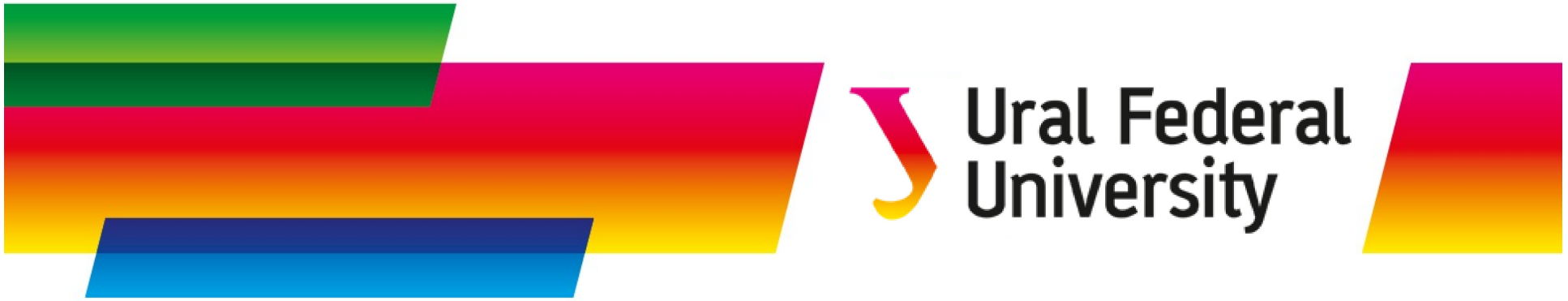


Detection and Dimension of Moving Object by Using Single Camera Applied to the Round Timber Measurement

Yuriy Chiryshev, Anastasiya Atamanova, Artem Kruglov, Svetlana Zavada[1]
[1]Engineering School of Information Technologies, Telecommunications and Control Systems, UrFU

Yuriy Chiryshev
j.v.chiryshev@urfu.ru

12th International Symposium Advances in Artificial Intelligence and Applications (AAIA'17)
Prague, Czech Republic, 3 - 6 September, 2017

Ural Federal University

Preamble

- Logs passing through the conveyor.
- The input data for the measurement algorithm is digitalized video sequence obtained from the camera which is mounted over a conveyor.
- Data on objects of interest is obtained from single view

Visual Sensor

Image Sequences

Ural Federal University

Image Processing Algorithm

The specifics of the task:

- Strict restrictions to the algorithm speedup (real time mode);
- Background dynamic changing (due to the moving parts of the conveyor)
- Flat contrast of the scene;
- Overlap of the objects of interest

The algorithm consists of three parts:

- Images Segmentation And Object Detection
- Determination of direction and velocity of the objects
- Estimation of the 3D objects properties by 2D projections “Shape from Silhouette”

Images Segmentation And Object Detection

- Simple difference between background pixel and the current pixel:

$$|\mu(x,y) - X_t(x,y)| - \sigma(x,y) > p$$

$$\mu = \frac{1}{n}\sum_{t=1}^{n} X_t(x,y) \qquad \sigma^2 = \frac{1}{n}\sum_{t=1}^{n} X_t^2(x,y) - \left(\frac{1}{n}\sum_{t=1}^{n} X_t(x,y)\right)^2$$

- Background updated with the infinite impulse response:

$$\mu_{t+1}(x,y) = (1-\alpha)\mu_t(x,y) + \alpha X_t(x,y)$$

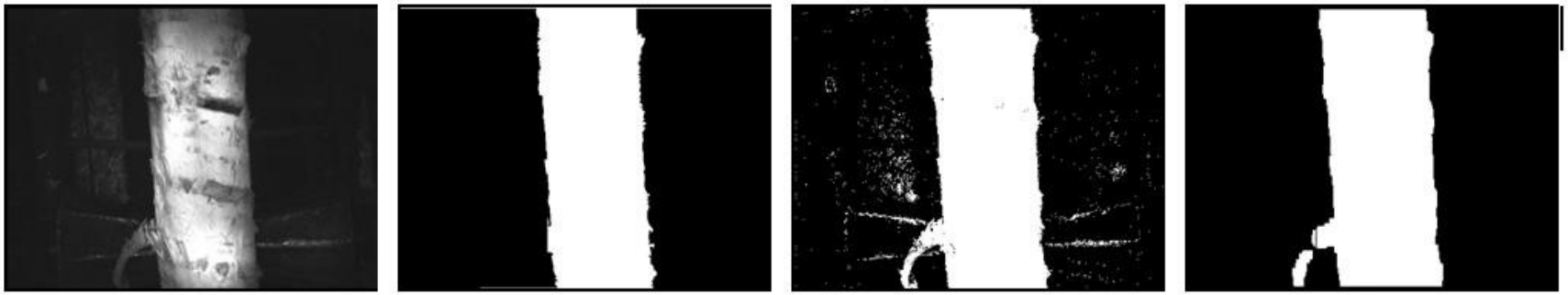
$$\sigma_{t+1}(x,y) = (1-\alpha)\sigma_t(x,y) + \alpha|X_t(x,y) - \mu_t(x,y)|$$

- **What are the optimal threshold *p* and parameter α?**

a) Background b) Background Mean Square Deviation c) Input Frame d) Background Model Subtraction Result (Log Silhouette)

Ural Federal University

Images Segmentation And Object Detection



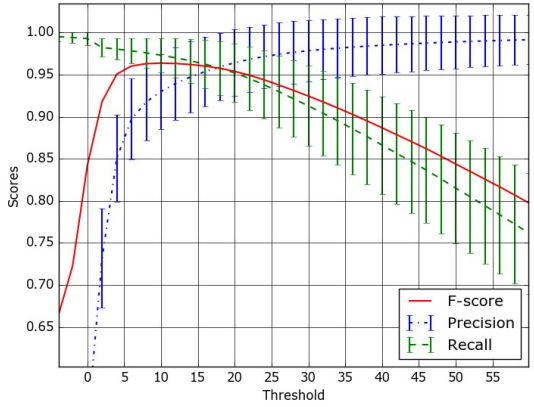
a) Input b) Ground Truth c) Output d) After Noise Filtration (Morphological Filter)

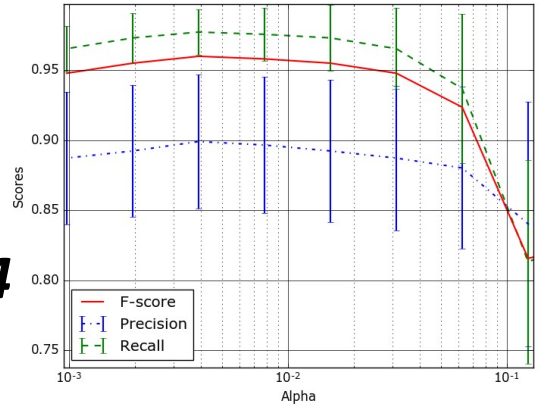
- Metrics:

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F_\beta = (1 + \beta^2)\frac{Precision \cdot Recall}{(\beta^2 + Precision) + Recall}$$

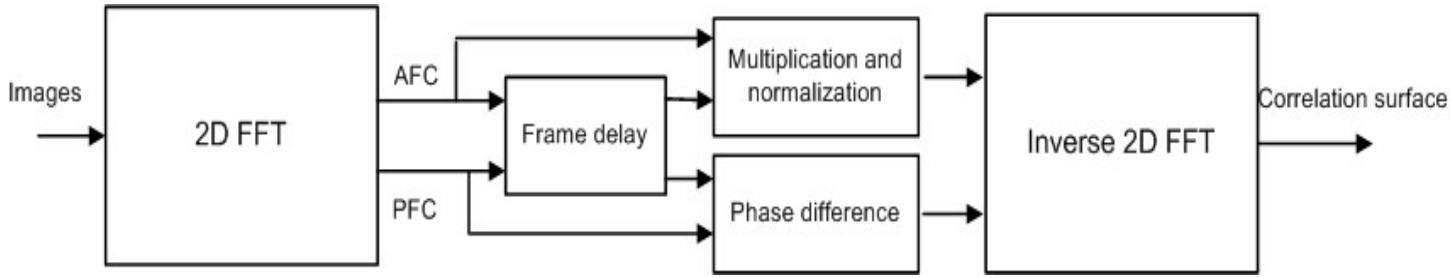




- **Optimal threshold *p=8* and parameter α*=0,004***

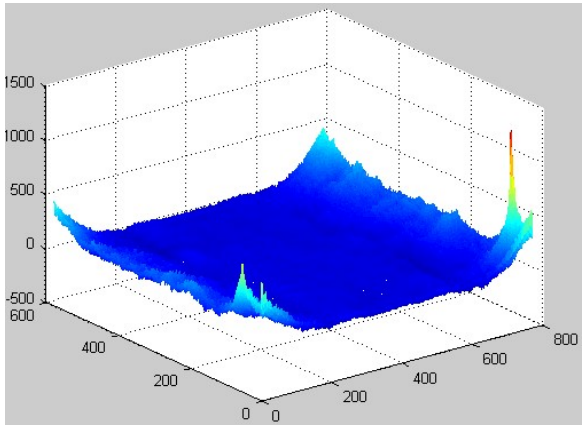
Determination of direction and velocity of the objects

- Phase Correlation Method:



Previous Image

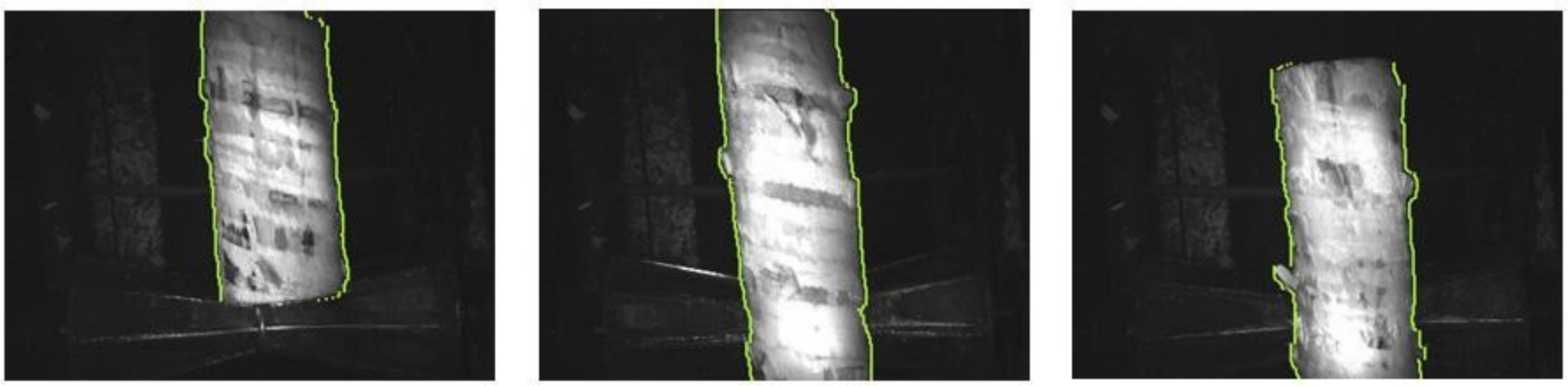
Current Image



Correlation Surfaces

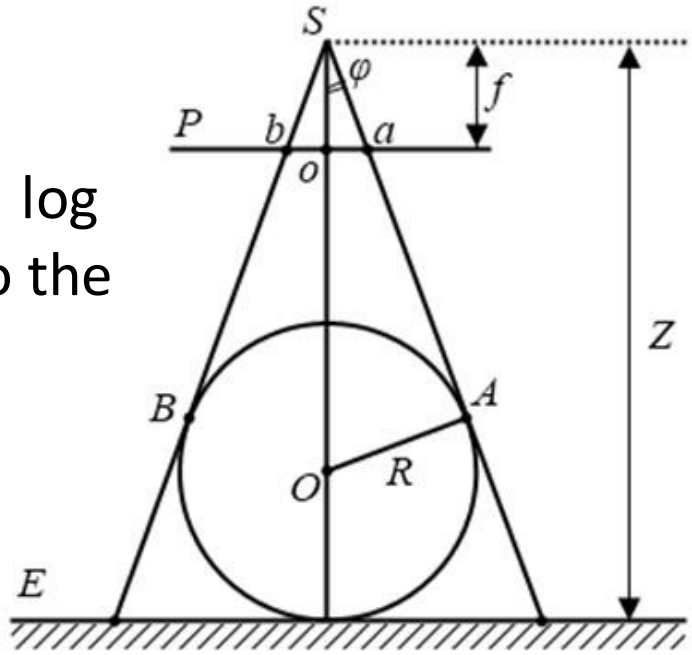
Ural Federal University

Contour extraction and parameters estimation



Video shooting model:

- Log is a generalized cylinder
- Camera is downward directed to observe log in such a way that image plane is parallel to the conveyor plane
- Internal and external parameters for the camera are given



$$R = \frac{Z}{1 + \sqrt{1 + \frac{f^2}{r^2}}} \qquad V = \frac{\pi}{3}\sum_{i=0}^{n}(R_i^2 + R_i r_i + r_i^2) l_i$$

Ural Federal University

Contour extraction and parameters estimation

Problem of the real log boundaries recovery from noisy input data

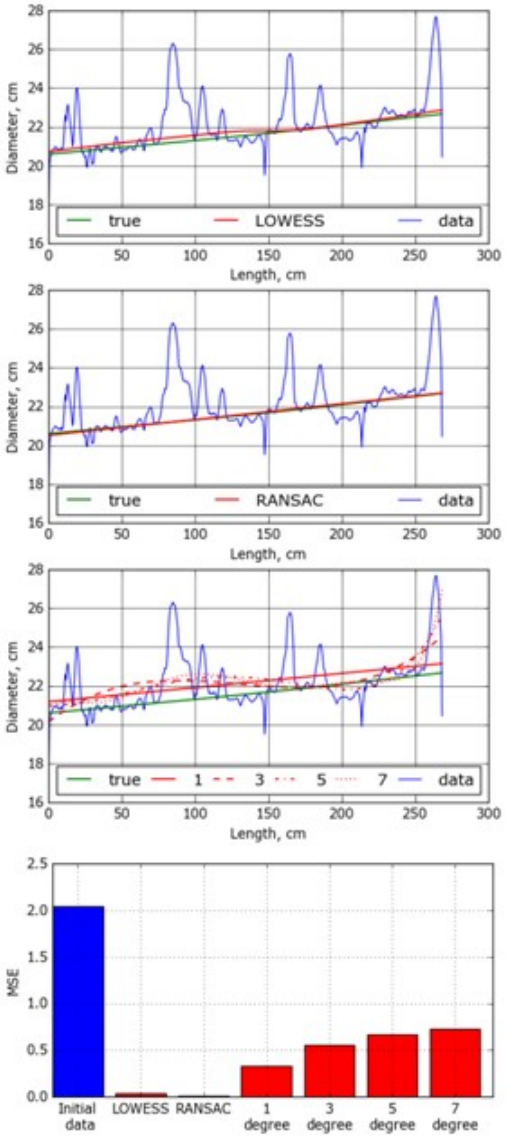
- Regression analysis :

$$\sum_{i=1}^{n}(y^*(x_i) - y(x_i))^2 \to \min_y$$

- Metrics:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i^* - Y_i)^2$$

Method	MSE(σ_{MSE})
Initial data (before smoothing)	1,781(0,153)
LOWESS	0,115 (0,097)
RANSAC	***0,045(0,041)***
Polynomial (1 degree)	0,271 (0,107)
Polynomial (3 degree)	0,395 (0,139)
Polynomial (5 degree)	0,585 (0,065)
Polynomial (7 degree)	0,726 (0,041)



Ural Federal University

Conclusion and Future Work

Results:

- Logs segmentation allow to conclude that the image can be separated into background and foreground regions by using quite simple subtraction methods.

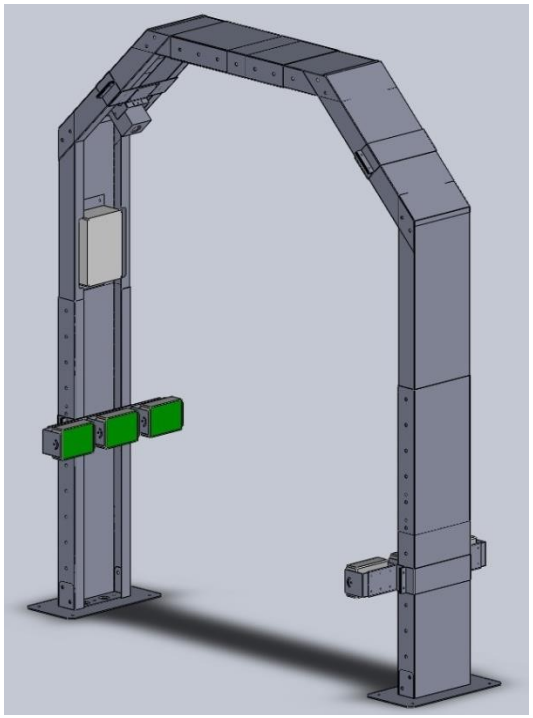
- At pixel level algorithm provides quality of the detection at the rate of 96,9% true positive rate with $2{,}9 \cdot 10^{-2}$ false positive rate;

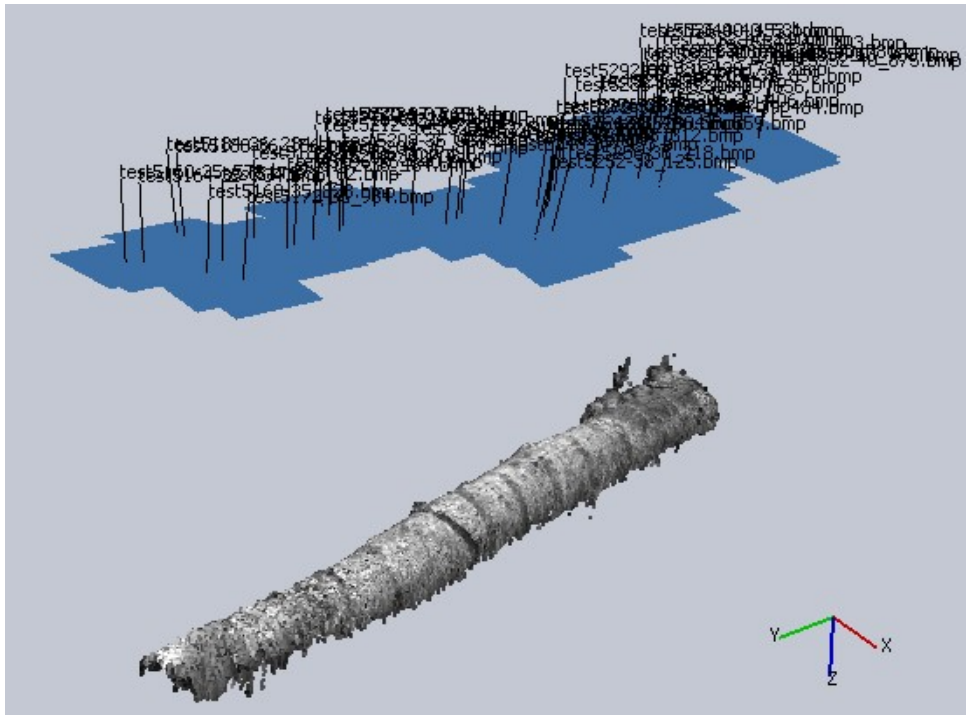
- Log reconstruction experiment show that the RANSAC has the best performance among the observed methods.

- The operation speed of the algorithm provides processing of the video sequence of 384x288 frame size at 25 frames per second (Intel Core i7, 2800 Mhz, 6Gb DDR RAM, GeForce GTS 450.)

Ural Federal University

Conclusion and Future Work





- From “single view” to “stereo camera”
- From “Shape from Silhouette” to “Shape from Motion”
- From “Curve fitting” to “Point Cloud and Parametric Reconstruction of Log Surface”
- Estimate the quality characteristics of logs, such as crook, ovality and buttswell.

Ural Federal University

Thank you for attention

Yuriy Chiryshev
j.v.chiryshev@urfu.ru