meshes from range images. In *Proc. SIGGRAPH* (1994), p. 318. 2