8th IEEE International Conference on Industrial Informatics

# INDIN 2010

July 13-16, 2010

Osaka University Nakanoshima Center, Osaka, JAPAN

Osaka Convention & Tourism Bureau

# Transport by Throwing - a bio-inspired Approach

Martin Pongratz*, Friederich Kupzog†, Heinz Frank‡ and Dennis Barteit§
*†Vienna University of Technology
Institute of Computer Technology, Gusshausstrasse 27-29, A-1040 Wien
*pongratz@ict.tuwien.ac.at, †kupzog@ict.tuwien.ac.at
‡§Reinhold Wuerth University
Department of Electrical Engineering, Daimlerstr. 35, D-74653 Kuenzelsau
‡frank@hs-heilbronn.de, §barteit@hs-heilbronn.de

***Abstract*—In modern industrial production fast and easily reconfigurable transportation systems are necessary. A viable bio-inspired approach to this is throwing and catching of transportation goods. In order to catch a thrown object the catching device has to be moved to the right position on time. This requires a fast and accurate acquisition of flight position and a prediction system for the interception point. The main topic of this paper is the development and comparison of two prediction models for the flight trajectory of a thrown tennis ball. The position acquisition, that is the base for the prediction, is based on a binocular vision system similar to two-eyed humans. The impact of the vision systems frame rate on the error of the prediction is reviewed as well. Future prediction is planned to be done based on a bio-inspired approach using a small set of reference throws.**

## I. Introduction

Throwing and catching is a fast and flexible approach to transportation . Accurate throwing and very accurate catching is required to enable this approach. External influences on a thrown objects trajectory like wind or a changing objects orientation boost demands on the catching instance. An accurate object tracking system is required [1] [2]. Findings on humans with normal and weak stereopsis emphasize the need for high quality tracking [3]. This tracking information has to be used for predicting the interception position. A model of the flight has to be developed and fitted into the measured positions of the object. Based on this model the interception position can be predicted. If the prediction is more accurate than the coverage of the catching device the object will be caught.
Similar to most work regarding the topic of catching [1] [4] also this approach is dealing with a tennis ball as thrown object. Approaches based on a monocular vision system [4] and based on binocular vision system [1] [5] have already been done. In this context usage of binocular vision systems equal a bio-inspired approach as most predators in nature feature two eyes. Prior research based on binocular vision systems differs in the distances the object is actually thrown. Generally this distance has been orders of magnitudes smaller than the throwing distance of 3 m used for the experiments presented. In contrast to small scale catching [1] [5] no gripper is used to verify the quality of the prediction system. An impact position verification system (compare [4] and [6]) based on a touch-kit is used to continuously meter the quality of the prediction. This systems detects the position of the tennis ball in the interception plane enabling better evaluation of the prediction quality than the binary result of a successful/unsuccessful catch.

## II. Application

Individualization of products has put focus on flexible production systems and their reconfiguration. In order to save cost the time of reconfiguration has to be minimized. One main aspect of this procedure is the reconfiguration of the transportation system. Conveyors have to be disassembled and reassembled which takes a long time. A very fast reconfigurable transport approach is to throw and catch objects [7]. Reconfiguration of such a system is limited to assigning a new target to the throwing instance and a new object origin to the catching instance. No mechanical reconfiguration is necessary. An individual sequence of production steps for each part is possible which also enables dynamic load balancing in the production facilities in case of an erroneous machine.

## III. Modeling the Flight

Flight properties of a thrown object depend mainly on the shape and surface of the object. A tennis ball is used for the work presented. The symmetric properties of the ball simplify throwing and modeling the flight of the ball. Effects of slow rotation of the object, in case of a ball also called spin, are minor for highly symmetrical objects like a ball. In contrast aerodynamic effects on rotating non-symmetrical objects are influencing the flight to a large degree. In case of high speed spin of highly symmetrical objects the Magnus effect has to be considered [8]. Neglecting all forces influencing the flight other than gravity and drag, the flight of a ball can be described by

$$\vec{v}(t+\Delta t) = \vec{v}(t) + \vec{a}(t) * \Delta t$$

$$\vec{a}(t) = -\frac{\vec{v}(t)}{|\vec{v}(t)|} * k * |\vec{v}(t)|^2$$

$$k = \frac{\rho * c_W * A}{2}$$

Where $\vec{v}(t)$ is the velocity at the instant $t$, $\vec{a}(t)$ equals the acceleration at the instant $t$, $\Delta t$ is the timely granularity of the calculation and $k$ is the aerodynamic factor which is calculated

978-1-4244-7300-7/10/$26.00 ©2010 IEEE

based on the air density $\rho$, the air drag coefficient $c_W$ and the cross section surface of the object $A$. Calculation of the flight trajectory can only be done iteratively based on the initial velocity $v(t = 0)$ as the influence of the air drag is nonlinear. The factor $k$ depends on the object and air density, varies from tennis ball to tennis ball and also depends on the spin of the ball [8]. In order to predict the objects trajectory based on measured positions the requirement for iterative calculation is a big challenge. Different sets of initial velocity $v(t = 0)$ and $k$ need to be tested and the best combination can be used to predict the future flight. Real-time requirements and accurateness within pre-specified bounds can not be obliged. These essential limitations can be avoided by using other models than the on presented above.

### A. Polynomial Model

The most simple approach to fit functions to the measured positions of the object is using polynomial functions. Subsequently predicting the future trajectory is enabled based on the fitted functions. The order $n$ of the polynomial function can be derived from the goodness of the fit. A higher order provides better goodness but the sensitivity to measurement errors of the acquired positions demands a lower order. For this reason a suitable compromise has to be found. The function describing the position $p$ of the object in the $i$-th $(x, y, z)$ spatial direction is estimated as following

$$p_i = p_0 + p_1 * t + p_2 * t^2 + ...p_n * t^n$$

Weighting the importance of the individual measured positions enables to incorporate accuracy variation depending on the objects distance to the camera set.

### B. Spatial separated physical Model

A more refined model of the flight is a spatial separated simplification of the model presented in the introduction of this chapter. Separating the movement in the spatial directions introduces an error due to the nonlinear property of the air drag [8]. This error depends on the ratio of the velocity in the spatial directions. If the movement mainly occurs in one direction the error introduced is negligible. The movement in each direction, according to this model, can be described by the differential equations

$$a_x = \dot{v_x} = -k * v_x^2$$

$$a_y = \dot{v_y} = -k * v_y^2 + g$$

$$a_z = \dot{v_z} = -k * v_z^2$$

if the $y$-direction is aligned with the direction of gravity. Symbols used are acceleration $a$, velocity $v$, aerodynamic factor $k$ and gravity $g$. Solving these equations for the positions along the spatial directions results in

$$x = x_0 + \frac{1}{k} * ln\,(1 + k * t * v_{x,0})$$

$$y = y_0 + ln(\frac{cosh(\sqrt{g * k} * (t - t_0))}{cosh(\sqrt{g * k} * t_0)})$$

$$z = z_0 + \frac{1}{k} * ln\,(1 + k * t * v_{z,0})$$

For evaluation of both models the functions are fit to the measured positions of the tennis ball in the early flight phase. Linear Least Squares are used for fitting the polynomial model while Nonlinear Least Squares are used for fitting the spatial separated model to the data. Subsequently for both models presented above the the equation

$$z(t) = -32.2$$

is solved for $t$. This equals calculating the time of the impact of the tennis ball, that has a diameter of $64.4$ mm, on a plane in the $z = 0$ plane. The position of the ball in x- and y-directions at that instant is the predicted impact position.

## IV. Experimental Setup

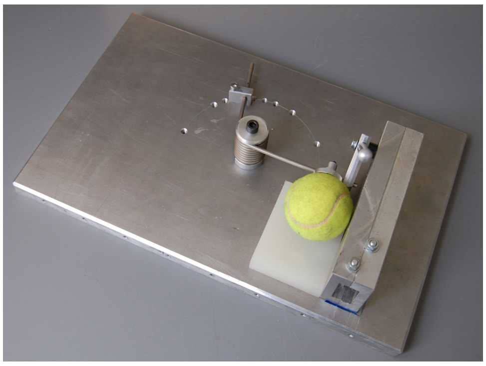

Fig. 1. Simple throwing device based on a leg spring

TABLE I
DST TOUCH KIT PROPERTIES

| | |
|---|---|
| Input Method | Finger and stylus input |
| Accuracy | 1.0% |
| Active Area | 727.15 mm × 408.05 mm |
| Resolution (h x v) | 16k ×16 k (maximum resolution) |
| Response Time | 20 ms for tap input |
| Minimum Touch Impact | 50 mN · s |
| Glass Thickness | 2.2 mm (±0.2 mm) |

The setup used in this work consists of a throwing device, an impact position verification system and the binocular vision system with a PC workstation. Based on a pre-streched leg spring the throwing device (Figure 1) accelerates the tennis ball to a velocity of roughly 10 m/s. The distance between the initial position of the tennis ball and the plane of the position verification system is 3 m. This results in an flight time of $\approx$ 300 ms. The throwing device is mounted onto a table with an inclination of 7° from the horizontal. A ball is thrown towards the plane where the impact position verification system is mounted. This systems consists of a Dispersive Signal Technology (DST) touch-kit. This touch-kit is used to detect the tennis balls position within a plane. Main properties of the DST touch kit are presented in Table I.

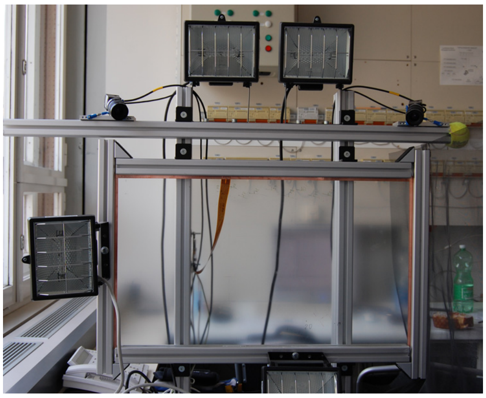

Fig. 2. DST touch-kit mounted into aluminum profiles

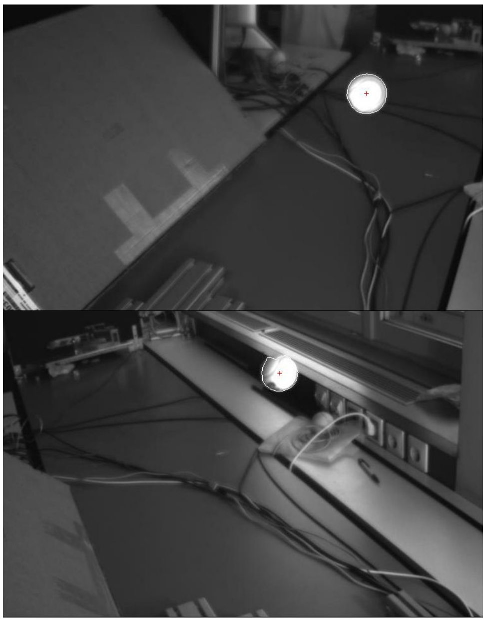

Fig. 3. Visual field and ball recognition in original image (top: right, bottom: left image

Figure 2 shows the DST touch-kit and the mounting in aluminum profiles. The interpretation of the DST touch-kit is done via mapping of the touch-kit to a screen with a resolution of $1600 \times 1200$ pixel. Calibration of the touch kit and the stereo vision system are done concurrent in one process. The calibration is based on 67 different images of the calibration sheet. For ten of those 67 calibration sheets the relation to the DST touch kit is known and this information is used to extract the position of the cameras to the touch kit. The vision system consists of two IDS Eye-1220-C gray scale cameras. Cameras are installed in a convergent setup and their visual field is ranging from the origin of the ball trajectory (throwing device) to approximately 50 cm from the impact plane (compare Figure 3 and Figure 2).

Main properties of the used cameras are shown in Table II. Both cameras are triggered synchronous by hardware via a microcontroller. Additional light is provided by four 500

TABLE II
$\mu$EYE-1220-M-GL PROPERTIES

| | |
|---|---|
| Interface | USB 2.0 |
| Resolution | $752 \times 480$ pixel |
| Sensor size | $1/3$ " |
| Maximum frame rate | 87 fps |
| Exposure time | 415 $\mu$s |
| Focal length | 6 mm |

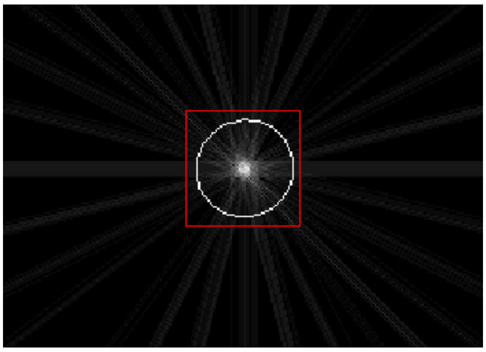

Fig. 4. Sample Hough transformation accumulator

W halogen floodlights (Figure 2). Video data is saved on the workstation and analyzed via Matlab and the AVI read interface dx_avi [9]. The ball is segmented via a background subtraction and the center of the ball in the image is found via a modified hough transformation [10]. Laplacian of Gaussian edge filtering is used to extract the edge image. A window of $7 \times 7$ pixels (enlarged by 3 pixels around each central point of the edge) is used to determine the potential radial direction of the arc. Both extreme ends of the line in the window are used to estimate the tangential direction in the central point of the window. Voting for the resulting radial lines of all edge points in the accumulator space enables extraction of the tennis balls center. A sample accumulator result is shown in Figure 4. The area inside the red square in Figure 4 is shown Figure 5 in detail. Also the extracted edge line from Figure 4 is drawn

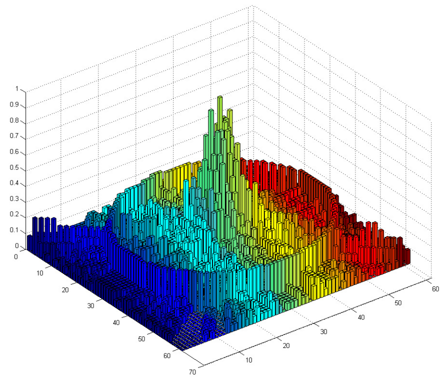

Fig. 5. Histogram of Hough transformation accumulator

into the histogram around the knoll. In the center of the knoll multiple spikes are visible. Filtering the accumulator with a filter of the size $7 \times 7$ pixels results in an individual maxima in the center of the knoll which is considered as the center of the tennis ball in the image. Based on this information the position of the ball in space is calculated through stereo triangulation based on the parameters acquired during the calibration using the stereo camera calibration toolbox [11]. In order to increase accuracy, stereo triangulation considers the lens distortion which is known from the camera calibration process, too.

## V. Prediction Results and Comparison

The impact position is predicted in a row of experiments for 20 throws using both models presented in III-A and III-B with a camera frame rate of 60 fps. The polynomial model is used up to the second order ($n = 2$) in order to give a good compromise between stability and sensitivity. The prediction error is presented in Figure 6 and Figure 7. Both components of the deviation from the impact position are shown. Comparing Figure 6 and Figure 7 shows that

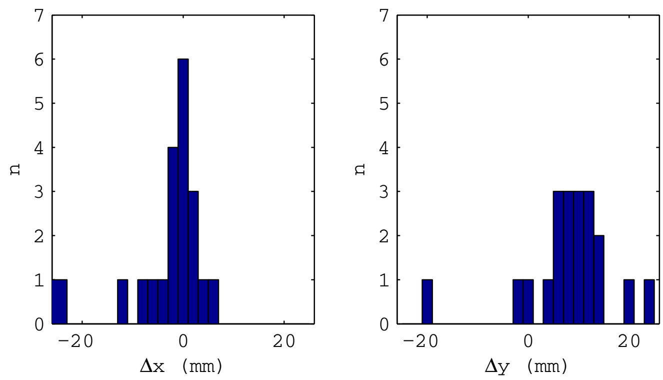


Fig. 6. Histogram of prediction errors based on the polynomial model (horizontal: $\Delta x$, vertical: $\Delta y$)

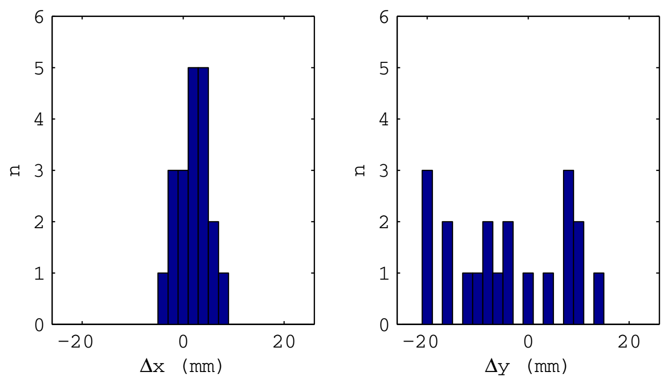


Fig. 7. Histogram of prediction errors based on the physical separated model (horizontal: $\Delta x$, vertical: $\Delta y$)

especially the vertical prediction of both models differs a lot. The polynomial models average deviation is 8.2 mm while the physical models average deviation is $-3.9$ mm. The corresponding numbers in horizontal direction are $-2.8$ mm and 2.0 mm for both models.

The histograms of the overall prediction error (distance between the predicted interception point and the actual impact position of the tennis ball on the plane $\Delta r$) is shown in Figure 8. The average deviation for both models presented is $\approx 10$ mm for the polynomial and $\approx 8$ mm for the physical model. Using Rayleigh distribution to model this errors deviation the prediction accuracy can be metered by the distance within 99.5 % of the throws are predicted. These numbers are 33.5 mm and 28.1 mm for both models. With other words: out of 200 throws in one case the prediction deviation is greater than 33.5/28.1 mm.

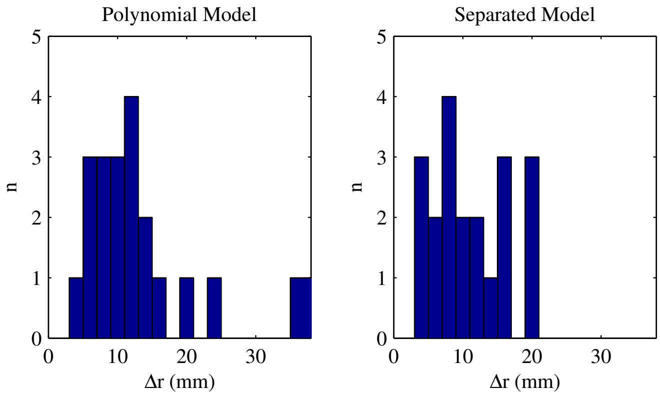


Fig. 8. Comparison of overall prediction errors

## VI. Frame Rate and Resolution Scaling

Besides the camera resolution the frame rate of the vision setup is one main parameter of the prediction accuracy. Higher frame rate and higher resolution equals higher cost for the vision setup. Also the processing requirements rise with the resulting increased data rate. Besides the standard resolution and frame rate used for the prediction results presented in the previous section the cameras allow to downscale the resolution from $752 \times 480$ to $376 \times 240$. This also reduces the bandwidth on the camera interface which is limiting the camera from operating at higher frame rates in general. As a result the achievable frame rates in the reduced resolution mode can be doubled.

Comparison of the 99.5 % prediction radii (Rayleigh distribution, compare last paragraph of previous section) for the three modes reduced resolution/standard frame rate, full resolution/standard frame rate and reduced resolution/doubled framerate is done in Table III for both models presented. While both models prediction increases as the resolution

TABLE III
Scaling Analysis

| | $376 \times 240$, 60 fps | $752 \times 480$, 60 fps | $376 \times 240$, 120 fps |
|---|---|---|---|
| | $\Delta r$ (mm) | $\Delta r$ (mm) | $\Delta r$ (mm) |
| pol. | 40.0 | 33.5 | 36.2 |
| phy. | 30.0 | 28.0 | 24.9 |

increases the combination of a increased frame rate with a reduced resolution shows different results. In case of the polynomial model the prediction accuracy decreases while it improves for the polynomial model. This behavior can be explained by the nature of the two models. While the higher number of measured positions, caused by the higher frame rate, offers measurement error rejection for the physical model the polynomial model is not able to follow the real trajectory of the object due to its missing relation to the physics of the flight.

## VII. Conclusion

Both models presented are viable for the task of predicting the impact position of tennis ball on a plane. The physical model shows a higher accuracy over all reviewed frame rates than the polynomial model. This behavior was expected. When taking a close look at the horizontal deviations presented in Figures 6 and 7 it is interesting to note that the physical model has a higher bias error than the polynomial model when suppressing the outliner in the polynomial model. A possible reason for this behavior lies in neglecting the spin of the tennis ball in the physics based model. Analyzing the video data a spin of $\approx 1000 \text{ min}^{-1}$ occurs. The influence of spin in this magnitude can not be neglected at such low throwing velocities (compare [8]).

The temporal development of the prediction during the flight has been left out of scope. In order to minimize the amount of energy necessary to position the catching device on time the prediction is required to be accurate also with only the information about the first ball positions used to fit the model in.

Additional research regarding higher frame rates with decreased image resolution (due to bandwidth restrictions) seems to be reasonable. Also non equal weighting of the positions calculated might lead to better prediction results as the accuracy of the position detection improves as the ball moves closer to the vision system due to the higher relative resolution.

Prediction so far has not been done at real time. For practical usage of this transportation approach real time requirements arise. Calculation of the prediction has not been optimized for calculation-time so far. Achieving a performance that fulfills the real-time requirements seems to be possible for this brute force approach. Another approach for solving the task of prediction is to use a set of reference throws and their corresponding impact position and to map the actual flight to this library and predict the interception point based on the memory of the prediction system. This scenario-based prediction, that is similar to the way humans and animals evaluate movements, is set as the goal for future research. In this context also analysis of more complex thrown objects has to be mentioned. Considering the rising calculation demands due to the relevance of the objects orientation and their prediction throughout the flight emphasizes the need of alternative and more efficient approaches to solve the task of prediction. Once more, nature can deal as a model.

## References


[1] A. Namiki and M. Ishikawa, "The analysis of high-speed catching with a multifingered robot hand," *Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on*, 2005.

[2] W. H. Massachusetts, W. Hong, and J. jacques E. Slotine, "Experiments in hand-eye coordination using active vision," in *Lecture Notes in Control and Information Sciences*. Springer-Verlag, 1995, pp. 130–139.

[3] M. G. S. G. Mazyn LI, Lenoir M, "The contribution of stereo vision to one-handed catching," *Experimental Brain Research, Volume 157, Number 3, August 2004 , pp. 383-390(8)*, 2004.

[4] D. Barteit, H. Frank, and F. Kupzog, "Accurate prediction of interception positions for catching thrown objects in production systems," *Industrial Informatics, 2008. INDIN 2008. 6th IEEE International Conference on*, 2008.

[5] Y. Imai, A. Namiki, K. Hashimoto, and M. Ishikawa, "Dynamic active catching using a high-speed multifingered hand and a high-speed vision system," *Robotics and Automation, 2004. Proceedings. ICRA '04. 2004 IEEE International Conference on*, 2004.

[6] D. Barteit, H. Frank, M. Pongratz, and F. Kupzog, "Measuring the intersection of a thrown object with a vertical plane," in *Industrial Informatics, 2009. INDIN 2009. 7th IEEE International Conference on*, June 2009, pp. 680–685.

[7] H. Frank, N. Wellerdick-Wojtasik, B. Hagebeuker, G. Novak, and S. Mahlknecht, "Throwing objects a bio-inspired approach for the transportation of parts," *Robotics and Biomimetics, 2006. ROBIO '06. IEEE International Conference on*, 2006.

[8] F. Alam, W. Tio, S. Watkins, A. Subic, J. Naser, P. Jacobs, T. McIntyre, M. Cleary, D. Buttsworth, D. Mee, R. Clements, R. Morgan, and C. Lemckert, "Effects of spin on tennis ball aerodynamics: An experimental and computational study," 2007. [Online]. Available: http://espace.library.uq.edu.au/view/UQ:120801

[9] "A directshow based avi read for matlab." [Online]. Available: http://cs-people.bu.edu/tvashwin/dx_avi/

[10] D. Scaramuzza, S. Pagnottelli, and P. Valigi, "Ball detection and predictive ball following based on a stereoscopic vision system," *Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on*, 2005.

[11] J.-Y. Bouguet, "Camera calibration toolbox for matlab," 02.06.2008. [Online]. Available: http://www.vision.caltech.edu/bouguetj/calib_doc/