MVA2015 IAPR International Conference on Machine Vision Applications, May 18-22, 2015, Tokyo, JAPAN

Mobile 3D Wood Pile Surveying

Christopher Herbon
HAWK Fakultät N
christopher.herbon@hawk-hhg.de

Benjamin Otte
HAWK Fakultät N
benjamin.otte@hawk-hhg.de

Klaus-Dietz Tönnies
Otto-von-Guericke-Universität Magdeburg
klaus.toennies@ovgu.de

Bernd Stock
HAWK Fakultät N
bernd.stock@hawk-hhg.de

Abstract

We present a novel method for photogrammetric wood pile surveying, which runs on mobile devices as well as on desktop computers. The demand for measurement techniques for wood piles has strongly increased in the last years. Unlike existing methods, our method is not limited to a single image and uses 3D reconstruction techniques on a set of images taken with a smartphone or digital camera. The reconstructed 3D model is then used to identify individual wood logs and perform photogrammetric surveying of the entire wood pile. An extensive evaluation is conducted on 246 data sets (7655 images) from the publicly available HAWK-wood database. For the wood log detection benchmark a true positive rate of 98.8% with a false positive rate of 0.7% is achieved. The volume computation showed an average absolute difference of 2.2% (contour volume) and 5.6% (solid wood volume).

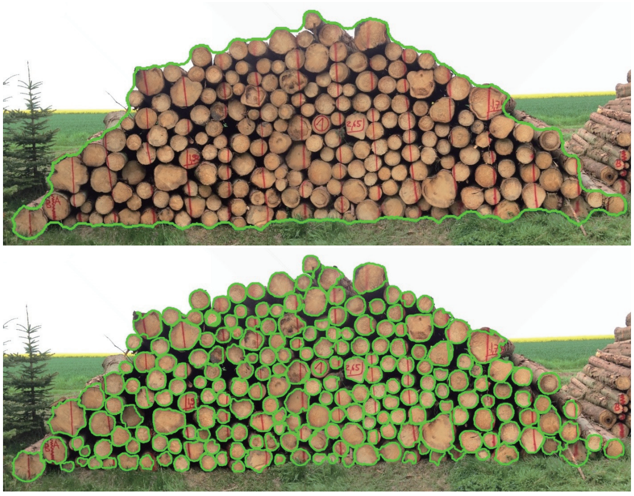

Figure 1. Wood pile. (a) Contour volume (green). (b) Individual wood logs / solid wood volume

1 Introduction

Wood piles are defined as stacked wood logs of a certain wood type, length, and quality. When trees are harvested, they are piled up on the side of the road to be counted, measured, sold, and dispatched. In order to determine cost and time efficient logistics for trucks to pick up and distribute the wood logs, forestry rangers must first gather information about the wood piles, which include the following for each pile:

- wood type and quality
- wood log length
- number of wood logs
- wood pile contour volume (figure 1(a))
- solid wood volume (figure 1(b))

While the first two parameters can be quickly determined upon visual inspection, the counting of the wood logs and the measurement of the wood pile contour volume, and the solid wood volume are error-prone, time consuming, and thus costly. Our novel method is capable of accurately determining these quantities by using multiple view reconstruction techniques with the wood pile width as a scale reference.

The largest obstacle for wood pile surveying is a varying wood type, wood quality, and especially varying geometric parameters of the wood pile, such as the arrangement of the wood logs and vegetation covering the front surface. State of the art methods focus on image based wood log detection and wood log location estimation but do not consider the wood pile itself as a meta-object. In this paper we use these methods as a foundation to reconstruct and measure wood piles.

When applying our method it is crucial for the user to receive immediate feedback on the measurement process. With no mobile internet available in many remote forestry areas a cloud service cannot be used. Our method works on handheld mobile devices, such as smartphones and tablets, and does not require a data connection of any kind.

2 Related research

Wood log detection has been an active area of research for over a decade. The fundamentals of wood log surveying were established by Fink [2]. Much research is based on the idea to perform wood log detection first and then use this information for detailed segmentation. For single images Gutzeit and Voskamp [3] show how individual wood logs can be detected and segmented by combining Haar-like features [10] and graph gut segmentation. A statistical model is built from the initial classification, which is then used to separate foreground from background. The approach still constrains the centroid of the wood pile to be located in the image center in order to correctly compute the statistical model.

Gutzeit and Voskamp's method [3] is further improved by [4] who extend this approach by an iterative classification, segmentation, and statistical modeling method. An initial model is built from a subset of

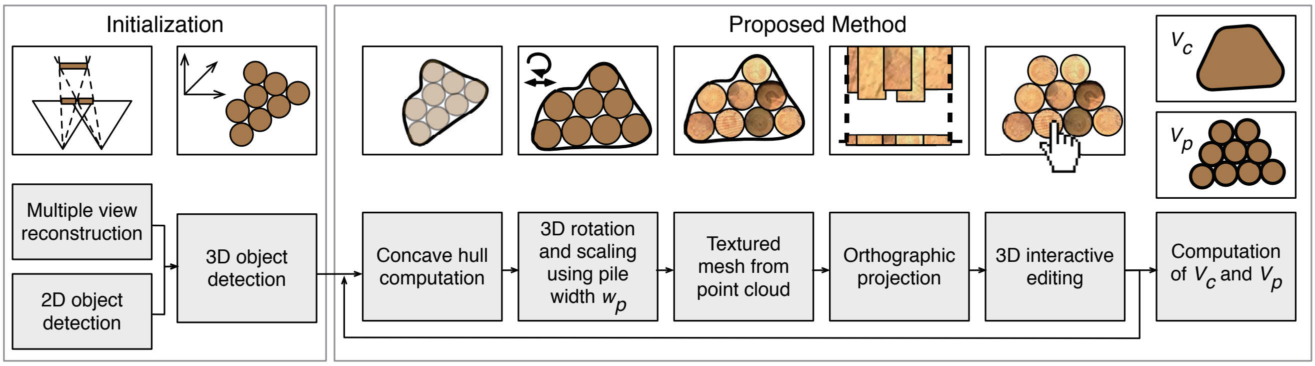

Figure 2. Method overview

wood log faces, which is then iteratively refined and used to detect the remaining wood logs.

While all these methods consider two dimensional information only, [5] uses camera calibration techniques [9] combined with a wood log detection method [4] to locate wood logs in 3D space and approximate their relative locations. This is achieved through quadric filtering and back projection of the detected two dimensional objects to 3D space. In our work we use this method as a basis for three dimensional wood pile reconstruction (see figure 2, *initialization*).

We evaluate our approach on the new publicly available *HAWKwood* database [6] (*Multi-Image Benchmark, M.1-M.3*). It includes 354 data sets in total with 7655 images, of which we use 246 data sets (6376 images). There exists a dedicated benchmark for each of the three parameters that we wish to compute: number of wood logs, wood pile contour volume, solid wood volume. For each data set the wood pile width and wood log length are given.

3 Wood pile as a meta object

Previous research has considered wood logs as individual objects. In our work we are interested in handling the wood pile as a meta object, meaning that we propose to handle the wood pile as a unit, which consists of wood logs, the space between wood logs, and parameters connected to the surroundings of the wood pile. By choosing this representation we can include additional information from the meta object, such as the front surface normals, and three dimensional orientation in our reconstruction and measurement.

By forestry convention we only measure the front of the wood pile, as it is considered to be representative for the entire wood pile when stacked according to forestry standards [7]. While the number of wood logs is a well defined quantity, the wood pile contour volume V_c and the solid wood volume V_s are explicitly defined by forestry conventions [6][7].

$$V_c = w_p l \frac{1}{k} \sum_{i=1}^{k} h_i \qquad (1)$$

The contour volume is denoted as V_c. Equation 1 shows its definition [6][7], where w_p is the width of the wood pile, l is the wood log length, and h_i is the height of the ith wood pile section. The section height h_i is determined by dividing the wood pile in equal sections and measuring the height in the center of the section (see [6] for a visual representation).

To compute the solid wood volume only, equation 2 uses the number of wood logs N, the wood log diameter d_i and the known, constant wood log length l. We can see that V_s is defined as the sum of all single wood log volumes.

$$V_s = \sum_{i=1}^{N} \left(\frac{d_i}{2}\right)^2 \pi l \qquad (2)$$

4 Method

4.1 Contributions

1. Our method provides a simultaneous solution for the determination of the number of wood logs, the wood pile contour volume, and the solid wood volume.

2. The proposed method is the first multiple view wood pile surveying technique.

3. All algorithms are specifically optimized for the use on mobile devices.

4. We provide results of extensive testing on several hundred data sets.

4.2 Proposed method

For each wood pile we compute a 3D reconstruction, consisting of a sparse point cloud and camera calibration data (obtained through structure from motion [9]). Furthermore we receive the 3D location and circular approximation (with radius r) of the wood log front faces through the method described in [5]. We denote the point cloud as Q and the camera poses as P_i with $i = 1 \ldots m$ and m being the number of registered images. The detected 3D objects (wood log faces) are defined as the set O, where each $O_i = \{x_i, y_i, z_i, r_i\}$ with the 3D center location $[x_i, y_i, z_i]$ of the wood log face and the radius r_i. The 3D wood log faces are used as an initial approximation of the wood pile front surface.

Structure from motion performs extrinsic calibration up to a scale factor only, which is usually unknown. Our experiments have shown that the scale of the wood pile is best recovered by the largest measurable length

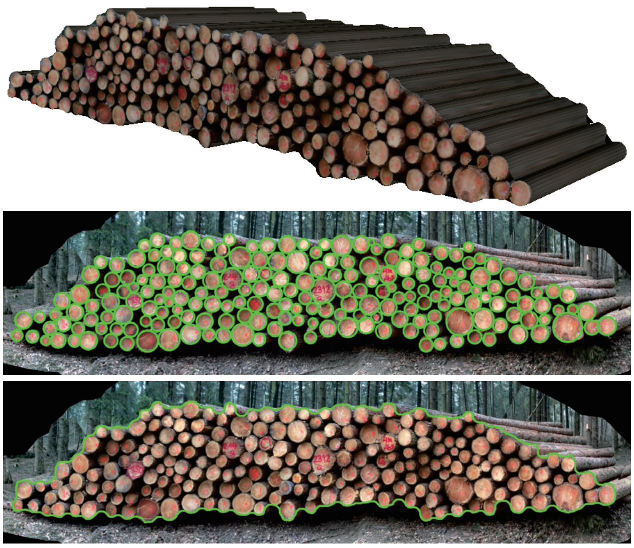

Figure 3. (a) 3D reconstruction. (b) Orthographic projection and detected objects. (c) Orthographic projection and contour volume V_c (green)

of the object. In our case this is the wood pile width. The accuracy of the wood pile width measurement is discussed in section 5.

In our method the first step for scale recovery is the computation of the concave hull of the wood pile, represented by the vertices of the wood log faces (see figure 2). For this we choose a k-nearest neighbor based method proposed by Moreira and Santos [8]. From an initial guess of $k = 3$, we iteratively refine the parameter k until a single concave contour is found, which consists of 3D vertices. The resulting concave hull is denoted as C. The points with the largest distance between them ($\mathbf{p}_{-\mathbf{x}}$ and $\mathbf{p}_{+\mathbf{x}}$) are considered the left and right edge of the pile. The Euclidean distance between $\mathbf{p}_{-\mathbf{x}}$ and $\mathbf{p}_{+\mathbf{x}}$ in $\mathbb{R}^3$ is denoted as $d = |\mathbf{p}_{+\mathbf{x}} - \mathbf{p}_{-\mathbf{x}}|$

For wood piles it can be generally assumed that the front surface is not strictly, but approximately planar (quasi-planar). Furthermore a wood pile always exhibits a width larger than its height or depth. We can determine the surface normal $\mathbf{n}$ of the quasi-planar front plane through a RANSAC-based plane fitting approach. The obtained $\mathbf{n}$ can be used to align the 3D front surface to the x-y-plane so that $\mathbf{n} \sim \mathbf{z}$, where $\mathbf{z} = [0, 0, 1]^T$ is the z-axis and $\sim$ denotes equality up to scale.

By performing principal component analysis (PCA) of the detected objects distribution in 3D, we can compute the eigen vector $\boldsymbol{\Upsilon}_{\mathbf{x}}$, which is the eigen vector with the largest eigen value λ. The vector $\boldsymbol{\Upsilon}_{\mathbf{x}}$ corresponds to the horizontal axis of the wood pile and may thus be used for alignment of the wood pile to the x-axis. From the rotated and scaled wood pile object a 2.5D mesh is created via delaunay triangulation [1] and texture is mapped onto it from the original images.

In [5] it is shown that the wood log detection rate (true positive rate) is approximately 98.0% with a false positive rate of 0.5%. Hence some wood logs may not be detected. A user needs to be able to add or remove wood logs from the three dimensional reconstruction. To prevent projective ambiguities, we perform orthographic projection of the wood pile along the z-axis (see figure 3(b)). This image can then be used as a visual representation for user input.

While it is trivial to remove falsely detected wood log faces from the reconstruction, the process of adding objects is more complex. [5] proposes to perform quadric filtering to determine a subset of vertices, which lie in the proximity of the true 3D location. In our case the filter parameters are determined through the size (r) and 2D location (x, y) of the object in the orthographic projection.

After the user finished editing, the pipeline is run again and the contour volume V_c and solid wood volume V_s are computed. The contour volume is determined via the orthographic projection of the 3D reconstruction. It is defined as the area of the projection of the front surface A_c (see figure 1(a)) multiplied with the known depth of the wood pile. Given the contour of the orthographic projection of the concave hull C_{op}, the area surrounded by this curve A_c can be determined through Green's theorem in equation 3, where $f(x, y)$ and $g(x, y)$ must be chosen explicitly. The contour volume V_c is defined as the multiplication of the known wood log length l with A_c.

$$A_c = \iint dA = \int_{C_{op}} (f(x, y)dx - g(x, y)dy) \quad (3)$$

$$V_c = A_c l \quad (4)$$

The last parameter of the wood log pile, that we are interested in, is the solid wood volume V_s. As shown by [6] and [7] V_s is defined as the sum of all wood log faces multiplied with the wood log length. Since we use a circular approximation for wood log faces, we can define the face area according to equation 5 and V_s through equation 6.

$$A_s = \sum_{i=1}^{N} \pi r_i^2 \quad (5)$$

$$V_s = A_s l \quad (6)$$

5 Experiments

Our method is based on [5], where experiments are performed on 65 data sets to determine the wood log detection rate. We extend these experiments to perform surveying on the wood pile as a whole. Experiments are conducted on the *HAWKwood* database [6], which consists of a total of 354 data sets and 7655 images for three different benchmarks, one for each of the parameters we wish to determine.

5.1 Number of wood logs

The detection of the number of wood logs is the basis of our proposed method. For our evaluation we use the 246 data sets with high overlap from the *HAWKwood M.1* benchmark, which are provided specifically for multiple view reconstruction. For 4 of these data sets results could not be obtained, as multiple view reconstruction failed, due to insufficient overlap of the images. For comparability we show results for fully automatic detection without interactive editing.

The detection rate *tpr* (true positive rate) and false positive rate *fpr* are defined by the following equations, where N_{true} is the number of correctly detected wood logs, N_{false} is the number of falsely detected wood logs, and N is the ground truth number of wood logs.

$$tpr = \frac{N_{true}}{N} \qquad fpr = \frac{N_{false}}{N} \tag{7}$$

	$\bar{x}$	s
tpr	98.8%	1.3%
fpr	0.7%	1.1%

Table 1. Automatic wood log detection results

We can see from table 1 that the *tpr* and *fpr* are very close to the results reported in [5].

5.2 Scale reference

As we mentioned previously, the width of the wood pile is the largest accurately measurable length of the object, and will thus be used as a scale reference. To determine its accuracy we have conducted 200 measurements of different wood pile widths by five different forestry rangers. It resulted in a very low standard deviation of $s_w = 0.12\%$ for wood piles in the range of 4.1m to 35.3m.

5.3 Solid wood volume

For the computation of the solid wood volume we use the automatic detection results combined with the interactive editing by the user. We can thus assume a *tpr* of 100% and a *fpr* of 0% for the evaluation of the contour volume and solid wood volume. The results for the *HAWKwood M.2* benchmark (71 real and 40 synthetic data sets with high overlap) show a lower drift for the real data than for synthetic data. Both results are suited for practical use, as the drift of the mean and the standard deviation are well within the limits specified by forestry conventions. When measuring by hand, the difference to the real value, according to [7], is usually up to 8%.

	$\bar{x}$	s
real	-5.6%	3.4%
synthetic	-7.0%	3.2%

Table 2. V_s results (difference to ground truth)

5.4 Contour volume

Similar to the solid wood volume, the computation of the contour volume was performed after interactive editing to ensure $tpr = 100\%$ and $fpr = 0\%$. The *HAWKwood M.3* benchmark provides 246 data sets (206 real and 40 synthetic) with high overlap, for which ground truth is provided by forestry standards for the real data sets and computed ground truth for the synthetic data. Table 5.4 shows that our method accurately computes the contour volume for real as well as synthetic data. It tends to slightly underestimate the volume in both cases, while the average difference for synthetic data is smaller due to a better segmentability for foreground and background.

	$\bar{x}$	s
real	2.2%	1.8%
synthetic	0.62%	0.05%

Table 3. V_c results (difference to ground truth)

6 Conclusion and perspectives

We have presented a novel method for automatic mobile wood pile surveying. The approach uses multiple view stereo techniques in combination with an existing 3D object detection approach to obtain the 3D locations of the detected wood log faces and derive the contour volume and solid wood volume. Extensive experiments have been performed on 246 data sets for three different benchmarks. We have proved our method to be suitable for real world applications with an average difference of 2.2% for the contour volume and -5.6% for the solid wood volume.

References

[1] Boris Delaunay. "Sur la sphere vide." *Izv. Akad. Nauk SSSR*, Otdelenie Matematicheskii i Estestvennyka Nauk 7.793-800, pp. 1-2, 1934.

[2] Florian Fink: "Foto-optische Erfassung der Dimension von Nadelrundholzabschnitten unter Einsatz digitaler, bildverarbeitender Methoden," *Dissertation Universittsbibliothek Freiburg*, 2004.

[3] Enrico Gutzeit and Jörg Voskamp. "Automatic segmentation of wood logs by combining detection and segmentation." *Advances in Visual Computing*, Springer Berlin Heidelberg, pp. 252-261, 2012.

[4] Christopher Herbon, Klaus Tönnies, and Bernd Stock. "Detection and segmentation of clustered objects by using iterative classification, segmentation, and Gaussian mixture models and application to wood log detection." *Pattern Recognition*, Springer International Publishing, pp. 354-364, 2014.

[5] Christopher Herbon, Benjamin Otte, Klaus Tönnies, and Bernd Stock. "Detection of Clustered Objects in Sparse Point Clouds Through 2D Classification and Quadric Filtering." *Pattern Recognition*, Springer International Publishing, pp. 535-546, 2014.

[6] Christopher Herbon. "The HAWKwood Database." `http://arxiv.org/abs/1410.4393` *arXiv preprint arXiv:1410.4393*, 2014.

[7] Horst Kramer, Alparslan Akça. "Leitfaden zur Waldmesslehre." Sauerländer, 2008.

[8] Adriano Moreira and Maribel Yasmina Santos. "Concave hull: A k-nearest neighbours approach for the computation of the region occupied by a set of points." *GRAPP (GM/R)*, pp.61-68, 2007.

[9] Pierre Moulon, Pascal Monasse, and Renaud Marlet. "Adaptive structure from motion with a contrario model estimation." *Computer Vision - ACCV 2012*, Springer Berlin Heidelberg, pp. 257-270, 2013.

[10] Paul Viola and Michael Jones. "Rapid object detection using a boosted cascade of simple features." *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition CVPR 2001*, Vol. 1, IEEE, 2001.