FILLED
Video data based fill level detection of agricultural bulk freight

Fabian Graefe   Walter Schumacher
Institut für Regelungstechnik
Technische Universität in Braunschweig

Raul Queiroz Feitosa   Diogo Menezes Duarte
Department of Electrical Engineering
Catholic University of Rio de Janeiro

Keywords: machine vision, fill level detection, stereo vision, overload process, automatic overloading.

Abstract: For automation of a continuous overloading process between two vehicles in motion, two information are essential. On the one hand there is the relative position between the vehicles to be known. On the other hand the loading point within the load space of the transport car has to be determined. Often a non optimal usage of the transport capacity is obtained without moving the overload swivel. In order to optimize the filling process by moving load point the distribution of the freight with in the load space has to be measured during the overload process. In this article the Institut für Regelungstechnik of the Technische Universität in Braunschweig introduces the system FILLED for video data based fill level detection of agricultural bulk freight such as chaffed corn or grass.

1 INTRODUCTION

Nowadays more and more powerful machines become available in agriculture. Therefore speeds in harvesting and harvesting tonnage increase. Keeping the modern fast continuous overloading machines in a proper relative position during the whole harvest imposes a significant stress upon the drivers. In order to disburden the drivers of a harvesting combination of a forage harvester and a transport unit a cooperation of the Institut für Landmaschinen und Fluidtechnik and the Institut für Regelungstechnik of the Technische Universität in Braunschweig deals with a system for automation of this overload process. In this context the Institut für Regelungstechnik is developing the system FILLED for video based fill level detection. Most of the fill level sensor systems available are bound to the storage device. Any modification of the transport car requires an adaptation of the sensor systems that imply in additional costs. Usually agricultural bulk freight such as chaffed corn or grass does not fill the load space utterly without variation of the load point. For automatically adjusting the load point it is necessary to know the distribution of the freight within the load space. Therefore any type of fill level sensor which provides just one local fill level information is of less use. Moreover sensors using ultra sonic methods are unsuitable due to their sensitivity for air motion and temperature. Radar based sensors systems usually encounter problems detecting organic material. Tests with laser scanners failed due to the cloud of spraying freight particles in the air. This work proposes a video based fill sensor system that has the potential to overcome all the mentioned troubles. Cameras provide the possibility to obtain a huge amount of 2D data in a single shot. The measurement range can be simply adjusted by tuning of focal length. By using a second camera 3D information can be computed using well known stereo techniques. Moreover the fast increasing processing power in addition with powerful mathematically methods of image processing makes video data based sensor systems become

Figure 1: Forage Harvester

Graefe F., Schumacher W., Queiroz Feitosa R. and Menezes Duarte D. (2005).
FILLED - Video data based fill level detection of agricultural bulk freight.
In *Proceedings of the Second International Conference on Informatics in Control, Automation and Robotics - Robotics and Automation*, pages 439-442
DOI: 10.5220/0001182704390442

more and more feasible. This article presents the key features of an automatic video based fill level measurement system (FILLED) that has been developed by the Institut für Regelungstechnik of the Technische Universität in Braunschweig. The system computes a 3D-model of the freight content using stereo vision. In the following the way FILLED derives the fill level from the image data will be explained. In section 2.1 we will provide a close look at the methods that are used for image capturing. Section 2.2 will introduce the means of image segmentation. The usage of stereo vision will be explained in section 2.3. In the end the strategy for improving the update rate will be pointed out in section 2.4. Section 3 will provide concrete information about the prototype of FILLED.

2 SYSTEM DESCRIPTION

The system FILLED measures the fill level by means of a stereo vision analysis. A 3D-model of the freight pile surface and the freig car's upper rims is calculated. To be more precise the local height of the freight pile is detected at a sufficiently high amount of points all over load area. The car's rims are used for determination of the freight car's coordinate system. The points measured are related to the car's coordinate system. FILLED processes data in three steps:

- Image capturing
- Segmentation
- Stereo vision

2.1 Image Capturing

The object to be measured is in motion during measurement process. This requires that the images of each stereo pair are captured simultaneously. A critical aspect of most stereo systems is the correspondence problem. It consists of finding the projections of points in 3D space on both cameras as accurately as possible. To simplify this task the cameras are geometrically calibrated by using some appropriated off-line procedure during initialization (Online procedures are thinkable (P.H.S. Torr, 2004)) Before coming into operation a calibration cycle consisting of:

- comparison of both image histograms of one shot and
- correction of the exposure time and gain are

attained. This procedure is performed once during initialization. Further the cameras must be calibrated geometrically by using some appropriated off-line procedure. It is advisable to repeat this calibration procedure from time to time.

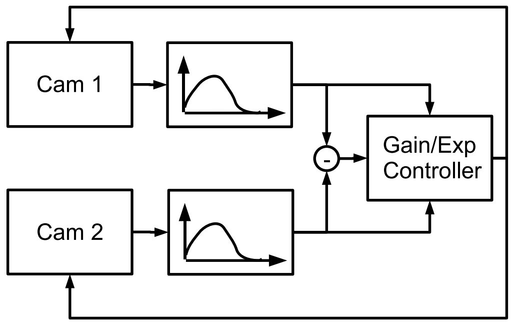

Figure 2: hardware control structure

2.2 Segmentation

In this step the pixels representing the heap on the image are identified. The segmentation is executed in two steps. First the regions showing the freight car are separated from the rest of the image. The second step of segmentation locates the image regions containing the freight heap. With the car's rims painted red the borders of the car can be extracted with a color segmentation procedure applied (Will Schroeder, 2003). In order to reduce the amount of points markes for De Hough Analysis the remaining datasets are processed with a canny (Canny, 1986) edge detection procedure. Using the De Hough transformation on the resulting bitmap the mapped rims are converted into lines. This way the limits, the coordinatesystem and thus the interior of the car is defined(Jon Orwant, 2000).

The second step locates the heap on the freight car. For segregating the heap from the car's interior the local frequencies of the textures are analyzed. Therefore a two dimensional Fourier transformation of a kernel around a designated pixel $P(x, y)$ is computed. If $F(u, v)$ is a Fourier coefficient of the transformed image sample, all points with mean Fourier coefficient $\overline{F}^*$ of less than the threshold T_F within a certain frequency range $\Delta\mathbb{F}$ are suppressed.

$$\Delta\mathbb{F} : \{F(u, v) \in r_{F,min} < r_F < r_{F,max}\} \tag{1}$$

$$r_F = \sqrt{(u^2 + v^2)} \tag{2}$$

$$P(x, y) = \begin{cases} 1 & \overline{F}^* > T_F \\ 0 & \text{else} \end{cases} \tag{3}$$

2.3 Stereo Vision

For setting up a 3D model of the payload surface a stereo vision technique is applied to the image region covering the freight heap. The stereo vision procedure can be divided into three steps image pre-processing, point matching and reconstruction of the points in space.

2.3.1 Pre-Processing

The RGB image is initially converted into gray-scale. After that an algorithm is applied to compensate non-uniform illumination and to provide a roughly uniform contrast all over the image. Figure 3 shows shadows projected from the freight cars walls in the regions close to the corners of the freight car. It can

Figure 3: freight car being loaded

also be seen that the top of the freight heap is brightly illuminated. The procedure to compensate for such effects bases on the fact that the texture information of the heap is concentrated in higher frequencies, while natural illumination variation is represented by low frequencies. By applying a linear Gaussian high pass filter the effects of varying illumination are eliminated.

$$f_{hp} = h_{ghp} * f_{in} \tag{4}$$

Where f_{hp}, h_{ghp},f_{in} and $*$ denote the output of the high-pass filter, the Gaussian kernel, the input image and the convolution operator. Now the local contrast f_{lc} is estimated by applying a linear low-pass filter to the absolute value of f_{hp}, according to the following equation:

$$f_{lc} = h_{glp} * |f_{hp}| \tag{5}$$

The local contrast is normalized by dividing every element of f_{hp} by the corresponding value of f_{lc}. So the final normalized image f_n is obtained by the last pre-processing operation given by:

$$f_n = \begin{cases} \frac{h_{hp}}{f_{lc}} & f_{lc} > 0 \\ 0 & f_{lc} = 0 \end{cases} \tag{6}$$

2.3.2 Point-Matching

In the second step pairs of corresponding points in the left and right camera images are located. These points will be used in the next step to reconstruct the 3D surface model. Points are first selected on the left image in such a way that they are as uniform distributed as possible over the heap. These points can be stored in a $3 \times N$ matrix $\mathbf{P_L}$, containing in each column the homogeneous coordinates $[x_{Li}, y_{Li}, 1]^T$ of the N points selected on the left. With the cameras mounted in such a way that the baseline is much shorter than the fixation point on the scene conventional correlation performs well. So let w_{Li} (resp. w_{Rj}) denote the vector obtained by scanning the window of size $(2n+1) \times (2n+1)$ centered at $P_{Li} = (x_{Li}, y_{Li})$ (resp. $(P_{Rj} = (x_{Rj}, y_{Rj})$) on the left (resp. right) image one row at a time. The similarity between the points at P_{Li} and P_{Rj} is then given by

$$c(P_{Li}, P_{Rj}) = \vec{w}_{Li} \cdot \vec{w}_{Rj} \tag{7}$$

where the symbol "$\cdot$" represents the internal product operator. The best match P^*_{Ri} of P_{Li} on the right image is given by:

$$P^*_{Ri} = \arg\max\left(\vec{w}_{Li} \cdot \vec{w}_{Rj}\right) P_{Rj} \tag{8}$$

To reduce the processing time the search is restrict to an area A_{Li} of size $(2m+1) \times (2m+1)$ with $m > n$ on the right image centered at position (x_{Li}, y_{Li}). Further improvements can be achived exploring the epipolar geometry. As (Emanuele Trucco, 1998) states the search for the matching point P_{Ri} can be restricted to the segment of its epipolar line contained in the area A_i within the right image.
In the third step the object points in 3D space are reconstructed. (Emanuele Trucco, 1998) proposes an extended triangulation that provides an optimal approximated solution for the object points in space. With the results of the geometry analysis the X and Y-axis of the car's coordinate system are identified parallel to the car's rims. The Z-axis is a vector perpendular to the X and Y-axis.

2.4 Performance Aspects

One basic problem dealing with video data analysis is the high amount of data which has to be processed. To keep processing time below an acceptable value the De Hough algorithm and the regions of interest are limited to convenient ranges. This is done by assuming that disparities between images of a single camera in consecutive frames as well between images of both cameras acquired at the same time are very low.

3 PERFORMANCE EVALUATION

FILLED has been tested with image data taken during harvesting process. Figure 3 shows a typical image that was used for testing. The image capture hardware consists of two cameras each connected with a PCI-Framegrabberboard. Two progessive scan cameras (Pacific FAC 9820, IDS), equipped with a $\frac{1}{2}$" CCD

640×480, RGB-color, progressivescan sensor have been implemented. Standard wide angle lenses with a focal length of 4.8 mm were attached. The De Hough space has computed with a resolution of $\Delta\varphi = \frac{\pi}{100}$. The R_{DH} axis has been mapped on to 255 values. Good results for stereovision preprocessing could be obtained using a 40×40 kernel for h_{hp}. The matching point procedure has been computed within a range of 120×120 pixels around the designated point P_{Li}. The application will be implemented in C/C++ running on a standard office PC with a 3GHz Pentium 4 processor using Windows 2000. With the functions of image segmentation that are already implemented on the target system we obtain a effective update rate $F_{upd} \geq$ 1Hz. The turnaround time of the modules of stereo vision still implemented in matlab m-code falls below 40 seconds. If FILLED is implemented completely on the target system we expect an overall update rate of $F_{upd} \geq 0.1$Hz.

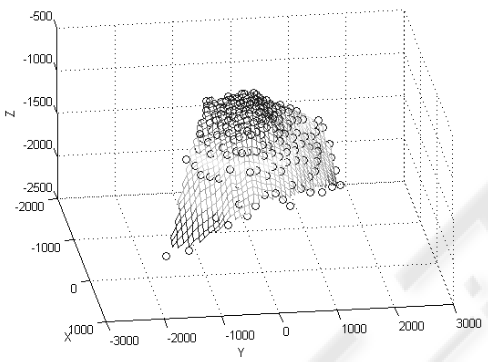

Figure 4: 3D Model of the freight heap

Figure 4 shows the 3D model of the bulk freight heap. This dataset was derived from the car shown in figure 3 as the result of the stereovision analysis. The surface is reconstructed with regard to the camera coordinate system. The axis are scaled in milimeters.

4 CONCLUSION

This work introduces a novel sensoring approach to measure the fill level of agricultural bulk freight. We believe that the general procedure used in the present FILLED prototype can adequately adapted to work with most kinds of bulk freight, and that an assignment of FILLED for many types of transport devices and for different materials would become feasible. A further ongoing research on shape from X methods investigates the possibility of using just a single camera instead of a stereo vision combination. Moreover usage of a calibrated camera system will provide the possibility to measure the overloaded volume. Although it becomes obvious that this could be a practical method of fill level measurement, there are still problems to be solved. With regard to the speed of filling - in our case the harvester processed aproximately 150 tonns per hour - an update rate of 0.1Hz would be hardly enough for controlled loading. However tests evidenced that it is possible to follow the filling process. An optimization of the implementation of the mathematical processes as well as the usage of specific hardware for array mathematic could be an approach to increase the measurement speed. Due to the fact that no reliable online calibration method could be implemented yet, the state of the calibrated optical system must not changed during measurement. This inhibits the usage of autofocus objectives or objectives which provide a variable magnification. Using the inline calibration algorithm mentioned before requires initialization runs from time to time to ensure that the epipolar system data are valid. This means that the overloadprocess has to be interrupted, which is unfavoured for economical reasons. We think that online calibration methods such as the eight point algorithm (P.H.S. Torr, 2004) could be a possible solution for this problem.

ACKNOWLEDGMENT

The authors would like to thank the DFG (German Research Foundation) for their financial support of this project.

REFERENCES

Canny, J. (Nov 1986). *A Computational Approach to Edge Detection*. IEEE Transactions on Pattern Analysis and Machine Intellegence, Vol 8, No 6.

Emanuele Trucco, A. V. (1998). *Introductory Techniques for 3-D Computer Vision*. Prentice Hall (25. Mrz 1998).

Jon Orwant, Jarkko Hietaniemi, J. M. D. (April 2000). *Algorithmen mit Perl*. ISBN 3-89721-141-6, 1. Auflage.

P.H.S. Torr, A. F. (2004). *Invariant Fitting of Two View Geometry or In Defiance of the eight point algorithm*. IEEE Transactions on Pattern Analysis and Machine Intellegence, 26(5), pages 648-651.

Will Schroeder, Ken Martin, B. L. (2003). *VTK Toolkit 3rd Edition An Object-Oriented Approach To 3D Graphics*. ISBN 1-930934-12-2, Kitware.